A few words about function spaces; what are they good for.

A function space is what its name says it is; a vector space whose elements are functions with domain in some set, values in another. In the most common examples (and we’ll see nothing transcending the absolutely most common examples), all functions take either real or complex values (one usually keeps the real and complex cases separate), the domain is either \( \mathbb{N} \) (or \( \mathbb{N}_0 = \mathbb{N} \cup \{0\} \)) or a subset of \( \mathbb{R}^n \). The set of functions is closed under addition (defined as usual) and scalar multiplication (defined as usual); in other words is a real or complex vector space. Because we are in analysis we need to have a notion of convergence for life to make sense so that these spaces are given a topology which may or may not be metric. We will see exclusively or almost exclusively topologies coming from a norm; in other words metric topologies. Some of the basic definitions and examples are in your Homework #3, but here they are again. I’ll stick to the real case, but most everything we do will also be valid for complex vector spaces (of complex valued functions) with some obvious modifications. We begin at a moderately general level.

Definition 1  A \textit{(real) normed vector space} is a pair \((V, \| \cdot \|)\) where \( V \) is a vector space over \( \mathbb{R} \) and \( \| \cdot \| \) denotes a map from \( V \) to \( \mathbb{R} \), where \( \|x\| \in \mathbb{R} \) denotes the value of the map at \( x \in V \), satisfying the following three properties:

1. \( \|x\| \geq 0 \) for all \( x \in V \), and \( \|x\| = 0 \) if and only if \( x = 0 \).
2. \( \|cx\| = |c| \|x\| \) for all \( c \in \mathbb{R}, x \in V \).
3. \( \|x + y\| \leq \|x\| + \|y\| \) for all \( x, y \in V \).

A map \( x \mapsto \|x\| \) satisfying the three properties above is called a \textit{norm}, so a normed space is a space in which a norm has been defined. Sometimes one uses subscripts to distinguish one norm from another. Many authors like to use the symbol from the space as a subscript for the norm of the space; if the space is denoted by \( V \), the norm is denoted by \( \| \cdot \|_V \). We follow the usual abuse of language saying simply “\( V \) is a normed space” when we mean \( V \) is a vector space in which one has defined a norm \( \| \cdot \| \) such that \((V, \| \cdot \|)\) is a normed space.

Lemma 1  Let \( V \) be a normed vector space. If \( x, y \in V \) define \( d(x, y) = \|x - y\| \).

Then \( d \) is a distance function for \( V \).

Proof. Exercise.

We will always consider a normed space as a metric space with this distance, called the \textit{norm distance}.

Definition 2  A \textit{Banach space} is a complete normed space. In other words, a space in which all Cauchy sequences converge. In even more words, the normed space \( V \) is a Banach space if whenever \( \{x_n\} \) is a sequence of points of \( V \) such that for every \( \epsilon > 0 \) there is \( N \) such that \( \|x_n - x_m\| < \epsilon \) whenever \( n, m \geq N \), then there exists \( x \in V \) such that \( \lim_{n \to \infty} \|x_n - x\| = 0 \).
1.1 Some interesting but a bit specialized facts.

I’ll probably skip this in class. Suppose $V$ is a vector space and $\| \cdot \|_1, \| \cdot \|_2$ are both norms in $V$. Both define metrics, both define the concept of an open set. Because open sets are always paired with closed sets (a set is closed if and only if its complement is open), and because closed sets can be exclusively defined in terms of the concept of convergence of sequences (a set is closed if and only whenever a sequence of points in the set converges to a point in the space, that point is in the set), it is fairly easy, trivial, no sweat, to prove that these norms define the same family of open sets if and only if they both have the same family of convergent sequences: convergence in one norm is equivalent to convergence in the other. Because we are in what is really a premium metric space environment, namely a metric space, we can go one step further. Because a sequence $\{x_n\}$ converges to $x$ if and only if $\{x_n - x\}$ converges to 0; it is enough that if a sequence converges to 0 in one norm, it does so in the other norm. The following theorem summarizes all this and a bit more.

**Theorem 2** Let $\| \cdot \|_1, \| \cdot \|_2$ be two norms in the vector space $V$. The following conditions are equivalent.

1. A subset $U$ of $V$ is open with respect to $\| \cdot \|_1$ if and only if it is open with respect to $\| \cdot \|_2$. This means: If $U \subset V$, then for every $x \in U$ there is $r > 0$ such that $B_1(x, r) = \{y \in V : \|y - x\|_1 < r\} \subset U$ if and only if for every $x \in U$ there is $\rho > 0$ such that $B_2(x, \rho) = \{y \in V : \|y - x\|_2 < \rho\} \subset U$.

2. For every sequence $\{x_n\}$ in $V$ the sequence $\{x_n\}$ converges with respect to the norm metric defined by $\| \cdot \|_1$ if and only if it converges with respect to the norm metric defined by $\| \cdot \|_2$. In this case, the limits are the same.

