[Warning: This is unedited uncorrected weekly notes]

week. [page]

1. Where are we, things to think about, sets and set operations, functions cardinality, inverse and forward images; posets and losets [03]

2. Cantor's description of Q, countable dense subsets of R, Cantor-Shroder-Bernstein, Axiom of choice [15]

3. AC, Zorn, Cardinality, Hamel basis [23]

4. Understanding of R, construction of R, R from Q. [35]

5. Dedikind cuts, Peano Arithmetic, (N_0, S) to $(N_0, +, \cdot, \leq)$, N_0 to Z to Q; Peano again, Metric spaces. [43]

6. Holder, Minkowski, limit points, Open/Closed sets, connected spaces, complete spaces, Cantor construction of R. [60]

7. Cantor's R, open/closed sets, interior/closure, compact sets, small sets, R is not union of small sets [71]

8. Cantor intersection theorem, Baire, Nowhere differentiable functions, completion of metric spaces. [87]

9. completion, category, pseudo metric, isometry, Banach's contraction mapping [98]

10. Banach, Contractions in R, linear contractions in R^n, inverse function theorem, differential equations and Picard, integral equations and iterations, compact spaces, continuity [107]

11. continuous functions, homeomorphisms, separable spaces, compactness, Arzela-Ascoli [123]

12. Home assignments, Arzela-Ascoli [132]

13. Residual sets, totally bounded sets, $\epsilon-$nets, pre compactness, Arzela-Ascoli, Stone-Weierstrass, Space filling/Peano curves, Cantor sets [142]

14. Complex $n-$dimensional space, Fourier series, exponentials, $C[0,1]$ as infinite dimensional C^n; Orthonormal vectors and Fourier series, Bessel, Dirichlet, Plancherel, $\sum(1/k^2)$ [156]

15. Plancherel jacobi identity, Home assignment. [174]

Home Assignments [187 – 236]

[187 – 191 – 195 – 199 – 203

207 – 211 – 215 – 219 – 222

226 – 228 – 231 – 234 –

Midsem 237 semestral 238-239

where are we:

So far we learnt some basic properties of real numbers,

convergence of sequences (helped to 'approximate' unknown numbers with known fractions; discovered new numbers like Euler's constant, convergence helped us to make precise statement like 'how does this number look like for large values of n' as in Stirling formula for $n!$ or Walli's product for $\pi/2$),

convergence of series (helped us to add infinitely many numbers which in turn helped us to make sense of power series — natural extension of the high school concept of polynomial — gives us a way to discover functions),

continuity and differentiation of functions from R to R (a class of nice functions whose graph can hopefully be drawn, functions which can be understood better, which can be used in our descriptions of physical phenomena and so on),

integration of functions (helps in calculating areas; later will also be used to discover functions by knowing certain equations obeyed by their derivatives; helps also to discuss convergence of series; allows us to define new functions and so on),

and then we learnt functions of several variables.

This last topic will be further continued in Calculus III to develop various useful stories. In our course we shall discuss Analysis.

Basically we discuss: sets; numbers; sets of numbers. At the end, if we have time, we shall discuss some Fourier series. Actually, questions concerning Fourier series lead to the (discovery and) development of set theory by Cantor.

But then, did we not discuss already sets of numbers — compact sets, Cantor set and so on? Yes. But there was a gap in our understanding. We did not discuss what sets are and their properties.

We did not know if there are real numbers at all — we only postulated that there is a set with certain operations (addition and multiplication); worked with such a set after making a mental picture of it as a line. Even if there is such a system, we have not proved that there is only one such system. All these issues need to be settled.

Things to think about:

Before we get into the subject, a word of advise is in order.The first thing you should develop is respect for the words you use. Many a times I see that you are using technical words and I get the impression that you do not know its meaning. Probably you know meanings, but are careless. Or probably lack discipline. It is for you to think.

You should try to figure out negation of sentences. For example, not many could write what is meant by f is not continuous. There is no maths in this, really speaking. For example consider,

Every student in this class has a red pen.

You seem to think its negation is the following.

Every student in the class has a green pen.

Every student has no red pen.

Student in other class have red pen.

and so on.

You are under the correct impression that matters are simple, but under the wrong impression that you need not think.

Consider the sentence: $x = 5$. Its negation is simple. $x < 5$ or $x > 5$. This sentence has no 'quantifier'. When you have phrase 'every student' or 'there exists a student' you are using quantifier. Negation of statements involving quantifiers are tricky unless you practice. If there are two quantifiers you need to think a little more. Try negating: for every student there is a problem in this set which is difficult

Many times you seem to think that explanation can be taken as proof. For instance consider the following statement.

BVR does not know Tamil because he is from Andhra Pradesh.

This is of the form: 'S because T', where S and T are statements. You should realize that T is only a plausible explanation for the statement S, but not its proof. In daily use we do accept such explanations as proof, but not in maths. After all, there are persons who satisfy statement T and know Tamil too. Thus T does not allow you to conclude S. In other words, knowing T is true does not mean S is true.

It is possible to define precisely what is meant by proof. Basically, if you make a sequence of sentences and tell me it is proof then (i) each sentence in that sequence must follow from earlier sentences or must be a hypothesis and (ii) the last sentence is the one you are proving. Think about it. I do not want to convert this into Logic class. Moreover knowing such a definition will not help you. The only thing that helps is practice. So practice. Do not try to convert exam into practice session.

Sets:

A set is any collection of objects. If S is a set and x is an object in this set then we write $x \in S$. If x is not an object in this set we write $x \notin S$.

This definition is good enough for us. The reason for making this statement is that if you literally take the above definition you will end up in problems.

You must have heard about the barber paradox. Consider only males. A barber in a town declared that he shaved those and only those who have not shaved themselves. Decide whether the barber shaved himself or not. Thus if you consider the set S of all persons in the town whom the barber shaved, you are unable to decide whether the barber himself is a member of this set or not. Thus even though S is defined as a collection of objects, we are unable to decide if a specific object (namely, the barber) is in this set or not.

So sometimes you see a definition like: a set is a well defined collection of objects. But then, what is meant by well defined? You have to start from somewhere. Thus for us a set is a collection of objects.

There are other such paradoxes too. Let k be the least natural number that can not be described in less than hundred letters. Does this make sense?

Or, Consider the collection S of all sets. This is clearly a collection of objects. Is this a set? If so it should belong to itself. Do you believe it?

Set operations:

If S and T are two sets, then they are same and we write $S = T$; if they have the same objects. In other words $x \in S$ implies $x \in T$ and $x \in T$ implies $\in S$.

We say S is a subset of T and write $S \subset T$; if $x \in S$ implies $x \in T$. Thus every object which is in S is also in T. We also say that T is a superset of S and write $T \supset S$.

If S and T are two sets then their union $S \cup T$ is the collection of those objects which are either in S or in T. Similarly their intersection $S \cap T$ consists of all objects which are in both S and T.

$$S \cup T = \{x : x \in S \text{ or } x \in T\}; \quad S \cap T = \{x : x \in S \text{ and } x \in T\}.$$

We define $S - T = \{x : x \in S; \quad x \notin T\}$, that is, it consists of all objects which belong to S but which do not belong to T.

In case we have a grand set Ω under consideration and if we are only considering objects that belong to Ω, then for $\Omega - S$ we write S^c and is called complement of S. Since objects which are in Ω are under discourse, it is called universe. Sometimes it is also called universal set, but there is nothing universal about it; the word only means that it is the universe now and all objects we are talking about are in this set (whether we say so or not).

Thus, unlike union and intersection, complement must be with reference to something. It may be specified or understood from the context.

Theorem (DeMorgan's laws):

$$(A \cup B)^c = A^c \cap B^c; \quad (A \cap B)^c = A^c \cup B^c.$$

Proof:

$x \in (A \cup B)^c$
$\Rightarrow x \notin (A \cup B)$ (def. complement)
$\Rightarrow x \notin A$; and $x \notin B$ (def union)
$\Rightarrow x \in A^c$ and $x \in B^c$ (def complement)
$\Rightarrow x \in A^c \cap B^c$ (def intersection).
These arrows can be reversed.
The other equality can be proved similarly.

Many times we use the word 'similar'. You should not blindly copy it like a faithful elementary school student. This phrase is an abbreviation for the following. 'I have done the proof once, it is similar to the earlier part and so I omit writing it (probably to spare the reader from getting bored).' As a result you should verify and earn the eligibility to use this phrase.

A set which has no objects is called empty set (that such a thing exists is an axiom). It is denoted by $\emptyset$.

We use numerals: $0, 1, 2, 3, \cdots$ We can precisely define them as follows, you need not keep this abstract way in mind. You can think of them as you have been doing all along.

$$0 = \emptyset; \quad 1 = \{\emptyset\}; \quad 2 = \{\emptyset, \{\emptyset\}\}; 3 = \{\emptyset, \ \{\emptyset\}, \ \{\emptyset, \{\emptyset\}\}\}$$

In general k is the set consisting of all the previous sets. Thus everything is a set. There is nothing before zero. So zero is empty set. Before one there is zero. So one consists of just the empty set.

The purpose of all this is just to convince that everything can be defined using only one symbol: empty set. (Every thing grew from empty set!)

N stands for the set $\{1, 2, 3, \cdots\}$.

Let A and B be two sets. Their Cartesian product $A \times B$ is defined as the set of all ordered pairs, with first element of the pair being an object from A and the second element being an object from B.

$$A \times B = \{(x, y) : x \in A; y \in B\}.$$

Of course I have used the word ordered pair (x, y) and you might wonder what the hell it is. If you are wondering, here it is

$$(x, y) = \{x, \{x, y\}\}$$

It is a set consisting of two objects. First object is x. Second object is the set consisting of x and y. Of course you can keep your mental picture in tact and ignore this definition. This definition of ordered pair is only to reinforce the feeling that ordered pair can be defined precisely. (everything is a set!)

functions and cardinality:

Let X and Y be sets. A function on X to Y is a rule that associates with every element of X one element of Y.

You can keep this in mind and there is nothing wrong with this definition. However you may suddenly get a doubt: what is this 'rule' and 'association' business? What is rule and what is not rule? What is association and what is not association?

If you get such a doubt, there is no need to worry. Here is something that helps. A function is a subset $R \subset X \times Y$ such that

$$\forall x \exists ! y \ \ (x, y) \in R.$$

Here $\exists!$ is an abbreviation for 'there is a unique'. Thus for each $x \in X$ there is exactly one y such that $(x, y) \in R$. If you want to denote your function by f then for a given $x \in X$, this unique point y is denoted by $f(x)$, value of the function f at the point x. Thus R associates with every x one y.

If you go back in time, you realize that a function f is a rule as said earlier and 'graph' of the function f is the set of all points (x, y) such that $y = f(x)$. Now if you want to be rigorous and not use unnecessary words like rule/association etc, then you turn things around. *By definition* the graph is regarded as function. Think about it.

But let me assure you that you should keep in mind the same good old idea of function. What all I said is an explanation, in case you feel you are cheated.

