Notes for 1 Feb Jan (Wednesday)

1 Recap

1. We finished compactness and the fact that the Heine Borel theorem implies that compactness in $\mathbb{R}^n$ is equivalent to every infinite subset having a limit point.
2. We showed that every perfect set is uncountable and constructed an example (the Cantor set) of a perfect set having no open interval in it.
3. We defined connectedness. (A set $E \subset X$ is connected if and only if it canNOT be written as a disjoint union of two *relatively* open subsets. For example, $E = (-\infty, 1] \cup [2, \infty)$ is not connected because it is a disjoint union of two *relatively* open subsets. (Note that $(-\infty, 1]$ is not an open subset of $\mathbb{R}$ but it is a relatively open subset of E because it is equal to $(-\infty, 1.5) \cap E$.

2 Sequences

Firstly, recall that a sequence of elements from a set$X$ is a map $f : \mathbb{Z}_+ \to X$ written as $x_1, x_2, \ldots$. A subsequence, written as $x_{n_1}, x_{n_2}, \ldots$ is a subset of a sequence where $n_1 < n_2 < n_3 \ldots$, i.e., you compose f with an injective order-preserving map from Z_+ to itself.

While we will mostly be interested in sequences of real numbers, we will try to be as general as possible and do things over general metric spaces (X, d).

A sequence x_n is said to converge to an element x (written as $x_n \to x$) if for every $\epsilon > 0$ there exists a natural number N_ϵ such that $d(x_n, x) < \epsilon \ \forall \ n > N_\epsilon$. The number x is called the limit of the sequence. If x_n does not converge to anything, it is said to diverge. (Like the themes of my lectures.) A sequence is said to be bounded if $d(x_n, p) < M \ \forall \ n$ for some p and M.

Here are some basic properties of convergence of sequences in general metric spaces :

1. A sequence converges to x if and only if every neighbourhood of x contains all but finitely many terms of the sequence.
   Pf : Exercise.
2. Limits are unique.
   Pf : Suppose there are two limits x and y of a sequence. Then $d(x, y) \leq d(x, x_n) + d(x_n, y) \ \forall \ n$. In particular, choosing n to be so large that $d(x_n, x), d(x_n, y) < \frac{\epsilon}{2}$, we see that $d(x, y) \leq \epsilon \ \forall \ \epsilon > 0$. This means that $d(x, y) = 0$. Therefore $x = y$.

3. Convergent sequences are bounded.
   Pf : Exercise.

4. If $E \subset X$ and x is a limit point of E, then there exists a sequence $x_n \in E$ such that $x_n \to x$.
   Pf : Since x is a limit point of E, for every natural $n > 0$, there exists a point $x_n \neq x \in E \cap B_{1/n}(x)$. This means that $d(x_n, x) < \frac{1}{n} \ \forall \ n > 0$. Thus $x_n \to x$.

5. A sequence converges to x if and only if every subsequence converges to x.
   Pf : Assume that $x_n \to x$. Suppose x_{m_i} is a subsequence. Given $\epsilon > 0$, there exists an N_ϵ such that $d(x_n, x) < \epsilon \ \forall \ n > N_\epsilon$. In particular, there exists an I_ϵ such that $i > I_\epsilon \to m_i > N_\epsilon$. Therefore for all $i > I_\epsilon$, $d(x_{m_i}, x) < \epsilon$.
   Assume that x_n does not converge to x. Then there exists an $\epsilon > 0$ such that there is no N_ϵ (no matter how large) satisfying $d(x_n, x) < \epsilon \ \forall \ n > N_\epsilon$. This means that for every natural $i > 0$ there exists a natural m_i such that $d(x_{m_i}, x) \geq \epsilon$. Therefore x_{m_i} cannot converge to x.

The last property is quite useful to prove divergence of sequences. For example, $-1, 1, -1, 1, \ldots$ diverges.

For sequences in $\mathbb{R}^k$, we have some more properties that tell us how addition and multiplication behave with sequences. Suppose $x_n \to x$ and $y_n \to y$ in $\mathbb{R}^k$.

1. $x_n + y_n$ converges to $x + y$.
   Pf : Choose n to be so large (i.e. $> N$) that x_n, y_n are $\epsilon/2$ close to x, y. Then $\|x_n + y_n - x - y\| \leq \|x_n - x\| + \|y_n - y\| < \epsilon$ for all $n > N$.

2. αx_n converges to αx.
   Pf : Exercise.