3. If $\{x_n\}$ is a sequence in $V$, then $\lim_{n \to \infty} \|x_n\|_1 = 0$ if and only if $\lim_{n \to \infty} \|x_n\|_2 = 0$.

4. There exist constants $a, b, b > a > 0$ such that

$$a \|x\|_1 \leq \|x\|_2 \leq b \|x\|_1$$

for all $x \in V$

A pair of norms satisfying any and thus all of the conditions of the theorem are said to be equivalent. The condition most commonly used to define equivalence is (1). Notice that it is an equivalence condition (that is obvious also from any of the other conditions), (1) can also be written in the form

$$\frac{1}{b} \|x\|_2 \leq \|x\|_1 \leq \frac{1}{a} \|x\|_2$$

so the condition is symmetric in the two norms. The proof of the theorem will be mostly left as an exercise. A lot of it is immediate, and/or done in the remarks preceding the statement. The only part that might give some trouble is that the third condition implies the fourth. That is done best by contradiction. Assume 3 holds, but 4 is false; say that there is no constant $b > 0$ such that $\|x\|_2 \leq b \|x\|_1$ for all $x \in V$. Then for every $b > 0$ there is some $x_b \in V$ such that $\|x_b\|_2 > b \|x_b\|_1$. Specialize to $b = n \in \mathbb{N}$; for every $n \in \mathbb{N}$ there is $x_n \in V$ such that $\|x_n\|_2 > n \|x_n\|_1$. Let

$$y_n = \frac{1}{\|x_n\|_2} x_n, n \in \mathbb{N}.$$
Then $$\|y_n\|_1 = \frac{\|x_n\|_1}{\|x_n\|_2} < \frac{1}{n} \rightarrow 0,$$
so that $$y_n \rightarrow 0$$ with respect to $$\| \cdot \|_1$$. But $$\|y_n\|_2 = 1$$ for all $$n$$, so $$y_n \not\rightarrow 0$$ with respect to $$\| \cdot \|_2$$, contradicting our assumption. Similarly one sees there is $$c > 0$$ such that $$\|x\|_2 \leq c\|x\|_1$$ for all $$x \in V$$; taking $$a = 1/c$$ we are done.

Referring to the examples of norms given in Homework 3 for $$\mathbb{R}^n$$, it is very easy to see they are equivalent by verifying that property (1) holds for any pair of them. For example, one has by the Cauchy Schwarz inequality if $$x = (x_1, \ldots, x_n)$$

$$\|x\|_1 = \sum_{k=1}^{n} |x_k| = \sum_{k=1}^{n} 1 \cdot |x_k| = \left(\sum_{k=1}^{n} x_k^2 \right)^{1/2} \left(\sum_{k=1}^{n} x_k^2 \right)^{1/2} = \sqrt{n}\|x\|_2;$$

using the fact that the square root of a sum of positive numbers is less than or equal than the sum of the square roots,

$$\|x\|_2 = \left(\sum_{k=1}^{n} x_k^2 \right)^{1/2} \leq \sum_{k=1}^{n} |x_k| = \|x\|_1.$$ 

We proved that $$\| \cdot \|_1$$ and $$\| \cdot \|_2$$ are equivalent.

A celebrated theorem now states:

**Theorem 3** If $$V$$ is a finite dimensional vector space, then ALL norms are equivalent.

**Proof.** Because equivalence of norms is an equivalence relation, one way of proving this is to single out one norm, and prove that all others are equivalent to it. Because $$V$$ is finite dimensional it has a finite basis. Let $$n = \dim V$$ and let $$\{x_1, \ldots, x_n\}$$ be a basis of $$V$$. If $$x \in V$$ it can be written in the form $$x = \sum_{k=1}^{n} c_k x_k$$ for a unique choice of real numbers $$c_1, \ldots, c_n$$ and it is easy to see (exercise!) that

$$\|x\|_V = |c_1| + \cdots + |c_n|$$

defines a norm in $$V$$. Suppose now $$\| \cdot \|$$ is another norm in $$V$$. Then, if $$x = \sum_{k=1}^{n} c_k x_k \in V,$$

$$\|x\| = \left\| \sum_{k=1}^{n} c_k x_k \right\| \leq \sum_{k=1}^{n} |c_k| \|x_k\| \leq \left(\max_{1 \leq k \leq n} \|x_k\|\right) \sum_{k=1}^{n} |c_k|,$$

that is,

$$\|x\| \leq b\|x\|_V$$

where $$b = (\max_{1 \leq k \leq n} \|x_k\|)$$). This is half of (1); we'll be done if we prove there is $$a > 0$$ such that $$a\|x\|_V \leq \|x\|$$ for all $$x \in V$$. This is a bit harder to do, but also more interesting. The key step is to realize that the map