I should say two things. Firstly, as already said, the symbol $\exists! y$ is an abbreviation for there exists unique y. Thus

$$\forall x \exists ! y \ \ (x, y) \in R$$

means

$$\forall x [\{\exists y \ \ (x, y) \in R\} \ \ \wedge \ \ \{(x, y_1) \in R, (x, y_2) \in R \Rightarrow y_1 = y_2\}].$$

Second point is the following. Sometimes you see a different (not equivalent to ours) definition of function. It is a subset $R \subset X \times Y$ such that

$$(x, y_1) \in R, (x, y_2) \in R \Rightarrow y_1 = y_2.$$

Thus for any point x there can not be two y. However it is quite possible there is no y at all. Thus for each x there is at most one y such that $(x, y) \in R$. The set of x for which there is such y is called the 'domain' of the function and for x in the domain the unique y is thought of as $f(x)$. In a sense these are 'partial' functions. What we defined are 'total' functions. We stick to our definition.

f is one-to-one, or $1-1$ or injective if two different points have different images. That is

$$x_1 \neq x_2 \Rightarrow f(x_1) \neq f(x_2)$$

f is onto or surjective if image is all of Y. That is

$$\forall y \exists x \; (x, y) \in R.$$

f is bijective if it is both one-one and onto.

Since you have seen several examples of functions, there is no need for me to give examples here in a text book style.

Cardinality:

The number of elements or cardinality of empty set is 0; $|\emptyset| = 0$. If X is a non-empty set and there is a bijection from X to $\{1, 2, \cdots, n\}$ then we say that X has n elements; $|X| = n$

For this definition to make sense you should first convince yourself that the following can not happen: X has bijection with $\{1, 2, \cdots, 25\}$ and also X has a bijection with $\{1, 2, \cdots, 33\}$. Because if this happens then we would be saying $|X| = 25$ and also $|X| = 33$! Yes, such a thing as above can not happen. (Why?)

A set which is either empty set or is in bijection with $\{1, 2, \cdots, n\}$ for some $n \in N$ is said to be a finite set. A set which is not finite is said to be infinite.

If X is in bijection with N, we say that X is countably infinite. If X is either finite or countably infinite, we say that X is countable. If X is not countable then we say that X is uncountable.

Several properties of countable sets were discussed last year. We showed that the set of real numbers R is uncountable. You should recall those.

We have only associated 'numbers' with finite sets. There are numbers associated with infinite sets too. If X is countably infinite, we say $|X| = \aleph_0$ (aleph zero). We say $|R| = c$ (c stands for 'continuum').

How do you compare sets? The good old primitive way. Associate with each student the chair he/she is sitting. If some 'unassociated chairs' are leftover after associating chair to each student, then we say there are more chairs than students. If 'unassociated students' are leftover after associating a student to each chair, then we say there are more students than chairs.

Of course, you might think we have made so much progress, we can as well 'count' the number of students and number of chairs and then compare these numbers. But the point is you are using too many unnecessary words in this counting. Moreover this (so called) progress stops with finite sets and is useless while comparing infinite sets.

Let X and Y be any sets, finite or countable or uncountable; anything. We say Y has more elements than X (or X has less elements than Y); write $|X| \leq |Y|$; if there is an injection (one-one function) on X into Y. It is a peculiar situation. We have not defined 'number of elements' in X or Y. But we defined what it means to say X has less elements than Y.

Common sense suggests

$$|X| \leq |X|$$

$$|X| \leq |Y|; \quad |Y| \leq |Z| \Rightarrow |X| \leq |Z|$$

$$|X| \leq |Y| \quad |Y| \leq |X| \Rightarrow |X| = |Y|.$$

The first two statements above are easy to prove using composition of functions. The last statement is true but not trivial. It is known as Cantor-Bernstein-Schroder Theorem. We shall prove it soon.

But are there larger and larger sets? Yes, given any set X we can cook up a set Y such that $|X| < |Y|$. This means $|X| \leq |Y|$ and $|X| \neq |Y|$. Here is how.

Let X be any set. Then $P(X)$ denotes the collection of all subsets of X. There is an injection on X into $P(X)$, namely, associate with $x \in X$ the singleton set $\{x\}$ in $P(X)$. Thus $|X| \leq |P(X)|$. However there is no bijection. Indeed, if there is a bijection, say $f : X \to P(X)$ then let us define the following subset of X

$$S = \{x \in X : x \notin f(x).\}$$

Since S is clearly a subset of X, there must be a $s \in X$ such that $f(s) = S$.

Assume $s \in S$. The definition of the set S tells you $s \notin f(s)$. But $f(s) = S$. So $s \notin S$. contradiction to the assumption.

Assume $s \notin S$. The definition of S tells you $s \in f(s)$. But $f(s) = S$. Thus $s \in S$. contradiction to the assumption.

But s must be either in S or in S^c. The above argument says none of these happen. This completes the proof that there is no bijection of X with $P(X)$.

Theorem: $|X| < |P(X)|$.

The above theorem is due to Cantor and the argument is known as Cantor's diagonal argument. If you use a symbol $\aleph$ for $|X|$ then it is customary to denote $|P(X)| = 2^{\aleph}$.

Of course, we should make sure that such notations do not contradict existing notations. For example suppose X is finite and $|X| = 5$. Then the above notation tells us $|P(X)| = 2^5$. But we already have a meaning for 2^5 and it is 32. So we must convince ourselves that indeed $|P(X)| = 32$. Yes, you can do it.

Generally empty set needs special consideration (thought). In the above proof what happens if X is empty set? Same argument works. The only function from X to anything is $R = \emptyset$, simply because whatever be Y, we have $X \times Y = \emptyset$.

Recall $P(X) = \{\emptyset\}$. In other words $P(X)$ has one element (namely, empty set). Obviously, if you take this element as y, there is no x such that $(x, y) \in R$. Handling empty set is a little troublesome. You are encouraged to think (and free to ask) but not encouraged to worry about empty set.

inverse and forward images:

Let $f : X \to Y$. For $B \subset Y$, we define inverse image of B under f as follows.

$$f^{-1}(B) = \{x \in X : f(x) \in B\}.$$

If you think of function as $R \subset X \times Y$,then this simply means

$$f^{-1}(B) = \{x \in X : \exists y (y \in B \quad \wedge \quad (x, y) \in R)\}.$$

Of course, there is no need for you to keep thinking function as subset of product space.

Simalarly, for $A \subset X$ we define its forward image under f by

$$f(A) = \{f(x) : x \in A\}.$$

Here is a useful property of inverse images.

Theorem:

$$f^{-1}(B_1 \cup B_2) = f^{-1}(B_1) \cup f^{-1}(B_2)$$

$$f^{-1}(B_1 \cap B_2) = f^{-1}(B_1) \cap f^{-1}(B_2).$$
$$f^{-1}(B^c) = [f^{-1}(B)]^c.$$
$$f(A_1 \cup A_2) = f(A_1) \cup f(A_2).$$
In general, $f(A_1 \cap A_2) \neq f(A_1) \cap f(A_2)$.

You should be able to prove this.

posets and losets:

Let X be a set (not empty). A relation on X is simply a subset of $X \times X$. These are actually called binary relations, but we need not bother. We call them relations.

We say a relation is a partial order if

$$\forall x \quad (x, x) \in R$$

$$\forall x, y, z \quad (x, y) \in R \;\; (y, z) \in R \Rightarrow (x, z) \in R$$
$$\forall x, y, \quad (x, y) \in R \;\; (y, x) \in R \Rightarrow x = y.$$
If the first statement holds, we say that the relation is reflexive. If the second statement holds we say that the relation is transitive. If the third statement holds we say that the relation is anti-symmetric. But you need not remember these technical words at this moment.

If xRy holds, we say x and y are comparable and write $x \leq y$. Thus with this new notation we have (i) $x \leq x$; (ii) $x \leq y$ and $y \leq z$ implies $x \leq z$ and finally (iii) $x \leq y$ and $y \leq x$ implies $x = y$.

In some books, a different definition is used. They, in essence, define $<$ instead of $\leq$. Thus you may see the following definition: a relation R' is a partial order if it satisfies
(i) $\forall\, x, y, z \quad (x, y) \in R', (y, z) \in R' \Rightarrow (x, z) \in R'$
(ii) $\forall\, x \quad (x, x) \notin R'$.

Both these definitions are equivalent. If there is R' as above then first observe that $(x, y) \in R'$ and $(y, x) \in R'$ can not hold simultaneously for any two points x and y. If they both did, then condition (i) implies $(x, x) \in R'$ contradicting (ii). Define

$$(x, y) \in R \longleftrightarrow x = y \quad \vee \quad (x, y) \in R'$$

Then this relation obeys our three rules.

Conversely, suppose we have a binary relation R satisfying our rules. If we define R' by

$$(x, y) \in R' \longleftrightarrow x \neq y \quad \wedge \quad (x, y) \in R$$

Then this satisfies the two conditions of the alternative definition. We follow our definition, thus keep in mind, we are talking of $\leq$.

If the following condition also holds then the partial order is called a linear order.

$$\forall x, y \quad x \leq y \quad \vee \quad y \leq x.$$

Thus any two elements are comparable.

A set with a partial order is called a partially ordered set or simply a poset. A set with a linear order is called a linearly ordered set or simply a loset.

The set of all subsets of the real line with $A \leq B$ if $A \subset B$ is a poset but not a loset.

The following sets are losets. I give only sets, the order is the usual order coming from real numbers.

(i) $X = R$
(ii) $X = [0, 1]$
(iii) $X = [0, 1] \ \cup \ [2, 3]$
(iv) $X = [0, 1) \ \cup (2, 3]$.
(v) $X = [0, 1) \ \cup [2, 3]$
(vi) $X = Q$ the set of rationals .

You might wonder why I am giving these examples rather than saying: take any subset of R. Each set above illustrates a property.

In the first example, there is no first element (something smaller than every other thing is first element) and it has no last element (some thing which is larger than every other thing is called last element). The second set has both of these.

In the first two sets given any two different elements, there is an element strictly in between them. The third set does not have this property. For example between 1 and 2 which are in this set, there is none in between.

In all these three examples, every subset which is bounded above has a supremum. In the fourth example the subset $[0, 1)$ is bounded above but it has no sup. Just to make clear the terms we used, recall from calculus course, $A \subset X$ is bounded if there is an $s \in X$ such that $a \leq s$ holds for all $a \in A$. Such an element s is called an upper bound. An element s is supremum of A if it is an upper bound and it is smallest, that is, if t is an upper bound of A then $s \leq t$.

The fifth example appears deceptively similar to the fourth example. It appears that the subset $[0, 1)$ has no sup, In fact, this subset has sup.

The sixth set is countable and also has interesting properties which we shall discuss next.

Theorem (Cantor's Characterization of Q): Let $(X, \leq)$ be a countable linearly ordered set (non-empty) which has no first element; no last element; between any two distinct elements there is another element.

Then X is isomorphic to Q, that is, there is a bijection $f : X \to Q$ such that $x \leq y$ iff $f(x) \leq f(y)$.

You should note that we are denoting partial order by $\leq$ both on X and Q.

Proof:

We start with an observation. Let S and T be two finite subsets of X with $S < T$. This means $s < t$ for every $s \in S$ and every $t \in T$. Then there is $a \in X$ with $S < a < T$. Since S and T are finite sets, take maximum of S and minimum of T and pick an element between these. Note that finiteness of the sets is important.