3. If we are working over $\mathbb{R}$, then $x_n y_n$ converges to xy.
   Pf : $|x_n y_n - xy| = |x_n(y_n + y - y) - xy| = |x_n(y_n - y) + (x_n - x)y| \leq |y_n - y||x_n| + |x_n - x||y|$. Now choose $n > N$ such that $|x_n|, |y| < M$ and $|y_n - y| < \frac{\epsilon}{2M}$, $|x_n - x| < \frac{\epsilon}{2M}$. Then $|x_n y_n - xy| < \epsilon$.

4. $x_n \to x$ if and only if each coordinate converges.
   Pf : Exercise.

5. If we are working over $\mathbb{R}$, then $\frac{1}{y_n}$ converges to $\frac{1}{y}$ if $y \neq 0$ and $y_n \neq 0$.
   Pf: Note that $|\frac{1}{y_n} - \frac{1}{y}| = |\frac{y - y_n}{y y_n}|$. Since $y \neq 0$ and $y_n \neq 0$, choose $n > N_1$ to be so large that $|y_n| \geq |y| - |y_n - y| \geq |\frac{y}{2}|$. Then $|\frac{1}{y_n} - \frac{1}{y}| \leq \frac{2}{|y|}|y_n - y|$. Now choose $n > N_2 > N_1$ so that $|y_n - y| < \epsilon \frac{|y|}{2}$. □

3 Subsequences

Two extremely important results about subsequences are the following-

Theorem 1. *1. Every sequence p_n in a compact set $E \subset X$ of a metric space contains a convergent subsequence p_{n_k} (whose limit is in E).*

2. *Every bounded sequence p_n in $\mathbb{R}^k$ contains a convergent subsequence p_{n_k}.*

Proof. 1. Zeroethly, if p_n does not consist of infinitely many distinct terms, then trivially it has a convergent subsequence (Why?).
Firstly, any limit point p of the set $\{p_1, p_2, \ldots\}$ is a limit of some subsequence (thus justifying the name "limit point"). Indeed, in every neighbourhood $B_{1/k}(p)$ there exists a point $p_{n_k} \neq p$ lying in $E \cap B_{1/k}(p)$. By definition this means that $p_{n_k} \to p$ as $k \to \infty$.
Secondly, suppose no point p of E is a limit point of the $\{p_1, p_2, \ldots\}$. Then for any p there exists a neighbourhood $B_{r_p}(p)$ that does not contain any points of E (other than p itself). Now $E \subset \cup_p B_{r_p}(p)$. Therefore by compactness, E is covered by finitely many sets $B_{r_1}(p_1), B_{r_2}(p_2), \ldots, B_{r_l}(p_l)$. But since E is infinite, this is not possible. (If every neighbourhood $B_{r_i}(p_i)$ contains only point from E, namely, p_i, and these neighbourhoods cover E, then how come E has infinitely many points ?)

2. By the Weierstrass theorem, a bounded sequence has a limit point. By the above reasoning, that limit point is indeed the limit of a subsequence. (By the way, this theorem is sometimes called the Bolzano-Weierstrass theorem.)

□

Here is another (admittedly technical) theorem.

Theorem 2. *The subsequential limits of a sequence $\{p_n\}$ in a metric space X form a closed subspace $E \subset X$.*

Proof. We just need to prove that E contains all of its limit points. Well, suppose p is a limit point of E. This means that for every natural $k > 0$, the neighbourhood $B_{1/k}(p)$ contains a subsequential limit $l_k \neq p$ of the sequence, i.e., $p_{k,n_m} \to l_k$ as $m \to \infty$. Intuitively, this means that there should be points of the sequence arbitrarily close to l_k (which is contained in $B_{1/k}(p)$. I leave it as an exercise (using the triangle inequality) to show that this implies that there exists an element of the sequence, which we denote as p_{n_k} that is in $B_{1/k}(p)$. Thus by definition of convergence, p is the limit of the subsequence p_{n_k}. Hence $p \in E$. □

4 Cauchy sequences and completeness

Now we shall define the notion of Cauchy sequences in general metric spaces. Indeed, p_n is said to be Cauchy if for every $\epsilon > 0$ there exists an N_ϵ such that $n, m > N_\epsilon$ implies that $d(p_n, p_m) < \epsilon$, i.e., the elements get close to each other if we go far into the sequence.

We take a small digression before returning to the topic of Cauchy sequences. There is a very convenient geometric notion that we shall use later on. If $E \neq \phi \subset X$, then let $\sup_{x,y \in E} d(x, y)$ (if it exists) is called the diameter of E and written as $diam(E)$. If the supremum does not exist (which can only happen if E is unbounded), then $diam(E)$ is said to be infinity, i.e., $diam(E) = \infty$. So for instance, a sequence is Cauchy if and only if as $N \to \infty$, the $diam(\{p_N, p_{N+1}, \ldots\}) \to 0$. Here is a technical result about diameters.