$$x \mapsto \|x\| : V \rightarrow \mathbb{R}$$

is continuous with respect to the $$\| \cdot \|_V$$ norm. This is an easy consequence of (2); in fact, let $$x_0 \in V$$. To prove continuity at $$x_0$$ we have to show that for
every $\epsilon > 0$ there is $\delta > 0$ such that if $\|x - x_0\|_V < \delta$, then $\|\|x\| - \|x_0\|\| < \epsilon$. It is an immediate (or easy) consequence of the triangle inequality that

$$\left\| \|x\| - \|x_0\| \right\| \leq \|x - x_0\|_V,$$

thus taking $\delta = \epsilon/b$ (as in (2)) we see at once that if $\|x - x_0\|_V < \delta$, then

$$\left\| \|x\| - \|x_0\| \right\| \leq \|x - x_0\|_V < \epsilon.$$

With the norm $\|\cdot\|_V$, the space $V$ is essentially the same thing as $\mathbb{R}^n$ with the norm $\|\cdot\|_1$; because this norm is equivalent to the Euclidean norm $\|\cdot\|_2$ (as proved above) the Heine-Borel theorem is valid; the set

$$S = \{x \in V : \|x\|_V = 1\}$$

is closed and bounded, hence compact (with respect to $\|\cdot\|_V$). Because the function $x \mapsto \|x\|$ is continuous with respect to $\|\cdot\|_V$ it assumes a minimum value $\alpha$ on the compact set $S$; there is $x_0 \in S$ such that

$$\|x\| \geq \|x_0\| = \alpha$$

for all $x \in S$. The important thing now is that because $x_0 \neq 0$ (since $0 \neq 1 = \|x_0\|_V$), we have $\alpha = \|x_0\| > 0$.

Assume now $x \in V$, $x \neq 0$. Then $x/\|x\|_V \in S$, hence

$$\left\| \frac{1}{\|x\|_V}x \right\| \geq \alpha,$$

but

$$\left\| \frac{1}{\|x\|_V}x \right\| = \frac{\|x\|}{\|x\|_V}$$

so that we proved $\|x\| \geq \alpha \|x\|_V$ for all $x \in V$, $x \neq 0$. But the inequality is trivially true if $x = 0$, thus we have completed the proof of the equivalence of the two norms.

The implication coming from this theorem is that if the vector space is finite dimensional, for the purpose of notions of convergence, continuity, completeness (it is very easy to see that if two norms are equivalent, they have the same Cauchy sequences), it doesn’t really matter what norm you use. Use the most convenient one. Sometimes it is a Euclidean norm, other times a different norm. It allows you to adjust the norm to the proof.

1.2 What are they good for

Function spaces appear now all over in analysis. They play a very big role in the theory of differential equations, in numerical analysis, and in modern physics, quantum mechanics, for example. We might see a little of this in this course.

2 The Theorem of Arzela-Ascoli.

The function space we have in mind in these notes is $C(K)$. Let $K$ be a compact metric space, then

$$C(K) = \{f|K \to \mathbb{R} : f \text{ is continuous}\}.$$
The norm in $C(K)$ is defined by

$$\|f\| = \sup_{x \in K} |f(x)| = \max_{x \in K} |f(x)|.$$ 

It is very easy to see that this defines a norm, so $C(K)$ is a normed space. We have

**Theorem 4** $C(K)$ is a complete normed space, hence a Banach space.

**Proof.** Exercise.

As is actually true in every infinite dimensional space, the Heine Borel theorem fails to hold. That is, there are closed, bounded subsets that are not compact; in fact every set with non empty interior fails to be compact. We have to assume $K$ is an infinite set; otherwise $C(K)$ is actually finite dimensional and thus Heine Borel will hold. Can you prove that $C(K)$ is infinite dimensional if and only if $K$ is an infinite set?

As it turns out, the Heine Borel theorem is ALWAYS false in an infinite dimensional normed vector space. But for many applications it is quite important to know exactly which are the compact subsets of a given normed vector space; the Arzela-Ascoli Theorem provides an answer for $C(K)$.

**Exercise 5** Consider $C([0, 1]) = \{ f : [0, 1] \to \mathbb{R} : f \text{ is continuous} \}$ Consider the sequence $\{f_n\}$ of elements of $C([0, 1])$ where

$$f_n(x) = x^n$$

for $n = 1, 2, \ldots$.

1. Prove that the sequence $\{f_n\}$ cannot have any convergent subsequence (not in the norm of $C([0, 1])$).

2. Use this to prove that the closed unit ball of $C([0, 1])$ is not compact.
for all \(x, y \in [0, 1]\). Each function \(f_n\) is uniformly continuous, but there is no \(\delta\) that works for all.