In case one of the sets is empty, the hypothesis that there are no end points makes this possible. For example if T is empty, then you take max of S, say s^* and using the fact that there is no last element, take a such that $s^* < a$. This will do because you need not satisfy anything w.r.t. T. (If both S and T are empty?, we do not need this, but).

The technique used below is called *back and forth* argument. First let us fix an enumeration.

$$X = \{x_1, x_2, x_3, \cdots\}; \quad Q = \{q_1, q_2, q_3, \cdots\}.$$

We shall re-enumerate the sets

$$X = \{a_1, a_2, a_3, \cdots\}, \quad Q = \{b_1, b_2, b_3, \cdots\},$$

in such a way that the map $f(a_i) = b_i$ is the required isomorphism. This will be so if we make sure that for each k

$$f(a_i) = b_i \text{ is order preserving on } \{a_1, \cdots, a_k\} \text{ onto } \{b_1, \cdots, b_k\}. \qquad (*)$$

After all the fact that we have enumerated the sets will tell you that the map is a bijection. Any instance of verification that the map preserves the order depends on just two elements.

Step1: Put $a_1 = x_1$ and $b_1 = q_1$.

Step 2: Put $b_2 = q_2$. If $b_2 < b_1$ consider the first i such that $x_i < x_1$ and declare this x_i as a_2. That there are no end points makes this possible. If $b_1 < b_2$ consider the first i such that $x_1 < x_i$ and declare this x_i as a_2.

Note that $\{a_1, a_2\}$; $\{b_1, b_2\}$ is an order preserving listing, that is, $f(a_i) = b_i$ is order preserving.

Step 3: Put a_3 to be the first unused x_i. Thus if the x_i we have chosen in step 2 is x_2 then $a_3 = x_3$ and if the x_i chosen in step 2 is not x_2, then $a_3 = x_2$. Let

$$S = \{b_i : i \leq 2, a_i < a_3\}; \quad T = \{b_i : i \leq 2, a_3 < a_i\}.$$

Choose the first unused q_i such that $S < q_i < T$ and set this q_i as b_3. This is possible by the observation made at the beginning.

Note that $\{a_1, a_2, a_3\}$, $\{b_1, b_2, b_3\}$ is an order preserving listing.

In general, if we have listings $\{a_1, a_2, \cdots, a_{2k}\}$ and $\{b_1, b_2, \cdots, b_{2k}\}$; order preserving, then we proceed as follows.

step $(2k+1)$: Put a_{2k+1} to be the first unused x_i. Let

$$S = \{b_i : i \leq 2k, a_i < a_{2k+1}\}; \quad T = \{b_i : i \leq 2k, a_{2k+1} < a_i\}.$$

Note that if $b_i \in S$ and $b_j \in T$, then $a_i < a_{2k+1} < a_j$ so that $b_i < b_j$ — remember the existing listing is order preserving. Choose the first unused q_i such that $S < q_i < T$ and set this q_i as b_{2k+1}. Note that the listing $\{a_1, a_2, \cdots, a_{2k}, a_{2k+1}\}$ and $\{b_1, b_2, \cdots, b_{2k}, b_{2k+1}\}$ is order preserving.

step $(2k+2)$: Put b_{2k+2} to be the first unused q_i. Let

$$S = \{a_i : i \leq 2k+1, b_i < b_{2k+2}\}; \quad T = \{a_i : i \leq 2k+1, b_{2k+2} < b_i\}.$$

As earlier if $a_i \in S$ and $a_j \in T$, then $a_i < a_j$. Choose the first unused x_i such that $S < x_i < T$ and set this x_i as a_{2k+2}.

The way we have listed, *all* of X is listed as a's and all of Q is listed as b's. In fact, x_1 appears at step 1; x_2 appears at least by step 3; x_3 appears at least by step 5 etc. Pause and think if you can explain 'etc' . Also at each stage (*) holds. This completes the proof.

The proof actually shows something more. Let X be a set as in the theorem with a linear order. Given any two elements $x \in X$ and $q \in Q$ we can get an order preserving bijection f so that $f(x) = q$. In particular, such an order preserving bijection is not unique. You have some freedom.

Theorem 2 (countable dense subsets of R): Let A and B be two countable dense subsets of R. Then we can get a bijection φ of R to itself so that both φ and its inverse are continuous which moreover sends A to B. More precisely, there is $\varphi : R \to R$ such that $\varphi(A) = B$; φ is a bijection, φ is continuous; φ^{-1} is continuous.

A function $\varphi : R \to R$ which is a bijection; continuous, inverse is also continuous is called a homeomorphism of R with itself. In other words, it is an isomorphism for the 'concept of continuity'.

Proof: First observe that A and also B are sets satisfying the hypotheses of Theorem 1. So fix an isomorphism $f : A \to B$. Define, for each x and y,

$$A_x = \{a \in A : a < x\}; \quad B_y = \{b \in B : b < y\}.$$

(i) For each x, $A_x \neq \emptyset$. This is because A being dense, there are points in a with $a < x$. Also $a \in A_x$ $a' \in A$ and $a' < a$ imply that $a' < x$ so that $a' \in A_x$. Further, $\sup A_x = x$. This is because if $x' < x$, then by denseness of A, there are points a with $x' < a < x$. Thus nothing smaller than x is an upper bound of A_x. Of course x is an upper bound and hence it is $\sup A_x$. Finally, $x \notin A_x$.
Similar statements holds for B_y.

Define

$$\varphi(x) = \sup f(A_x) = \sup\{f(a) : a \in A_x\} = \sup\{f(a) : a < x\}.$$

This supremum is sensible because the set under consideration is non-empty from (i). Also A being dense, we can get $\alpha \in A$ with $x < \alpha$ and clearly $f(\alpha)$ is an upper bound for the set under consideration. We have used the fact that every non-empty set of reals which is bounded above has a supremum.

(ii) φ is a strictly increasing function.
Indeed if $x < x'$ then by successively using denseness of A, we get points $a_1, a_2 \in A$ such that $x < a_1 < a_2 < x'$. Thus every point of $f(A_x)$ is smaller than $f(a_1)$ so that $\varphi(x) \leq f(a_1) < f(a_2) \leq \varphi(x')$.

(iii) If $x \in A$, then $\varphi(x) = f(x)$.
Indeed, for every $a \in A_x$, we have $a < x$ so that $f(a) < f(x)$. Hence $\varphi(x) \leq f(x)$. If $y < f(x)$, we can choose $b \in B$ with $y < b < f(x)$. Then $a = f^{-1}(b) \in A_x$ and $y < f(a)$. Thus nothing smaller than $f(x)$ is an upper bound of $f(A_x)$, So $\varphi(x) = f(x)$.

(iv) Given any number y, there is an x such that $\varphi(x) = y$.
Indeed put $x = \sup f^{-1}(B_y) = \sup\{f^{-1}(b) : b < y\}$. By (i) B_y and hence $f^{-1}(B_y)$ is non-empty. It is bounded above because, B being dense there are points $b \in B$ with $y < b$ and $f^{-1}(b)$ is an upper bound for $f^{-1}(B_y)$. We show that $\varphi(x) = y$.

Need to show that $\sup f(A_x) = y$. If $a \in A_x$, then $a = f^{-1}(b)$ for some $b < y$ so that $f(a) = b < y$ and hence y is an upper bound for $f(A_x)$. If $y' < y$, then by denseness of B get $b \in B$, with $y' < b < y$. Then, definition of the point x tells that, $a = f^{-1}(b) \in A_x$ and $y' < b = f(a)$ showing that anything smaller than y is not an upper bound for $f(A_x)$.

Thus φ is a strictly increasing map of R in view of (ii). It is onto R in view of (iv). And $f(A) = B$ in view of (iii). This is the required map.

Just note that any increasing bijection φ of R is continuous and its inverse is also continuous. Being bijection it is first of all strictly increasing. For reals $a < b$, we have

$$\varphi^{-1}(a, b) = (\varphi^{-1}(a), \varphi^{-1}(b)); \quad \varphi(a, b) = (\varphi(a), \varphi(b)).$$

These equalities are enough to show required continuity.

For example, there is a homeomorphism of R to itself which transports set of rationals to the set of algebraic numbers. There is a home that transports the set of rationals to the set of non-rational algebraic numbers.

We can use the Cantor's theorem to setup homeomorphism as above taking one Cantor set to another. We shall do this later.

Cantor-Shroder-Bernstein Theorem:

Let X and Y be sets. Let $f : X \to Y$ be an injection and $g : Y \to X$ be an injection. Then there is a bijection $\varphi : X \to Y$. In other words, if $|X| \leq |Y|$ and $|Y| \leq |X|$ then $|X| = |Y|$.

The original proof of Cantor used well-ordering principle. After several refinements, here is a proof.

Let us think of $f(x)$ as child of x. Thus if $f(x) = y$, then y is child of x or equivalently x is immediate ancestor of y. Similarly, $g(y) = x$ would mean child of y is x or equivalently y is immediate ancestor of x. Since f and g are injective an immediate ancestor, if exists, is unique.

Suppose that $x \in X$ has an immediate ancestor y_1 and y_1 has an immediate ancestor x_2 and x_2 has an immediate ancestor y_3 and so on. In other words, we start with $x \in X$;

$$g(y_1) = x; \quad f(x_2) = y_1; \quad g(y_3) = x_2; \cdots .$$

This is the ancestral chain for x. It is quite possible that the ancestral chain of x continues for ever or the ancestral chain of x stood at a finite stage. Again there are two possibilities in the later case, x may have an even number $(0, 2, 4, \cdots)$ of ancestors or an odd number $(1, 3, 5, \cdots)$ of ancestors. Similarly if you start with $y \in Y$ its ancestral chain may continue for ever or may stop with even number of ancestors or add number. Accordingly we partition the sets

$$X = X_\infty \cup X_e \cup X_o; \quad Y = Y_\infty \cup Y_o \cup Y_e.$$

We define $\varphi(x) = f(x)$ in case $x \in X_\infty$ or $x \in X_e$. And $\varphi(x) = g^{-1}(x)$ if $x \in X_o$. This last clause makes sense because x has an immediate ancestor.

We show that this does.

We show that φ maps in a bijective way X_∞ to Y_∞; X_e to Y_o; X_o to Y_e. This shows that φ is globally bijective too — since we have partitions of the sets X and Y. To realise there is something here, remember the x^2 map is injective on $(-\infty, 0]$; injective on $[0, \infty)$ but not injective on $(-\infty, \infty)$.

Note that if the ancestral chain of x infinite then so is that of $f(x)$. Indeed the ancestral chain of $f(x)$ has x and all its ancestors. Hence $f(x) \in Y_\infty$. Conversely if $y\infty$, it has an immediate ancestor x. Clearly $x \in X_\infty$ and $f(x) = y$ showing that φ — which is same as f — is a bijection between X_∞ and Y_∞.

If $x \in X_e$ then x has an even number of ancestors. Since ancestors of $f(x)$ consists of x and ancestors of x we see $f(x) \in Y_o$. conversely if $y \in Y_o$, then y has an immediate ancestor, say x. The ancestors of x consists of ancestors of y except x itself. Thus $x \in X_e$ and of course $f(x) = y$. Thus φ, which is same as f is again a bijection between X_e and Y_o.