Theorem 3. *1. $diam(\bar{E}) = diam(E)$ where recall that $\bar{E}$ is the closure of E (recall further that the closure is the smallest closed set containing E and that it is simply E along with all of its limit points).*

2. If K_n is a collection of compact sets such that $K_n \subset K_{n+1}$ and if $\lim_{n\to\infty} diam(K_n) = 0$ then $\cap K_n$ consists of a single point.

Proof. 1. Since $E \subset \bar{E}$, of course $diam(E) \leq diam(\bar{E})$. So we need to prove the opposite inequality to conclude equality. Given $x, y \in \bar{E}$, if we manage to prove that for every $\epsilon > 0$ there exist $a_\epsilon, b_\epsilon \in E$ such that $d(x, y) \leq d(a, b) + \epsilon$, then of course $d(x, y) \leq diam(E) + \epsilon \ \forall \ \epsilon > 0$ implying that $d(x, y) \leq diam(E) \ \forall \ x, y \in \bar{E}$ thus allowing us to conclude the result.
Indeed, since the closure consists of E along with its limit points, given an $\epsilon > 0$ and $x, y \in \bar{E}$, surely there exist points $a_\epsilon \in E \cap B_{\epsilon/2}(x)$ and $b_\epsilon \in E \cap B_{\epsilon/2}(y)$. Now $d(x, y) \leq d(x, a_\epsilon) + d(a_\epsilon, y) \leq d(x, a_\epsilon) + d(a_\epsilon, b_\epsilon) + d(b_\epsilon, y) < \epsilon + d(a_\epsilon, b_\epsilon)$

2. By previous results, we know that $\cap K_n \neq \phi$, i.e., $\exists p \in \cap K_n$. We just need to prove that there is no point $q \neq p$ in $\cap K_n$. Indeed, suppose $\exists q \neq p \in \cap K_n$. This means that $diam(K_n) \geq d(p, q) \neq 0$. A contradiction.

□

Lastly, we prove an extremely important result. (Indeed, sometimes people like to define real numbers using the last property of the following theorem.)

Theorem 4. *1. In any metric space (X, d), every convergent sequence is Cauchy.*

2. If $E \subset X$ is a compact set, and $p_n \in E$ is a Cauchy sequence in E, then it converges to a limit $p \in E$.

3. In $\mathbb{R}^m$, every Cauchy sequence converges.

Proof. 1. Indeed, if $p_n \to p$, then choose an N so that $n, m > N \to d(p_n, p), d(p_m, p) < \frac{\epsilon}{2}$. Then $d(p_n, p_m) < d(p_n, p) + d(p, p_m) < \epsilon$.

2. p_n has a convergent subsequence that converges to a point $p \in E$. Here is a general statement - If a subsequence of a Cauchy sequence converges to a point p, then the sequence itself converges to p. Indeed, choose an N so large that $d(p_{n_k}, p) < \frac{\epsilon}{2} \ \forall \ k \geq N$ and $d(p_i, p_j) < \frac{\epsilon}{2} \ \forall \ i, j > N$. Now for all $n > max(N, n_N)$ we see that $d(p_n, p) < d(p_n, p_{n_N}) + d(p_{n_N}, p) < \epsilon$.

3. We just need to show that the given Cauchy sequence p_n is contained in a compact set E. (Then the previous result shows that it indeed converges.) Since the sequence is Cauchy, its diameter is finite (why?). So it is bounded. So it has a convergent subsequence. (Bolzano-Weierstrass.)

□

The importance of Cauchy sequences in mathematics cannot be overstated. Notice that real numbers are so special, that to know whether a sequence converges to something or not, you do *not* have to know what its limit is. You simply need to check whether it is Cauchy or not. So there is no need for clever guesswork as to what its limit may be.

This property is so useful that it is given a name -
A metric space (X, d) is said to be *complete* every Cauchy sequence converges.
(Why is this expected to be useful ? Very often, if you are asked to solve an equation on a computer (whether it is the root of a polynomial, the eigenvectors of a matrix, an ODE, or even worse, a PDE), you usually try to come up with an iterative algorithm to solve it. Theoretically speaking (even practically for that matter), you need to know whether the algorithm converges at all (and if it does, whether it does to the right limit or not). For these things, you do not need to know anything about what it might converge to. You just need to prove that you have a Cauchy sequence in a complete metric space. (How do you cast an ODE or a PDE in this framework ? Well, stay tuned is all I can say for now.)