For a second example, consider the set of functions \(g_n : [0, 2\pi] \to \mathbb{R}\) defined by \(g_n(x) = \sin nx\). Once again, each function in the set is uniformly continuous, and, given \(\epsilon > 0\), for each \(n\) there is \(\delta > 0\) such that \(x, y \in [0, 2\pi]\), \(|x - y| < \delta\), implies \(|g_n(x) - g_n(y)| < \epsilon\). But if we look at the functions in question we see that the larger \(n\) is, the more oscillatory they become. For example the value 1 is assumed only once by \(g_1\), but it is assumed 1000 times by \(g_{1000}\). This shows that even more so than for the previous example, one \(\delta\) cannot work for all these functions.

The examples we just saw are examples of sets of continuous functions that are not equicontinuous. Maybe it is time to make a definition. For the sake of simplicity, I’ll assume the functions are real valued. But equicontinuity is a property that can be defined for sets of functions from one metric space to another (and more).

**Definition 3** Let \(X\) be a metric space and let \(S\) be a set of functions in \(X\) form \(X\) to \(\mathbb{R}\). We say that \(S\) is **equicontinuous** iff for every \(\epsilon > 0\) there is \(\delta > 0\) such that if \(x, y \in X\) and \(d(x, y) < \delta\), then \(|f(x) - f(y)| < \epsilon\) for all \(f \in S\).

**Lemma 6** Let \(X\) be a metric space and let \(S\) be an equicontinuous set of functions from \(X\) to \(\mathbb{R}\). Then every \(f \in S\) is uniformly continuous.

**Proof.** Too trivial to bother. \(\square\)

**Exercise 7** Prove: If \(X\) is a metric space, every finite set of continuous real valued functions on \(X\) is equicontinuous.

Very roughly a set will be equicontinuous if the functions in the set do not have too many oscillations. Also very roughly, if we are working on a compact domain (as we will) there isn’t enough room for a continuous function to have more than a finite number of oscillations, except if these oscillations get smaller and smaller. But if you have an infinite number of functions on a compact set, you can get each function to be more oscillatory than a preceding one, and end with a non-equicontinuous set.

We are ready to state a version of the Arzela-Ascoli theorem. I’ll keep it simple and state the version that is most commonly used in applications.

**Theorem 8** **Arzela-Ascoli** Let \(-\infty < a \leq b < \infty\) and let \(f_n : [a, b] \to \mathbb{R}\) be continuous for \(n = 1, 2, 3, \ldots\). If the set \(\{f_n : n \in \mathbb{N}\}\) is uniformly bounded and equicontinuous, then the sequence \(\{f_n\}\) has a uniformly convergent subsequence.

Before going into the proof, let us see what the hypothesis are, and what the conclusion should be.

**Hypotheses**

- \(f_n \in C([a, b])\) for \(m \mathbb{N}\).
- **Uniformly bounded** means that there is \(M \geq 0\) such that \(|f_n(x)| \leq M\) for all \(x \in [a, b], n \in \mathbb{N}\). Equivalently:

\[
\|f_n\| \leq M \quad \forall n \in \mathbb{N}.
\]
Equicontinuous means that for every $\epsilon > 0$ there is $\delta > 0$ such that if $x, y \in [a, b]$ and $|x - y| < \delta$, then $f_n(x) - f_n(y) < \epsilon$ for all $n \in \mathbb{N}$.

Conclusion
We can find integers $n_k$, $1 \leq n_1 < n_2 < n_3 < \cdots$ such that the sequence \{\(f_{n_k}\)\} converges uniformly; that is, converges in the norm of $C([a, b])$. Because $C([a, b])$ is complete, it suffices to prove that it is a Cauchy sequence in that norm.

Proof. We begin selecting a countable dense subset $D$ of the interval $[a, b]$, for example $D = \mathbb{Q} \cap [a, b]$. We order the elements of $D$ as a sequence, say $D = \{r_1, r_2, \ldots\}$.

Now we will use a technique known as the diagonal process, which is used in a lot of contexts. I’ll try to be as precise as I can. Let $M \in \mathbb{R}$, $M \geq 0$ be a uniform bound for all the functions; that is, let $M$ be as in (3). Consider the sequence of real numbers that we obtain evaluating all the functions at $r_1$, the first element of $D$; that is, consider the sequence of real numbers \{\(f_{n_k}(r_1)\)\}. This is a sequence in the interval $[-M, M]$; this interval being compact, it has a convergent subsequence. Normally we would use a notation such as \{\(f_{n_k}(r_1)\)\} to denote this subsequence, but we will have to take a subsequence of it, then a subsequence of the subsequence, and so forth, and our notation could become absolutely horrendous. So we will use a different way of indicating subsequences, with double indices. We denote the convergent subsequence by \{\(f_{1m}\)\}_{m=1}^{\infty}, where we consider \{\(1m\)\} a strictly increasing sequence of natural numbers. For example (it is a bit awkward, I know, but it could be worse) we might have \{\(1m\)\} = \{2^n\}, meaning

$$11 = 2, 12 = 4, 13 = 8, \ldots$$

Let $s_1 = \lim_{m \to \infty} f_{1m}(r_1)$. We now forget everything but the subsequence we just constructed; that is, we concentrate on \{\(f_{1m}\)\}. We evaluate all terms at $r_2$; the sequence of real numbers \{\(f_{1m}(r_2)\)\} is a sequence in the compact set $[-M, M]$, thus has a convergent subsequence. Let us denote it by \{\(f_{2m}\)\}_{m=1}^{\infty}.