Similar argument as above shows that $x \in X_o$ implies that $g^{-1}(x) \in Y_e$. Conversely if $y \in Y_e$ then $g(y) \in X_o$ and $\varphi(g(y)) = y$; showing that g^{-1} is a bijection between X_o and Y_e.

This completes the proof.

The proof is tricky, proof using well-ordering principle is straightforward. However to get a feel for the proof, it is better to look at instructive examples.

Let us consider the most trivial example.

$$X = Y = \{0, 1, 2, 3, \cdots\}; \quad f(x) = 2x; \quad g(y) = 2y.$$

Let us first consider ancestors of points in X. Only zero has infinitely many ancestors. Elements that have zero ancestors consists precisely the set O of odd numbers. Elements that have exactly one ancestor constitutes the set $2 \times O = \{2 \times 1, 2 \times 3, 2 \times 5, \cdots\}$.

The set $2^2 \times O$ is the set of points that have exactly two ancestors. In general $2^k \times O$ is the set of points that have exactly k ancestors.

The map φ exchanges O with $2 \times O$; exchanges $2^2 \times O$ with $2^3 \times O$ etc in the obvious way. Of course, $\varphi(0) = 0$.

Axiom of Choice:

There are several equivalent formulations of this axiom. Without converting this course into a course in Set Theory, here are some.

(1) Given any (nonempty) family $\mathcal{A}$ of disjoint non-empty sets, there is a set S such that $S \cap A$ is singleton for all $A \in \mathcal{A}$. In other words, you can make a set picking exactly one point from each of the given sets.

(2) Given any family of non-empty sets $\mathcal{A}$, there is a function f with domain $\mathcal{A}$ such that $f(A) \in A$ for each $A \in \mathcal{A}$. That is, given any family of nonempty sets we can associate with each set in that family one point from that set.

Before stating the next one, we need a definition. Let P be a poset. A set $C \subset P$ is called a *chain* if given any two elements $p, q \in C$ either $p < q$ or

$p = q$ or $q < p$. In other words $(C, <)$ is a loset. An *upper bound* for a chain C means an element $s \in P$ such that for each $p \in C$ either $p < s$ or $p = s$. In other words, it is just an upper bound in the sense we have defined for subsets of losets. The only difference is that, now we have a poset and the subset we are talking about is linearly ordered. Note that the upper bound itself need not belong to the chain, it belongs to the poset.

An element m of the poset is said to be a *maximal element* if $\neg(m < p)$ for all $p \in P$. Observe that we are not saying that $p \leq m$ for each p. We are only saying that there is nothing larger than m. After all, some elements p may not be 'comparable' with m.

For instance, let P be the collection of subsets of R having at most four elements, with usual inclusion order. The elements $\{1, 2, 3, 4, \}$, $\{8, \sqrt{2}, 3/4, 49\}$ are maximal elements in P. In fact any four element set is a maximal element. Here is a chain:
$\{1\}, \{1, 2, 3, \}$.
Here is another chain:
$\emptyset, \{3\}, \{3, 49\}, \{3, 49, 99\}, \{3, 49, 99, -31\}$.

The collection of all finite subsets of R is another poset. It has no maximal elements. Here is a chain in this poset.
$\emptyset, \{1\}, \{1, 2\}, \{1, 2, 3, \}, \{1, 2, 3, 4, \}, \cdots\cdots\cdots$.
This chain has no upper bound, because then that set must include all these points and hence can not be finite.

You can also consider the collection of countable subsets of R.

Here is an equivalent of the Axiom of choice.

(3) Let $(P, <)$ be a (non-empty) poset. Suppose that any chain $C \subset P$ has an upper bound. Then P has a maximal element.

Here is another equivalent of the axiom of choice. We say that a linear order is a well order if every non-empty subset has a first element. In other words, let $(X, \leq)$ be a loset. it is called a well-ordered set, simply woset, if $A \subset X$ and $A \neq \emptyset$ implies there is an element $a \in A$ such that $a \leq x$ for all $x \in A$. Note that $a \in A$ and also note that such an element (if exists) is unique. Here is another form of the axiom of choice.

(4) Every non-empty set can be well-ordered, that is, given a non-empty

set X, there is a binary relation $\leq$ on X so that $(X, \leq)$ is a woset.

Such a set as in (1) is called choice set, a function as in (2) is called a choice function. The statement (1) or (2) is called axiom of choice. The statement (3) is called Zorn's lemma. It was actually (in an equivalent form) proposed by Hausdorff in 1914, Kuratowski in 1922 but popularized by Zorn in 1935. Statement (4) is called the 'well-ordering principle'.

Here is an application of the Zorn's lemma which you already know.

Every vector space has a basis, that is, linearly independent set B which is maximal — if you put in another vector in B then the resulting set is no longer independent. (equivalently, every non-zero vector v can be uniquely expressed as a *finite sum* $v = \sum c_i v_i$ such that (i) the vectors v_i are distinct; (ii) $v_i \in B$ for each i and (iii) $c_i \neq 0$ for each i).

Consider the collection $\mathcal{P}$ of all independent subsets of the vector space. Order this collection by saying $A \leq B$ if $A \subset B$ for $A, B \in \mathcal{P}$. This collection is non-empty because singleton set consisting of a non-zero vector is an independent set. well, if your vector space consists of only zero and nothing else, then take B to be empty set, verify this does and end the proof.

This is a poset (Any collection of sets with inclusion as order is a poset). If you take a chain $\mathcal{C} \subset \mathcal{P}$, then union S of all sets in this chain is an upper bound. It obviously includes all sets in the chain and hence is an upper bound, if only we verify that this $S \in \mathcal{P}$. If you take finitely many vectors in this set S, then each of these vectors is in one set in the chain and hence all these vectors in *one* set (given finitely many elements of a loset, there must be one of them which is larger than all others). In other words these finitely many vectors belong to one independent set. So they are independent.

You should get feeling for the feature 'independence is a finitary property'. that is, a set of vectors is independent iff every finite subset is independent.

I have not given any references. You can consult any book on set theory and read the topics we are discussing like Halmos, Naive Set Theory. One of the best sources is the book Real Analysis by Hewitt and Stromberg. Generally most of the books contain much more (or much less) than what I intend to do. That is why I have not given any references so far. But please do consult any book you like from the library.

Axiom of Choice is abbreviated to AC and last time we introduced four statements.

The first one is existence of choice set: Given any family of nonempty sets which are pairwise disjoint, you can make a set which contains exactly one point from of each of the given sets. you would say 'is it not obviously true? why make an axiom?' You do not think an axiom is necessary at all. Let us name the second statement as existence of choice function; given a family of (non-empty) sets there is a function which associates with each set in the family one point of that set.

The content of the statement is brought out very beautifully by Bertrand Russel: Given lots of pairs of shoes you can provide an algorithm to select one shoe from each pair; but given lots of pairs of new socks you have no algorithm to select one from each pair.

Let us name the fourth statement: every set can be well-ordered as well-ordering principle. This would in particular imply that the set of real numbers can be well ordered. That is, you can prescribe an linear order on R such that every non-empty subset has a first point. You would probably say 'Is it not obviously false? why discuss this statement?'. You believe that this statement is blatantly false. The third statement: partially ordered set has maximal element if every chain has upper bound; is called Zorn's lemma (due, in another form, to Hausdorff and later Kuratowski, but popularised by Max Zorn).

However all are just the same.

Theorem: (1) existence of choice set (2) existence of choice function (3) Zorn's Lemma and (4) Well-ordering principle; are equivalent statements.

We have seen already an application of Zorn, to show existence of Basis for any vector space. You have learnt (finite dimensional vector spaces) in the first semester. Second semester you learnt groups. If you travel a little more on one of its roads, you will reach the Neilsen-Schreier theorem which says that every subgroup of a free-group is itself a free-group. This is proved using AC.

In this third semester you are learning rings. The fact that a non-trivial unital ring has non-trivial maximal ideals needs AC. Next semester you will learn fields. You will learn that very field can be put inside a bigger field F which is algebraically closed. This last phrase means every nontrivial polynomial in one variable x with coefficients coming from F will assume the value zero for at least one value of $x \in F$. This is achieved with AC.

So much for algebra. There are several applications in Analysis which is our interest. They will unfold in your later courses. We shall see one or two. But first few useful things are to be observed.

(1) Countable union of countable sets is countable.

Of course we have already done this rather quietly. For a moment assume that all the sets are infinite. You see that if $(A_n : n \geq 1)$ are the sets, then by definition of 'countably infinite', there exists a bijection between A_n and N. But what you need is to pick *one* bijection for *each* n (so that you can produce a map of the union to N). AC allows you this.

(2) For any two sets X and Y, either $|X| \leq |Y|$ or $|Y| \leq |X|$.

In other words there is an injection from X to Y or from Y to X. equivalently, any two sets can be compared. To prove this, let P be the set of all pairs (S, f) where $S \subset X$ and f is an injection on S to Y. This is non-empty because you can take any $x \in X$ and any $y \in Y$; put $S = \{x\}$ and $f(x) = y$. Then $(S, f) \in P$. Well, if one of them is empty you need not prove any thing (why?), so assume both are non-empty sets.

Define $(S, f) \leq (T, g)$ if $S \subset T$ and g extends f. The last phrase means that for $x \in S$ we have $f(x) = g(x)$. This is a partial order and so P is a poset. Every chain has an upper bound. Indeed if $\{(S_\alpha, f_\alpha) : \alpha \in \Delta\}$ is a chain then here is upper bound: take S as union of all the sets S_α. For $x \in S$, put $f(x) = f_\alpha(x)$ in case $x \in S_\alpha$. This is a good definition because if x is in two of the sets then the fact that we have a chain implies that you get the same values whichever f_α you use.

So by Zorn, get a maximal element (T, g). In case $T = X$ then you got an injection on X to Y. In case range g is all of Y, then g^{-1} provides an injection on Y to X. If neither happens, then you pick $x^* \in X - T$ and $y^* \in Y -$ range g. Put $T_1 = T \cup \{x^*\}$; put $g_1(x) = g(x)$ for $x \in T$ and $g_1(x^*) = y^*$. You see that $(T, g) < (T_1, g_1)$ contradicting maximality of (T, g).

This completes proof.

(3) Every infinite set contains a copy of N, that is, contains a countably infinite set.

Fix a choice function f for subsets of X, the given infinite set. This means for each non-empty subset $S \subset X$ we have $f(S) \in S$. Define

$$x_1 = f(X)$$

$$x_2 = f(X - \{x_1\})$$

$$x_3 = f(X - \{x_1, x_2\})$$

and in general

$$x_{k+1} = f(X - \{x_1, \cdots, x_k\}).$$

The set $(x_i : i \geq 1)$ does the job. Carefully note that these are distinct elements.

(4) If X is an infinite set and $F \subset X$ is a finite set, then $|X| = |X - F|$. That is, X and $X - F$ have the same cardinality.

Denote your finite set by

$$F = \{x_1, \cdots x_k\}.$$

Fix a choice function as above. Put

$$x_{k+1} = f(X - \{x_1, \cdots, x_k\})$$

$$x_{k+2} = f(X - \{x_1, \cdots, x_k, x_{k+1}\})$$

and in general for $n \in N$

$$x_{k+n} = f(X - \{x_1, \cdots, x_{k+n-1}\})$$

Here is a bijection on X onto $X - F$. If $x \notin (x_i : i \geq 1)$ put $h(x) = x$. Put $h(x_i) = x_{i+k}$ for $i = 1, 2, \cdots$. This does the job.

You only need to note that X is partitioned into two sets $A_1 = \{x_1, x_2, \cdots\}$ and $A_2 = X - A_1$. Similarly $X - F$ is partitioned into two sets $B_1 = \{x_{k+1}, x_{k+2}, \cdots\}$ and $B_2 = X - A_1$ (watch out, not $X - B_1$). The map h is an bijection of A_i and B_i.

(5) If X is an infinite set then $|X \times \{0,1\}| = |X|$.

That is, two copies of X has the same potency as one copy of X. Remember $X \times \{0,1\}$ consists of all pairs $(x, 0)$ for $x \in X$ and also all pairs $(x, 1)$ for $x \in X$.

To prove the statement, let P consist of all pairs (S, f) where $S \subset X$ and f is a bijection on $S \times \{0,1\}$ to S. This is non-empty. Indeed using (3) you get S which is a copy of N. Since $S \times \{0,1\}$ is also countably infinite there is a bijection f of $S \times \{0,1\}$ to S. This (S, f) is in P.

Define $(S, f) \leq (T, g)$ if $S \subset T$ and g extends f. This makes sense because $S \times \{0,1\} \subset T \times \{0,1\}$. This makes P a poset.

Every chain has an upper bound. If (S_α, f_α) is a chain then take S as the union of these S_α. If $(x, i) \in S \times \{0,1\}$ put $h(x, i) = f_\alpha(x, i)$ in case $x \in S_\alpha$. again the fact that we have a chain tells us there is no conflict in case x is in two of the sets S_α. It is a bijection too. If you take two distinct points (x, i) and (y, j) then there is one S_α in which both x and y are present etc.

Let (T, g) be a maximal element, exists by Zorn. Can $X - T$ be infinite? No because of the following reason. If it were infinite, use (3) and get a copy of N in $X - T$, denote it by A. Since $A \times \{0,1\}$ is a countably infinite set, fix a bijection φ on $A \times \{0,1\}$ onto A.

Take $T_1 = T \cup A$. Define g_1 on $T_1 \times \{0,1\}$ as follows. If $x \in T$, then $g_1(x, i) = g(x, i)$ and if $x \in A$, then $g_1(x, i) = \varphi(x, i)$. Note that φ takes you to A whereas g takes you to T. This helps you to show that g_1 is bijection. But then $(T, g) < (T_1, g_1)$ contradicting maximality of (T, g).

Thus $X - T$ is finite and hence by (4), T has the same cardinality as X. Now

$$X \times \{0,1\} \sim T \times \{0,1\} \sim T \sim X.$$

(6) If S, T are disjoint infinite sets, $|S| < |X|$ and $|T| < |X|$ then $|S \cup T| < |X|$.

recall $|S| < |X|$ means there is an injection from S to X but there is no bijection.

In view of (2) either $|S| \leq |T|$ or $|T| \leq |S|$, there is no loss in assuming that $|S| \leq |T|$. So there is an injection $f : S \to T$. Define g from $S \cup T$ to $T \times \{0, 1\}$ by $g(x) = (f(x), 0)$ if $x \in S$ and $g(x) = (x, 1)$ if $x \in T$. Note that ranges of S and T under g are disjoint. Easy to see that g is an injection showing

$$|S \cup T| \leq |T \times \{0, 1\}| = |T| < |X|.$$

where we used (5) for the equality. The first inequality is witnessed by g and the last inequality is hypothesis. You only need to note that if $|A| = |B|$ and $|B| < |C|$ then $|A| < |C|$.

(7) if X is infinite then $X \times X \sim X$.

Let P be the set of all pairs (S, f) where $S \subset X$ and f is a bijection on $S \times S$ onto S. This is partially ordered by inclusion. That is $(S_1, f_1) \leq (S_2, f_2)$ if $S_1 \subset S_2$ and f_2 is an extension of f_1. This means $f_2(x, y) = f_1(x, y)$ for $x, y \in S_1$. This makes P a poset.

P is non-empty because from (3) you can take a countably infinite set $S \subset X$ and use the fact that $S \times S$ is also countably infinite to get a bijection $f : S \times S \to S$. Then $(S, f) \in P$.

Every chain has upper bound. Indeed if (S_α, f_α) is a chain, here is the upper bound: S is the union of all the S_α. For $x, y \in S$, if they both belong to S_α we put $f(x, y) = f_\alpha(x, y)$. Since we have a chain, there is one α such that both x and y are in S_α. Since we have a chain, this definition does not depend on which α we take. That this is a bijection is routine. if you take two pairs (x_1, y_1) and (x_2, y_2) then there is one S_α in which all these four points x_1, x_2, y_1, y_2 are available and f_α is a one-one map on this $S_\alpha \times S_\alpha$ so that the f_α values of these two pairs are different, but these are f values as well. Also given any $a \in S$, get an α such that $a \in S_\alpha$ and hence there is an $(x, y) \in S_\alpha \times S_\alpha$ such that $f_\alpha(x, y) = a$ so that $f(x, y) = a$.

Take a maximal element (T, g). Since $T \subset X$ there are only two possibilities, either $|T| = |X|$ or $|T| < |X|$ — the identity map shows that $|T| \leq |X|$.

If $|T| = |X|$ we are done because

$$|X \times X| = |T \times T| = |T| = |X|.$$

Assume $|T| < |X|$. we show contradiction for maximality by producing (T_1, g_1) such that $(T, g) < (T_1, g_1)$.

Towards this end we first note that $X - T$ can not be finite. If it were, then X and T differ by finitely many points so that (4) tells $|X| = |T|$ which is not the case now.

Since $T \times T \sim T$ we conclude that T is infinite; we saw that $X - T$ is infinite. We are assuming that $|T| < |X|$. If we also have $|X - T| < |X|$ then (6) leads to contradiction $|X| < |X|$. Thus we must have $X - T \sim X$.

We are now ready to contradict maximality of (T, g). Since $X - T$ and X are of the same cardinality, pick a subset $S \subset X - T$ with $|S| = |T|$. You only need to fix a bijection φ on X onto $X - T$ and take $S = \varphi(T)$. Let $T_1 = T \cup S$. We shall now define a bijection g_1 on $T_1 \times T_1$ onto T_1 which extends g. Let us make two observations.

Firstly, $T_1 \times T_1$ is disjoint union of the four sets $A_0 = T \times T$; $A_1 = T \times S$; $A_2 = S \times T$; $A_3 = S \times S$. And also these four sets have the same cardinality simply because S and T have the same cardinality. Since $T \times T \sim T$ (Remember g) we conclude that all these four sets are equipotent with T and also with S.

Second observation is the following. We can express S as disjoint union of three sets S_1, S_2, S_3 all having the same cardinality as S. This is seen as follows.

We know that T is infinite. So take two points s and t from T and denote $T_0 = T - \{s, t\}$. consider the sets $A = \{s\} \times T$; $B = \{t\} \times T$; $C = T_0 \times T$. You see that A has same power as T; B has same power as T; T_0 which differs from T by a finite set has same power as T and hence C has the same power as $T \times T$ which has same power of T (Remember g). Thus we decomposed $T \times T$ into three disjoint sets each of power of T. But $T \times T$ has same power as T (Remember g). so we could decompose T into three sets of the power of T. But S and T are of the same power so we can decompose S into three sets each of the same power as S. Denote $S = B_1 \cup B_2 \cup B_3$.

To complete the proof define g_1 on $T_1 \times T_1$ as follows. On $A_0 = T \times T$ follow g to take you to T; on the other three sets A_1, A_2, A_3 which makeup the remaining part of $T_1 \times T_1$ use any maps taking you in a bijective way to B_1, B_2, B_3 respectively.

This completes the proof (by showing an element larger than the alleged maximal element and thereby establishing that T must have same cardinality as X and thereby $X \times X$ must have a bijection to X).

(8) Let X be an infinite set. Let $seq(X)$ denote the set of finite sequences of points from X. That is, things of the form $(x_1, x_2, \cdots x_k)$ where $k \geq 1$ is an integer and each $x_i \in X$. Then $seq(X) \sim X$.

Since $T \times T$ is of power T, we conclude, in particular, that a countable union $S = \cup T_i$ of disjoint sets $T_1, T_2, \cdots$ each of power T has power T. Indeed T being infinite, get by (3), an injection $f : \{1, 2, \cdots\} \to T$ and for each $i \geq 1$ get a bijection $f_i : T_i \to T$. Let now $s \in S$, then there is unique i such that $s \in T_i$. Put $g(s) = (f(i), f_i(s))$ gives a injection from S to $T \times T \sim T$.

Now X, $X \times X$, $X \times X \times X \cdots$ all have power of X and the earlier para tells you that their union is also of power X. But this union is precisely $seq(X)$.

Sometimes, for technical reasons one includes the empty sequence also in $seq(X)$. Empty sequence means sequence of length zero. There is only one such sequence, namely (). Even if you include this one element in the set $Seq(X)$ its power is still same as that of X.

Of course, some of you may have a psychological objection to include the empty sequence as a sequence at all. Do not worry, in that case, you do not have to do this. I only said: if sometime later somebody does some such thing you need not scratch your head. That is all.

(9) Let X be an infinite set. $Seq(Q) \times Seq(X) \sim X$.

Proof is already included in the above. for each fixed $a \in Seq(Q)$ the set of points in our set with first coordinate equal to a has power $Seq(X) \sim X$ and the number of possible a is countable.

(10) $f : X \to Y$ be surjection. Then $|Y| \leq |X|$.

Fix a choice function φ for non-empty subsets of X. For each $y \in Y$ the set $A_y = f^{-1}(y)$ is non-empty because f is surjection; so it makes sense to define $g(y) = \varphi(A_y)$. Then g is a one-one map of Y to X and $|Y| \leq |X|$.

Now we shall return to Analysis.

We showed that every vector space has a basis. We made no fuss about the underlying field. So let us use that freedom. Consider R as a vector space over Q, the field of rational numbers. Thus we have the following.

(11) The vector space R over the field Q has a basis. That is, there is a set $B \subset R$ such that every $x \in R$ can be uniquely expressed as a finite sum $\sum q_i b_i$ where $b_i \in B$ are distinct and $q_i \in Q$. Here uniqueness is interpreted as: if there are two such sums representing x the b s with non-zero coefficients are same in both expressions and those non-zero coefficients also agree.

In other words you can cheat by taking a finite sum and adding to it a term: $(0 \times b)$ with $b \in B$ which was not already there in the sum.

Such a Basis is called Hamel basis. How large is it?

(12) Any Hamel basis has the same cardinality as that of R.

Define a map $f : seq(Q) \times seq(B) \to R$ by
$f(q_1, \cdots, q_k; b_1, b_2, \cdots, b_l) = 0$ if $k \neq l$ and $= \sum q_i b_i$ if $k = l$.
Since B is a basis, this map is onto R. Thus (10) and (9) imply

$$|R| \leq |seq(Q) \times seq(B)| = |B|$$

But obviously $|B| \leq |R|$ and thus $|B| = |R|$. In using (9) we implicitly assumed that B is infinite. In fact, it is uncountable because if it were countable then we note that $seq(Q) \times seq(B)$ is also countable showing that the first inequality above itself is a contradiction. Just remember that the set of real numbers is not countable.