The idea now is that each $2m$ is an integer from the list \{\(1m\)\}, and $21 < 22 < \cdots$. Let $s_2 = \lim_{m \to \infty} f_{2m}(r_2)$. Now the sequence \{\(f_{2m}(r_1)\)\} is a subsequence of \{\(f_{1m}\)\} hence also converges to $s_1$. And so forth; in other words: INDUCTION! Assume that for some $n \geq 1$ we have constructed a sequence \{\(f_{nm}\)\}_{m=1}^{\infty} such that $\lim_{m \to \infty} f_{nm}(r_n) = s_n$, then it has a subsequence \{\(f_{(n+1)m}\)\}_{m=1}^{\infty} such that $\lim_{m \to \infty} f_{(n+1)m}(r_{n+1}) = s_{n+1}$ for some $n+1 \in [-M, M]$. Here is the picture usually associated with all this:

\[
\begin{align*}
&f_{11}(r_1) \quad f_{12}(r_1) \quad f_{13}(r_1) \quad \cdots \quad s_1 \\
&f_{21}(r_2) \quad f_{22}(r_2) \quad f_{23}(r_2) \quad \cdots \quad s_2 \\
&f_{31}(r_3) \quad f_{32}(r_3) \quad f_{33}(r_3) \quad \cdots \quad s_3 \\
&\vdots \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \vdots
\end{align*}
\]

Now if we consider just the functions, each row is a subsequence of the preceding row, so that the functions of, for example, the third row, if evaluated at $r_1$, also converge to $s_1$. We now consider the diagonal of the scheme, that is, we consider the sequence \{\(f_{nn}\)\}_{n=1}^{\infty}. I know this is awkward, but $(n+1)(n+1)$ as a subindex is $\geq n(n+1)$, because \{\(n+1m\)\} is a subsequence of \{\(nm\)\}. Moreover, $n(n+1) > nm$; $n(n+1)$ comes right after $nm$ in the $n$-th subsequence. Thus $(n+1)(n+1) > nm$ and the diagonal sequence is a bona fide subsequence of the original sequence \{\(f_n\)\}. It is also clear (as clear as anything in this context!) that \{\(f_{nn}\)\} is a subsequence of \{\(f_{1m}\)\}. In fact \(f_{nn}\) is the $n$-th element of the $n$-th subsequence, which is a subsequence of the $n-1$-st sequence, which is a subsequence of the $n-2$-nd sequence, which is. . . a subsequence of the first subsequence. It is not exactly a subsequence of \{\(f_{2m}\)\}, but \{\(f_{nn}\)\}_{n=2}^{\infty} is a subsequence of
\{f_{2m}\}. In general, \(\{f_{nn}\}_{n=k}^\infty\) is a subsequence of \(\{f_{km}\}_{m=1}^\infty\). The conclusion is:

\[
\lim_{n \to \infty} f_{nn}(r_k) = s_k
\]

for \(k = 1, 2, 3 \text{ ldots} \).

Notice that we have not used equicontinuity at all, so far. We haven’t even used continuity. We only used that the functions are uniformly bounded. But now it will come, because we want to prove now that the subsequence \(\{f_{nn}\}\) is a uniform Cauchy sequence. For this let us give \(\epsilon > 0\). Let \(\delta > 0\) be the \(\delta\) coming from the condition of equicontinuity for \(\epsilon/3\); that is, \(\delta > 0\) is such that if \(x, y \in [a, b]\), \(|x - y| < \delta\), then \(|f_n(x) - f_n(y)| < \epsilon/3\) for all \(n\); in particular, \(|f_{nn}(x) - f_{nn}(y)| < \epsilon/3\) for all \(n\). Because the set \(D = \{r_1, r_2, \ldots\}\) is dense, the family of open intervals \(\{(r_k - \delta, r_k + \delta)\}_{k \in \mathbb{N}}\) covers \([a, b]\); because \([a, b]\) is compact there is a finite subcovering: There is \(K\) such that

\[
[a, b] \subset \bigcup_{k=1}^{K} (r_k - \delta, r_k + \delta).
\]

Because the sequences \(\{f_{nn}(r_k)\}\) converge, hence are Cauchy sequences, for each \(k \in \mathbb{N}\), there exist \(N_k\) such that \(n, \ell \geq N_k\) implies that \(|f_{nn}(r_k) - f_{\ell\ell}(r_k)| < \epsilon/3\).

We have all the ingredients set up for the grande finale.

Let \(N = \max(N_1, \ldots, N_K)\). Let \(n, \ell \geq N\). If \(x \in [a, b]\), there is \(k\), \(1 \leq k \leq K\) such that \(x \in (r_k - \delta, r_k + \delta)\). Because \(n, \ell \geq N \geq N_k\), \(|f_{nn}(r_k) - f_{\ell\ell}(r_k)| < \epsilon/3\).

We thus have

\[
|f_{nn}(x) - f_{\ell\ell}(x)| \leq |f_{nn}(x) - f_{nn}(r_k)| + |f_{nn}(r_k) - f_{\ell\ell}(r_k)| + |f_{\ell\ell}(r_k) - f_{\ell\ell}(x)| < |f_{nn}(x) - f_{nn}(r_k)| + \frac{\epsilon}{3} + |f_{\ell\ell}(r_k) - f_{\ell\ell}(x)| < \epsilon/3 + \epsilon/3 = \epsilon/3.
\]

But \(|x - r_k|, \delta\); thus by the equicontinuity based choice of \(\delta\), \(|f_{nn}(x) - f_{nn}(r_k)| < \epsilon/3\) and \(|f_{\ell\ell}(r_k) - f_{\ell\ell}(x)| < \epsilon/3\), hence

\[
|f_{nn}(x) - f_{\ell\ell}(x)| \leq |f_{nn}(x) - f_{nn}(r_k)| + \frac{\epsilon}{3} + |f_{\ell\ell}(r_k) - f_{\ell\ell}(x)| < \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon.
\]

Since \(x\) is arbitrary and \(N\) does not depend on \(x\), we proved that \(|f_{nn}(x) - f_{\ell\ell}(x)| < \epsilon\) for all \(x \in [a, b]\) whenever \(n, \ell \geq N\). The subsequence \(\{f_{nn}\}\) converges uniformly.

Convergence in \(C(K)\) is the same as uniform convergence; what we proved in proving the Arzela-Ascoli Theorem is that the diagonal sequence is a Cauchy sequence in the \(C([0, 1])\) norm, thus converges in that norm by completeness.

As a consequence of the previous theorem, we obtain the following characterization of compact subsets of \(C([a, b])\).

**Theorem 9 (Arzela-Ascoli, version 2)** A subset of \(C([a, b])\) is compact if and only if it is closed, bounded and equicontinuous.

**Proof.** Assume \(S\) is closed, bounded and equicontinuous. By the Arzela-Ascoli Theorem, if \(\{f_n\}\) is a sequence in \(S\), it has a convergent subsequence. Because \(S\) is closed, the limit of the subsequence must be in \(S\). Thus \(S\) is sequentially compact, hence compact.

Conversely, assume \(S\) is compact. Then, of course, it is closed and bounded.

To see it is equicontinuous, let \(\epsilon > 0\). There exist then, by compactness, a finite number of functions \(f_1, \ldots, f_m \in S\) such that \(S \subset \bigcup_{k=1}^{m} B(f_k, \epsilon/3)\). Because we have only a finite number of functions, there is a common \(\delta > 0\) such that \(|x - y| < \delta\) implies \(|f_k(x) - f_k(y)| < \epsilon/3\) for all \(x, y \in [a, b], |x - y| < \delta\). If
Let $f$ take values in a metric space; say also consider a more general situation, in which one works with functions that define a countable dense subset (can you prove this?). The proof goes through almost the first part of the course that every compact metric space is separable; i.e., has a countable dense subset and the compactness. You might remember from the proof of Arzela-Ascoli you will see that there is no need to restrict the domain to a closed and bounded interval of the real line. If you look and understand the proof of Arzela-Ascoli you will see that there is a countable dense subset (can you prove this?).

\[ |f(x) - f(y)| \leq |f(x) - f_k(x)| + |f_k(x) - f_k(y)| + |f_k(y) - f(y)| \]
\[ \leq \|f - f_k\| + |f_k(x) - f_k(y)| + |f_k(y) - f(y)| < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon. \]

How does one verify in practice that a family of functions (or a sequence of functions) is equicontinuous? In a vast majority of cases it is done using the mean value theorem, as we are about to see. In concrete cases one might just do it by direct verification, but the theorem is not used so much in a concrete case, but in trying to prove a general theorem. I’ll try to illustrate this in the next section. Here is the standard result on how one verifies equicontinuity.

**Theorem 10** Let $S \subset C([a, b])$, $-\infty < a \leq b < \infty$. If each $f \in S$ is differentiable in $(a, b)$ and there is $C$ such that $|f'(x)| \leq C$ holds for all $f \in S$ and all $x \in (a, b)$, then $S$ is equicontinuous.