(13) There is a function $f : R \to R$ such that $f(x + y) = f(x) + f(y)$ for all $x, y \in R$ which is not continuous. In other words it is a (additive) group homomorphism.

You can make the range of f to be any non-trivial Q-subspace of R.

You can make f a bijection, so that it is a group isomorphism.

Fix a Hamel basis B.

Take $v \in B$. Here is f. For any $x \in R$, $f(x)$ is the coefficient of v in the expression of x as a finite rational linear combination of elements of B.

That this is a Q-linear map and hence additive is general vector space result. This is not continuous because it takes only rational values and is not the zero function. (Pause and think)

Take any Q-subspace of R. Take a basis for the subspace, say, H. Clearly $|H| \le |B|$ and so fix a surjection $\varphi : B \to H \cup \{0\}$. if $x = \sum q_i b_i$ define, $f(x) = \sum q_i \varphi(b_i)$. Again the fact that f is additive is a genera vector space fact. It is not continuous because if $v \in B$ with $\varphi(v) = 0$ then f takes the value zero on the set of rational multiples of v, that is, on a dense set. Range is also the given subspace.

To get an additive isomorphism, you only need to take a bijection φ of B to itself and define f on R by: if $x = \sum q_i b_i$ then $f(x) = \sum q_i \varphi(b_i)$. Of course if you take the identity map as your bijection φ then f is identity too. Take four elements from the basis, x_1, y_1, x_2, y_2 such that $y_1/x_1 \neq y_2/x_2$ and take a bijection φ of B with $\varphi(x_i) = y_i$ for $i = 1, 2$. Note that the resulting f would map x_i to y_i. It is not continuous because for any continuous map $f(x)/x$ (?) is a constant.

(14) There is an (additive) group isomorphism from R to R^2.

Think of both R and R^2 as vector spaces over Q, take Hamel bases, observe that both have the same cardinality (namely c), set up a bijection between them and extend by Q-linearity. Exactly like the above.

There are several things you can say using AC. But we stop here. We shall make some comments about the equivalence of the four statements.

Choice set existence implies choice function existtence. Indeed, given any family of sets $(A_\alpha; \alpha \in \Delta)$ put

$$B_\alpha = \{(\alpha, x) : x \in A_\alpha\}; \qquad \alpha \in \Delta.$$

These are disjoint because every point in B_α is a pair and its first coordinate is α. Take a choice set S, thus for each α we have $f(\alpha) \in B_\alpha$. Define $f(\alpha)$ to be the second coordinate of the unique point in $S \cap B_\alpha$.

Choice function implies choice set. If $(A_\alpha : \alpha \in \Delta)$ are disjoint sets and f is a choice function, that is, $f(\alpha) \in A_\alpha$ for each α then we take $S =$ Range f. Then S is a choice set for the given family of disjoint sets.

Zorn implies choice set. In fact let $(A_\alpha : \alpha \in \Delta)$ be disjoint sets. Let

$$P = \{S \subset \cup A_\alpha : |S \cap A_\alpha| \leq 1 \quad \text{for each} \quad \alpha\}.$$

This is a poset by defining $S_1 \leq S_2$ if $S_1 \subset S_2$. It is non-empty because if you take one of the sets and a point x from that set then S consisting of this single point is in P. Every chain has upper bound, namely, union of sets in the chain. So let T be a maximal element. If there is an A_α such that S has no point of A_α, pick one point from this A_α and let T' be the set consisting of points in T along with this extra point. Then $T' \in P$ and contradicts maximality of T. Thus for each α, we have $|S \cap A_\alpha| = 1$ showing that S is a choice set.

Existence of choice function implies Zorn. Shall only outline but not carry out the full proof. Take a poset P where every chain has an upper bound. We need to exhibit a maximal element.

Let us say that a chain C is maximal if there is no element of P larger than every element of C. That is

$$\neg \ \exists p \in P \ \ \forall x \in C \ \ (x < p).$$

The idea is to show that there is a maximal chain. For this several methods are available. Simplest (not necessarily the best) method is the following. Fix a choice function for subsets of P. That is a function f which associates with each non-empty subset of P a point in that set. Consider

$$S_0 = P; \qquad p_0 = f(S_0);$$

$$S_1 = \{x : p_0 < x\} quad \quad p_1 = f(S_1)$$
$$S_2 = \{x : p_1 < x\}; \qquad p_2 = f(S_2)$$

and so on

$$S_\infty = \{x : p_n < x \ \text{ for all } \ n\}; \qquad p_\infty = f(S_\infty)$$
$$S_{\infty+1} = \{x : p_\infty < x\}; \qquad p_{\infty+1} = f(S_{\infty+1})$$

etc. The p s so collected form a chain because at any time we are selecting a point larger than what we already have. if at some stage the set S_α is empty then collect all the p got so far. That is a maximal chain. The fact that we have arrived at empty set signals that there is nothing larger than all the selected p s. And of course, at some stage you do arrive at empty set.

Now this maximal chain has a upper bound, let it be a. This is a maximal element of P. Indeed if there is an element b with $a < b$ we have $x \leq a < b$ for every $x \in C$ contradicting maximality of the chain.

The above procedure of selecting points can be formalized but needs some work we shall not undertake. It is not trivial. It is easy to say 'so on' but difficult to explain what is this 'so on'.

Well-ordering principle implies choice set. Indeed if (A_α) is a disjoint family of sets then well order their union, and pick the least (in that well order) lament of A_α for each α. This set of points so selected gives choice set. That choice set implies well ordering principle is carried out in a fashion similar to the above construction.

It is in this form of existence of maximal chains that Hausdorff and later, Kuratowski formulated. However Zorn's name got stuck because he popularised. Max Zorn himself does not like it to be called Zorn's lemma! It is too late to change things now.

There are several other beautiful results in set theory, but we need to return to Analysis. But before doing so few words on history. I have already mentioned that all this had origins in problems concerning Fourier series.

Unfortunately even in scientific investigations we have fundamentalism — after all we are human beings. Just as the famous Mach and other scientists opposed Boltzman's ideas; jut as the Church opposed Galileo's ideas; here too we have Georg Cantor being opposed from several quarters. the famous Kronecker and Poincare opposed rather very very vehemently. So did the Church. After all God is infinite and hence (!) infinity is God. How can the infinity that represents integers different from the infinity that represents real numbers? There can not be many Gods, there is only one God! Worse, you are saying given any God there is a bigger God! Definitely not acceptable.

It was left to Hilbert to say: No one can dislodge us from the Paradise created by Cantor. Paul Erdos used to say: Keep your brains open.

We shall now return to Analysis. One of our objectives is to construct the set of real numbers —- this means to show a set with some operations satisfying whatever we assumed last year. Of course, any construction work needs using cement, bricks, water and so on and more important, dirtying our hands.

After all, you can not construct something out of nothing. So the first question is what do we have to start with. To simplify life let us ask: how do you construct real numbers if I give you the set Q of rational numbers. It is indeed easy. After all we know enough to be able to say

$$x \in R \Rightarrow \sup\{q \in Q : q \leq x\} = x.$$

This identifies real numbers as sup of a set of rational numbers. the only nuisance above is that this set depends on x. We should be able to describe real numbers using *only* rationals. There is nothing but set of rationals before us. You can not pretend that you already have real numbers and you are only describing them using rationals.

Getting a clue from the above, let us put things differently. Every real number x, *cuts* Q into two non-empty parts: those rationals q that are not above it ($q \leq x$) and those rationals that are above it ($q > x$).

Also every cut determines a real number, namely, sup of the lower part of the cut.

So what is a cut and how do we construct a complex system of real numbers starting from Q. Remember you need to prescribe a set; need to prescribe addition and multiplication and order on your set; and then show you have a complete Archimedian ordered field.

Of course, you can also ask how do you construct rational numbers? We do so using the set Z of integers. So how do you construct Z, the set of integers. We do so using the set of natural numbers $N = \{1, 2, \cdots, \}$ with the only operation being the successor operation $Sn = n + 1$.

This brings us to: who gives you the set N and the operation S on this set which hopefully tells us 'adding one'.

You can get either terrified or excited on the attitude you take. But be assured you need not construct real numbers in the exam. then why are we doing this? Don't you want to reassure yourself that real number system does exist (in flesh and blood) and all the things you learnt in Calculus have meaning.

Understanding R:

If you go close to the real line you see that you can add and multiply numbers. But in a superficial look at the line, you only see what comes after what. We shall now isolate some order-theoretic properties of R, whose presence in a loset helps us to recognize that that loset is none other than R itself.

This understanding will familiarize with structure of R. Also motivates and tells us 'how to proceed' to construct R.

Let $(X, <)$ be a loset. A non-empty subset $S \subset X$ is said to be *bounded above* if there is an $a \in X$ such that for all $x \in S$ we have $x \leq a$. Such an element a is also called an *upper bound* for the set S. An element $s \in X$, if exists, is said to be *supremum* of the set S if it is an upper bound and for any upper bound b we have $a \leq b$. In other words, supremum is the least upper bound. If there is supremum, then it must be unique. Note that neither upper bound nor supremum need belong to the set S.

Similarly, a subset $S \subset X$ is said to be *bounded below* if there is an $b \in X$ such that for all $x \in S$ we have $b \leq x$. Such an element b is also called a *lower bound* for the set S. An element $b \in X$, if exists, is said to be *infimum* of the set S if it is a lower bound and for any lower bound c we have $c \leq b$. In other words, infimum is the greatest lower bound. If there is an infimum, then it must be unique.

A set is said to be *bounded* if it is bounded below as well as above. You should not be frightened with several concepts being thrown in. You should notice that these concepts are not new. You saw them in the context of R. We know what is meant by bounded, supremum, infimum etc for subsets of real line. It just so happens that these are meaningful and useful in the context of losets too.

You should be careful with your intuition and not be swayed by appearance. For example consider the set $X = [0, 1) \cup [2, 3)$ with usual order. This is a loset. Take $S = [0, 1) \subset X$. You might think that this set has no supremum simply because 1 is not in our set. In fact 2 is its supremum as far as the loset X is concerned. In other words, in the loset before your eyes this element

2 is an upper bound for the set and there is no upper bound smaller than this.

In fact, as a loset the above set X is isomorphic to $[0, 2)$. The map $f(x) = x$ for $0 \leq x < 1$ and $f(x) = x - 1$ for $2 \leq x < 3$ sets up an order isomorphism of X with $[0, 2)$. Here is a simple observation regarding existence of infimums and supremums.

Theorem: Let X be a loset. The following are equivalent.
(i) Every non-empty bounded set has supremum.
(ii) Every non-empty set bounded above has supremum.
(iii) Every non-empty bounded set has infimum.
(iv) Every non-empty set bounded below has infimum.

Proof: If (i) holds then (ii) can be shown as follows. Take $A \neq \emptyset$ bounded above. Pick $a \in A$. Consider $B = \{x \in A : a \leq x\}$. This is non-empty because $a \in B$. Also B is bounded above by the bound of A. Moreover, B is bounded below by a. Easy to see that supremum of B, which exists by (i), works as sup of A. In fact, A and B have the same set of upper bounds.