**Proof.** By the mean value theorem, for every $f \in S$, $x, y \in [a, b]$, there exists $z$ (depending naturally on $f, x, y$) between $x$ and $y$ such that $f(x) - f(y) = f'(z)(x - y)$. Taking absolute value and using the hypothesis, we see that $|f(x) - f(y)| \leq C|x - y|$ for all $f \in S$, $x, y \in [a, b]$. Given $\epsilon > 0$, it suffices to take $\delta = \epsilon/C$ to get $|f(x) - f(y)| < \epsilon$ for all $f \in S$, $x, y \in [a, b]$ such that $|x - y| < \delta$.

3 Extensions and applications

If you look and understand the proof of Arzela-Ascoli you will see that there is no need to restrict the domain to a closed and bounded interval of the real line. The only properties of the interval $[a, b]$ that played a role where the existence of a countable dense subset and the compactness. You might remember from the first part of the course that every compact metric space is separable; i.e., has a countable dense subset (can you prove this?). The proof goes through almost identically for $C(K)$, where $K$ is an arbitrary compact metric space. One can also consider a more general situation, in which one works with functions that take values in a metric space; say $C(K, Y)$, where $Y$ is a metric space and we define $f \in C(K, Y)$ to mean $f : K \rightarrow Y$ is continuous. In this case equicontinuity and boundedness of the set $S$, and the compactness of $K$ is not quite enough. Part of the proof was to be able to extract convergent subsequences from a sequence of the form $\{f_n\}$, where $x \in K$ and $\{f_n\}$ is a sequence in $S$. The reason we could do it was because the functions were uniformly bounded; that is, all take values in an interval $[-M, M]$ and by Heine-Borel such an interval is compact. But for a more general space one needs to be more specific; for example, require that all functions take their values in a compact subset of $Y$.

Here is an application; it may not be the best but the best may be outside of our scope. Suppose we want to solve the differential equation

\[ y' = f(x, y) \]
with an initial value \( y(0) = y_0 \). If a differentiable solution exists in an interval \([a, b] \), then the fundamental theorem of calculus tells us that it will satisfy

\[
y(x) = y_0 + \int_0^x f(t, y(t)) \, dt.
\]

This changes the differential equation to an integral one; at first glance nothing much has been achieved but for many reasons integral equations are more manageable than differential ones. If \( f : \mathbb{R}^2 \to \mathbb{R} \) is continuous, and is Lipschitz in \( y \): there exists \( C \) such that

\[
|f(x, y_1) - f(x, y_2)| \leq C|y_1 - y_2|
\]

for all \((x, y_1), (x, y_2) \in \mathbb{R}^2\) then the general theory states that there exists a unique solution defined in some interval \([a, b]\) where (with the hypothesis as stated!) either \( b = \infty \) or \( \lim_{x \to b^-} |y(x)| = \infty \). From the integral form of the equation, we see that if we have the additional hypothesis that \( f \) is bounded, say by \( M \), then \( y \) is bounded by \( |y_0| + Mx \) and cannot have an infinite limit at a finite \( b \), thus \( b = \infty \) in this case.

What happens if \( f \) is not Lipschitz in \( y \)? As one learns even in our rather low level differential equations course, there are cases in which the solution is not unique. The standard example with a continuous \( f \) is \( f(y) = y^{2/3} \); that is: Solve \( y' = y^{2/3} \) satisfying \( y(0) = 0 \). The obvious solution is \( y \equiv 0 \). But so is \( y = 27x^3 \). So uniqueness fails; but does a solution exist?

I will sketch a proof of the following theorem:

**Theorem.** Let \( f : \mathbb{R}^2 \to \mathbb{R} \) be bounded, say \( |f(x, y)| \leq M \) for all \((x, y) \in \mathbb{R}^2\), and continuous. Let \( y_0 \in \mathbb{R} \). Let \( b > 0 \). There exists a differentiable function \( \phi : [0, b] \to \mathbb{R} \) such that \( \phi(x) = f(x, \phi(x)) \) for all \( x \in [a, b] \).

In other words, the initial value problem \( y' = f(x, y), y(0) = y_0 \), has a solution in every interval \([0, b]\).

To prove the theorem we will use a fact whose proof is outside of our current scope, namely that one can approximate \( f \) by very nice functions. Specifically, there exists a sequence \( \{f_n\} \) of differentiable functions on \( \mathbb{R}^2 \) (main fact: They are continuous and Lipschitz in \( y \)) such that for every compact subset \( C \) of \( \mathbb{R}^2 \), the restrictions \( \{f_n|_C\} \) converge uniformly to \( f \) on \( C \), and such that \( |f_n(x, y)| \leq M \) for all \((x, y) \in \mathbb{R}^2, n \in \mathbb{N}\).

Let’s accept this. Then for each \( n \in \mathbb{N} \) we can solve the initial value problem \( y_n' = f_n(x, y_n), y_n(0) = y_0 \) in all of \([0, \infty)\). That is, there exists a differentiable \( \phi_n : [0, \infty) \to \mathbb{R} \) such that \( \phi_n(x) = f_n(x, \phi_n(x)) \) for all \( x \geq 0; \phi_n(0) = x_0 \).

Let us restrict all this to the closed and bounded interval \([0, b]\) and consider the sequence \( \{\phi_n\} \). It is uniformly bounded; integrating the equation (as before) to get to the integral equation,

\[
\phi_n(x) = y_0 + \int_0^x f_n(t, \phi_n(t)) \, dt
\]

from which \( |\phi_n(x)| \leq |y_0| + Mx \leq |y_0| + Mb \) for all \( x \in [0, b] \). Let us set \( M = |y_0| + Mb \) for use later on. Moreover,

\[
|\phi_n'(x)| = |f_n(x, \phi_n(x))| \leq M
\]

for all \( x \in [0, b], n \in \mathbb{N} \); by Theorem 10, the sequence \( \{\phi_n\} \) is equicontinuous thus there is a subsequence \( \{\phi_{n_k}\} \) that converges uniformly on \([0, b]\), to some continuous function \( \phi \).

We still need to see that \( \phi \) solves the initial value problem. Now \( |\phi_n(x)| \leq \tilde{M} \) for all \( x \) implies that \( \phi \) satisfies the same bound. The set \( A = [0, b] \times [-M, M] \) is
compact and, by our assumption, \( \{f_n\} \), hence also \( \{f_{n_k}\} \), converges uniformly to \( f \) on this set. Given \( \epsilon > 0 \) there is thus \( K_1 \) such that if \( k \geq K_1 \), then 
\[
|f_{n_k}(x,y) - f(x,y)| < \epsilon/2 \quad \text{for all } (x,y) \in A.
\]
Because \( f \) is uniformly continuous on \( A \), there is \( \delta > 0 \) such that if \( (x,y),(x',y') \in A \) and \( |(x,y) - (x',y')| < \delta \), then 
\[
|f(x,y) - f(x',y')| < \epsilon/2.
\]
Because \( \{\phi_{n_k}\} \) converges uniformly to \( \phi \) on \( [a,b] \), there is \( K_2 \) such that if \( k \geq K_2 \) we have \( |\phi_{n_k}(x) - \phi(x)| < \delta \) for all \( x \in [0,b] \).

If now \( x \in [0,b] \) and \( k \geq \max(K_1,K_2) \), then \((x,\phi_{n_k}(x))\), \((x,\phi(x))\) are points in \( A \) satisfying (because \( k \geq K_2 \))
\[
|(x,\phi_{n_k}(x)) - (x,\phi(x))| = |\phi_{n_k}(x) - \phi(x)| < \delta, \quad \text{hence } |f(x,\phi_{n_k}(x)) - f(x,\phi(x))| < \epsilon/2.
\]

In addition, because \( k \geq K_1 \),
\[
|f_{n_k}(x,\phi_{n_k}(x)) - f(x,\phi(x))| < \epsilon/2.
\]

Thus, for \( x \in [0,b] \) and \( k \geq \max(K_1,K_2) \) we get
\[
|f_{n_k}(x,\phi_{n_k}(x)) - f(x,\phi(x))| \leq |f_{n_k}(x,\phi_{n_k}(x)) - f(x,\phi_{n_k}(x))| + |f(x,\phi_{n_k}(x)) - f(x,\phi(x))| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon;
\]
the sequence \( \{x \mapsto f_{n_k}(x,\phi_{n_k}(x))\} \) converges uniformly to \( x \mapsto f(x,\phi(x)) \) on \( [0,b] \) as \( k \to \infty \). Now equation (4) is of course valid for \( n = n_k, \ k = 1,2,3 \ldots \) and because of the uniform convergence we can take limits for \( k \to \infty \) to get that \( \phi \) satisfies the integral equation
\[
\phi(x) = y_0 + \int_0^x f(t,\phi(t)) \, dt.
\]

From this we see, setting \( x = 0 \), that \( \phi(0) = y_0 \) (this is, of course, also a consequence of the uniform convergence of \( \{\phi_{n_k}\} \) to \( \phi \)). Because \( f, \phi \) are continuous, it follows that the function \( x \mapsto f(x,\phi(x)) \) is continuous hence, by the fundamental theorem of calculus, \( \phi \) (as an integral of a continuous function) is differentiable and
\[
\phi'(x) = f(x,\phi(x))
\]
for all \( x \in [0,b] \). This proves that \( \phi \) solves the initial value problem.