Obviously (ii) implies (i).

Similarly (iii) and (iv) are equivalent.

Assume (ii) holds. We can argue (iii) as follows. Take $A \neq \emptyset$ which is bounded. Take the set B of all lower bounds of A. Since A is bounded, this is non-empty. Also every element of A is an upper bound of B. Use (i) to get $s = \sup B$.

We claim that s is infimum of A. In fact, every element of A being an upper bound of B, we conclude that $s \leq a$ for each $a \in A$. That is, s is a lower bound for A. Further, if x is a lower bound of A, then $x \in B$ by definition of the set B and hence $x \leq s$. Thus s is the greatest lower bound of A, that is, inf A.

Similarly, one shows (iii) implies (i).

Say that a loset is *boundedly complete* if any of the above conditions hold.

Let X be a loset. Sets of the form $\{x \in X : a < x\}$, $\{x \in X : x < b\}$, $\{x \in X ; a < x < b\}$ are called *open intervals*. Remember, open intervals in the real line are precisely of the form $(-\infty, a)$ or (a, b) or (b, ∞). This is exactly what we are saying for any loset.

A subset $D \subset X$ is dense, if every non-empty open interval contains a point of D. A loset is *separable* if there is a countable dense set D.

Just as we have a photograph of Q, the set of rationals in the theorem of Cantor, the following is a photograph of the real line.

Theorem (characterization of R): Let $(S, \leq)$ be a loset such that (i) it has no first point, no last point, between any two distinct points there is some other point; (ii) there is a countable set which is dense; (iii) every non-empty bounded subset has supremum. Then S is order isomorphic to R.

Proof: We imitate arguments involved in proving that there is a homeo of R that sends one countable dense set onto another such set. Instead of two countable dense subsets of R, we start with Q of R and countable dense set D of the loset X. Observe that D can not have a first point. In fact if $d \in D$ then the open interval $\{x \in X; x < d\}$ is non-empty (because X has no first point) and hence must contain a point of D. Similarly D has no last point. Thus D has no end points. Also given $a < b$ from D, the open interval $\{x \in X : a < x < b\}$ is non-empty because of hypothesis on X. And D being dense, there is $d \in D$ such that $a < d < b$.

Now we can appeal to Cantor and fix an order preserving isomorphism $f : D \to Q$ and repeat earlier arguments. In other words, take $x \in X$, observe that the set $\{\varphi(d) : d \in D; d < x\}$ is bounded above in R and its sup be denoted by $\varphi(x)$. First justify that when $x \in D$ then this sup indeed same as $\varphi(x)$ so that there is no clash of notation. This is order preserving isomorphism.

This completes proof.

It can be shown (you will do that later) that violation of any one of the features in the above picture will not give you R. In other words, given any one of the above conditions, there are losets that do not satisfy the given condition but satisfy all other conditions; and not order isomorphic to R.

We discuss one more characterization of the reals using its order properties. We need some definitions. Let X be a loset. A *cut* means a partition $X = L \cup U$ such that $a \in L, b \in U \to a < b$. We shall consider only cuts where both L and U are non-empty. There are exactly four possibilities: L has an upper bound in L (hence $= \sup L$) or does not have; U has a lower bound in U (hence $= \inf U$) or does not have. Accordingly we have four possibilities.

Both exist: A cut is said to be a *jump* if L has upper bound in L and U has a lower bound in U.

None exists: A cut is said to be a *gap* if L has no upper bound in L and U has no lower bound in U.

Exactly one happens: A cut is said to be *Dedikind cut* if L has upper bound in L but U has no lower bound in U OR U has a lower bound in U but L has no upper bound in L.

For example let $X = [0,1] \cup [2,3) \cup (3,4]$ with usual order. Then $L = [0,1]$ leads to a jump; $L = [0,1] \cup [2,3)$ leads to a gap; $L = [0, \frac{1}{2})$ or $L = [0, \frac{1}{2}]$ leads to Dedikind cut.

Notice that in describing the cuts above, we have only described L. This is enough because U is necessarily the complement of L. It is not necessary to carry the extra baggage, U.

Here is another photograph of the real line and to recognize that this is a photo of R we compare this with the previous photo.

Theorem: Suppose that $(X, \leq)$ is a loset such that (i) it has no first point and no last point (ii) it has a countable dense set (iii) every cut is a Dedikind cut. Then X is order isomorphic to R.

Proof: Need to verify X satisfies previous theorem. Take $a < b$. If there are no points in between, then $L = \{x : x \leq a\}$ and $U = \{x : b \leq x\}$ is gap.

Let $A \subset X$ be any non-empty set bounded above. Put

$$L = \{x \in X : x \leq a, \text{ for some } a \in A\} = \bigcup_{a \in A} \{x : x \leq a\}.$$

This gives a cut. Indeed, $\emptyset \neq A \subset L$. If b is any upper bound of A, using the fact that X has no last element get t such that $b < t$ to see $t \in L^c$. Thus both L and L^c are non-empty. If $x \in L$, then there is $a \in A$ such that $x \leq a$ and anything smaller than x is also smaller than a and hence is in L. So it is a cut. Hence this must be a Dedikind cut.

Suppose L has a sup $s \in L$. Since $A \subset L$, we see that s is an upper bound of L. As noticed by you, $s \in L$ tells that there is $a \in A$ with $s \leq a$. But s is sup of L tells $s = a$. In other words this upper bound s of A is in A and is hence sup A.

Suppose L^c has infimum $s \in L^c$. Since $A \subset L$ we see that s is an upper bound of A. Let $t < s$. As noted at the beginning of the proof, we can get $t < u < s$. Since s is inf L^c, we conclude that $u \in L$. Hence there is $a \in A$ with $u \leq a$. In particular $t < a$ showing that t can not be upper bound of A. Thus s is the least upper bound of A.

Thus every non-empty set bounded above has a sup.
This verifies conditions of the earlier theorem to complete the proof.

Construction of R:

We shall now proceed to the construction of the real number system. Before we do this, you should be convinced that it is necessary.

For example a number like 13/9 has a clear and concrete existence in our minds. On the other hands, some numbers are difficult to understand, difficult to believe they exist and so on. For example $\sqrt{2}$ does not have as concrete existence as 13/9. Of course, if you take some trouble you can visualise this number. For example (and this is not the only way), you can say, let us draw a concrete right angled triangle with two sides of unit length. Let us measure the diagonal.

If you consider

$$\sqrt{3}^{\sqrt{17}}$$

I am sure none of us have any idea what it is. We would even wonder if there is such a number at all. So it is necessary to convince ourselves that such numbers exist.

You might be deceived by your exposure to calculators. You might say, what is there, if I punch square root instruction and press the x^y function I can get this and show you. But this is illusion. What you get is an approximation. I am sure many of you probably do not even know what your calculator is showing. You just believe 'it *should* show what I asked for'.

You might also momentarily wonder: Did we not meet all numbers already? Did we not show last semester existence of square roots and powers. We did this in several steps, not only that, we developed several laws of indices and so on concerning expressions of the form a^b. Yes, you are right, we actually did all the hard work (thank God, it was over!).

But remember, all that was achieved using properties of the R, real number system, we announced at the beginning. Everything done so far used all the properties (and only those properties). Since we were clear headed and have been careful to list the properties we used, our job now is well focussed. We need to answer the question: is there a system satisfying those properties we listed. This is what we do now.

Thus constructing R simply means exhibiting a system satisfying those properties we listed. We need to exhibit a set R and define operations $+, \cdot$ along with elements 0, $1 \neq 0$ and relation $\leq$ such that those axioms hold. We shall recall those axioms only briefly now:

(I) deals with $+$ and 0; says we have a group;

(II) deals with $\cdot$ and $1 \neq 0$; says nonzero things form a group;

(III) deals with $+$ and $\cdot$; says these two things are friendly;

(IV) deals with $\leq$; says we have a loset;

(V) deals with $\leq$ and $+$ and $\cdot$; says they are friendly;

(VI) deals with subsets: says bounded non-empty subset has sup.

***R* from *Q*:**

So for the next couple of hours we should not use real numbers, we do not have them, we construct them. However we do have rational numbers before us and we use them. It comes as a surprise that real numbers are very simple and nothing profound or complicated. What plays a crucial role is the intuition we gained from our earlier discussion.

Recall that, every real number was earlier described as sup of the set of rational numbers below it. This single sentence (and nothing else) is at the heart of what we do now.

Let us say that a subset $x \subset Q$ is a cut if it satisfies the following:
(i) $\emptyset \neq x \neq Q$;
(ii) $p \in x$, $q \leq p$ imply that $q \in x$;
(iii) $p \in x$ implies there exists r with $p < r$ and $r \in x$.

Thus a cut is a non-empty proper subset of Q such that if you take an element in it then everything smaller than that is also there and somethings larger than that are also there. (We can not say everything larger than that is also there).

The collection of all cuts is denoted by R. This is our real number system.

Notice that sets are usually denoted by A, B, C and so on. But now we are using the symbols x, y, z and so on for cuts, which are some special subsets of Q. We denote elements of Q by p, q, r, s and so on.

We do easy things first. We define an order on R by saying

$$x \leq y \ \leftrightarrow x \subset y.$$

Since $x \subset x$ conclude that $x \leq x$.
If $x \leq y$ and $y \leq z$ then $x \leq z$ as a consequence of set inclusions.
If $x \leq y$ and $y \leq x$ then $x = y$.
Thus we have a poset. Now we show that it is a loset. Take x and y. Suppose $x \not\leq y$. Thus $x \not\subset y$. So pick $p \in x$ and $p \notin y$. The second property of p implies that nothing larger than p is in y (remember y is a cut). Thus everything in y must be smaller than p. The first property of p implies that everything smaller than p is in x (remember x is a cut). Thus everything which is in y must also be in x. Thus $y \subset x$, that is, $y \leq x$.
Did I use too much English? You can rewrite using symbols.

Thus R is a loset. We shall now show that this loset has sup property. Take $\emptyset \neq S \subset R$. Thus S consists of some cuts. Let S be bounded, say $x \leq z$ for all $x \in S$. We exhibit sup for the set S. Let x^* be the union of all the cuts x that belong to S — just keep in mind that cuts are certain subsets of Q.

Since S is not empty, take $x \in S$ and take $p \in x$ to see $p \in x^*$. Thus $x^* \neq \emptyset$. Since $x \subset z$ for each $x \in S$ we see that $x^* \subset z$. Since z is a cut take $q \notin z$. Clearly $q \notin x^*$. Thus $x^* \neq Q$.

Let $p \in x^*$, take $x \in S$ so that $p \in x$. But then everything smaller than p is in this x (remember x is a cut) and hence in x^* too.

Let $p \in x^*$, take $x \in S$ so that $p \in x$. But then a little larger than p is also in this x (remember x is a cut) and hence in x^* too.

These three observations show that x^* is a cut. Since $x \subset x^*$ for each $x \in S$, that is, $x \leq x^*$, we see x^* is an upper bound for S. If z is any upper bound of S, then for each $x \in S$ we have $x \leq z$, that is, $x \subset z$ so that $x^* \subset z$, that is, $x^* \leq z$. That is x^* is the smallest upper bound. Thus S has a sup.

Hope you appreciate how trivial things are, and how we exhibited sup in a painless way. We shall now go on to define addition and multiplication.

R **from** Q:

We are now in the process of constructing real number system using Q, the set of rational numbers. A cut x is a non-empty proper subset of Q such that for any of its elements, everything below is also there and somethings above are also there. R is the collection of all cuts.

We defined an order $x \leq y$ if $x \subset y$ and showed that with this definition R is a loset having least upper bound property.

We now define addition on R. Let $x \in R$ and $y \in R$. Then

$$x + y \;=\; \{p + q : p \in x; q \in y\}.$$

This set on right side is non-empty because, x and y are so and hence any $p + q$ with $p \in x$ and $q \in y$ would do. There are also points which are not here. Indeed if $p_1 \notin x$ and $q_1 \notin y$ then for any $p \in x$ we have $p < p_1$ and for any $q \in y$ we have $q < q_1$. Thus $p+q < p_1+q_1$ showing that $p_1+q_1 \notin x+y$.

Finally, if $p \in x$ and $q \in y$ then x being a cut things r a little larger than p are also in x so that things like $r + s > p + q$ are also in $x + y$. Thus $x + y$ as defined above is indeed a cut.

The equality $x + y = y + x$ follows from the fact that the set of $p + q$ is same as the set of $q + p$. Similarly $(x + y) + z = x + (y + z)$ follows simply because the set of numbers $(p+q)+r$ is same as the set of numbers $p+(q+r)$. See how we are using the corresponding rule for Q.

Let

$$0^* = \{p \in Q : p < 0\}.$$

This is clearly a cut. Note that zero itself is not in the set.

We claim this is zero element. Take any $x \in R$. We show $x + 0^* = x$. If we have $p \in x$ and $q \in 0^*$ then $q < 0$ so that $p + q < p \in x$. Since x is a cut we conclude $p + q \in x$. Thus $x + 0^* \subset x$. Conversely, let $p \in x$. Let $r > p$ and $r \in x$. Then $p = r + (p - r) \in x + 0^*$. Thus $x \subset x + 0^*$.

We shall show inverse. Let $x \in R$. Define

$$y = \{-q : q \in x^c ; q \text{ not least element of } \ x^c\}.$$

Here x^c, as usual, is the complement of x, that is, all rational numbers q such that $q \notin x$.

If $p \in x$ and $-q \in y$ then $q \in x^c$ tells $p < q$ so that $p + (-q) < 0$. Thus $x + y \subset 0^*$. Conversely, if we show that every negative rational is in $x + y$ it follows that $x + y = 0^*$ and completes the proof that y is inverse of x. For this it is enough — because $x+y$ is a cut — to exhibit a sequence of elements $p_n - q_n$ in $x+y$ such that every negative rational is smaller than one of these numbers.

We first make a construction. Take any $p \in x$ and $q \in x^c$. Set $p_0 = p$ and $q_0 = q$. Take mid-point $(p + q)/2$ which is a rational number. If this is in x then this is p_1 and $q_1 = q$. If this mid-point is in x^c then it is q_1 and $p_1 = p$. Now take mid-point of p_1 and q_1. continue and convince yourself you can write the inductive step to construct the sequence (p_n, q_n) so that each p_n is in x and each q_n is in y and $q_{n+1} - p_{n+1} = (q_n - p_n)/2$. Thus, by induction, we have

$$q_n - p_n = \frac{q-p}{2^n}.$$

Observe that given any rational number $r > 0$, there is n such that $q_n - p_n < r$. This is because of the following reason. By binomial theorem (for integers) $2^n > n$ so that by Archimedian property there is an n such that

$$\frac{q-p}{r} < n < 2^n; \qquad \text{or } \ q_n - p_n < r.$$

Every positive rational is larger than one of the $q_n - p_n$.

Let s be any rational number $s < 0$. The above argument shows that s is smaller than one of the $p_n - q_n$. If q_n is not the least element of x^c, this already implies that $p_n - q_n \in x + y$ and hence $s \in x + y$. if q_n is the least element of x^c then we can not say that $p_n - q_n \in x + y$. However, we can get n such that $s/2 < p_n - q_n$, by the above argument. Take a rational q_n^* such that $q_n < q_n^* < q_n - s/2$ (remember $-s > 0$). Then, of course $q_n^* \in x^c$ and is not its least element.

$$s = \frac{s}{2} + \frac{s}{2} < (p_n - q_n) + (q_n - q_n^*) = p_n - q_n^* \in x + y$$

completing the proof that $s \in x + y$. Thus $x + y = 0^*$.

We proved that R with addition so defined is a group (abelian). Let us see its relation with order. Let $x < y$ and $z \in R$. Wish to show $x+z < y+z$.

Of course $x \subset y$ tells that if $p \in x$ then $p \in y$ so that for any $q \in z$ we have $p+q \in y+z$ showing that $x+z \subset y+z$, that is, $x+z \leq y+z$. The only question is whether equality can hold here. If so, we would have $x+z = y+z$; add $-z$ to both sides, use associativity (already proved), replace $z+(-z)$ by 0^*, use $x + 0^* = x$ etc to see $x = y$. But this is false.

We need to define multiplication. This is extremely tricky. You can not blindly put, just like addition, $x \cdot y$ to consist of all pq with $p \in x$ and $q \in y$. Note we are using dot for the multiplication we plan to define, but for product of rational numbers we do not use dot. This will never be a cut, simply because all negative numbers below some stage will be in both x and y showing that all sufficiently large positive numbers are in this suggested set.

The root cause of the problem is that multiplication of two negative numbers is positive number. This suggests us to define product of positive numbers first. This is what we do now.

let $x > 0^*$ and $y > 0^*$. We define

$$x \cdot y \quad = \quad \{q \in Q : q \leq 0 \quad \text{or } q = p_1 p_2,\ 0 < p_1 \in x,\ 0 < p_2 \in y\}.$$

Thus all negative rationals are put in this set right away. This is a cut. In fact, this is clearly non-empty. If $r \notin x$ and $s \notin y$ then you can easily argue that $rs \notin x \cdot y$ so that the suggested set is not all of Q. Let $r \in x \cdot y$ and $s < r$ be rational. We need to show $s \in x \cdot y$.

If $s \leq 0$ then by definition $s \in x \cdot y$. So let $s > 0$. Then $r > 0$ too so that there is $p_1 \in x$, $p_2 \in y$ with $0 < p_1$ and $0 < p_2$ and $r = p_1 p_2$. Thus $0 < s/p_2 < r/p_2 = p_1$ and hence $s/p_2 \in x$. Thus

$$s = \frac{s}{p_2} p_2 \in x \cdot y.$$

Finally let $r \in x \cdot y$. Need to show some thing larger than that in $x \cdot y$. If this $r \leq 0$ then any $p_1 p_2$ will do. So let $r > 0$. Then $r = p_1 p_2$. Take a little larger than p_1 in x (remember p_2 is positive) etc to complete the proof.

Thus the suggested set is a cut and hence it is in R. Need to show $x \cdot y = y \cdot x$ and $(x \cdot y) \cdot z = x \cdot (y \cdot z)$. These follow as in the case of addition using the corresponding properties in Q. Shall now exhibit multiplicative unit.

$$1^* = \{q \in Q . q < 1\}.$$

This is clearly a cut — need to use between any two rationals there is another one etc to show things a little larger.

Take any $x > 0^*$. We show $x \cdot 1^* = x$. All rationals $r \leq 0$ are in both: left side by definition of product and right side by definition of the element 1^*. Need to show both sides have the same positive elements. If you take a positive element $p_1 p_2$ on left side then $0 < p_2 < 1$ implies that this is smaller than p_1 and hence is in x. conversely, if we take $p \in x$, then a little larger, say, $r > p$ is also in x and hence

$$p = r\,\frac{p}{r} \in x \cdot 1^*.$$

We now show inverse. Let $x > 0^*$. Let

$$y = \{q \in Q : q \leq 0\} \bigcup \left\{\frac{1}{q} : q \in x^c, q \text{ not least element of } x^c\right\}$$

All negative rationals are in y. If $q \in x^c$ then anything larger than that is in x^c so that any positive smaller than $1/q$ is also y. If $1/q \in y$ then q being not the least element of x^c take something smaller than that and one divided by that etc; to see some thing larger than $1/q$ is also in y. Thus y is a cut and $y > 0$.

Need to show now that $x \cdot y = 1^*$. All rationals $r \leq 0$ are in both sides. So let us show they have the same positive things too. Let $r \in x \cdot y$. Thus $r = p_1 p_2$ with $0 < p_1 \in x$ and $0 < p_2 \in y$. But definition of y tells $p_2 = 1/q$ with $q \in x^c$. But then $q > p_2$ showing that $p_1 p_2 < 1$ and hence is in 1^*.

Finally we need to show that $0 < r < 1$ implies $r \in x \cdot y$. Repeat the argument given for addition. Argue, using the same construction, that p_n/q_n goes closer and closer to one, that is, anything smaller than one is smaller than one of these. This is easily achieved by observing

$$0 \ < \ 1 - \frac{p_n}{q_n} \ = \ \frac{q_n - p_n}{q_n} \ \leq \ \frac{q_n - p_n}{p}.$$

and you can make $q_n - p_n$ as small as you please, see earlier calculation.

This completes the proof that y is multiplicative inverse of x.

Let us see the relation of order. We only have the following rule to be verified: $x > 0$ and $y > 0$ implies $x \cdot y > 0$. Hence we can proceed to verify this without completing our definition of multiplication. Just remember, we have not defined multiplication if one of x and y is not positive. Shall do so soon.

But, of course, verification of the above property is trivial because, $x > 0$ tells we have $0 < p_1 \in x$ and similarly $0 < p_2 \in y$ and hence $0 < p_1p_2 \in x \cdot y$ to complete the proof.

Let us complete the definition of multiplication.

$$x \cdot y = \begin{cases} - \quad [(-x) \cdot y] & \text{if} \quad x < 0, \quad y > 0 \\ - \quad [(x) \cdot (-y)] & \text{if} \quad x > 0, \quad y < 0 \\ \quad\; [(-x) \cdot (-y)] & \text{if} \quad x < 0, \quad y > 0 \\ \quad\; 0 & \text{if} \quad x = 0, \quad \text{or} \quad y = 0 \end{cases}$$

This may look complicated, but first think of real numbers and observe these are correct. That was the reason for this definition.

This completes the definition of multiplication. That this satisfies all requirements is routine patient verification. You will do in exercises. This verification does *not* need using cuts. You need simply properties already proved and some 'algebra' — I do not mean you have to use theorems from algebra!, I mean *you do* some algebra.

This completes construction of R. This method is due to Dedikind. Later you will se another method due to Cantor.

enter Peano:

Thus if we are granted Q, we could construct

$$(R, +\,,\ \cdot\,,\ \leq)$$

satisfying all the needed rules. I have been a little loose; I freely used whatever I needed about Q without enunciating beforehand what I expect Q to obey. But do not worry; if some one challenges, you can also enunciate, just read the above argument and see what is used and state those. That is all. This can be easily done. But let us not be too legalistic and spend time on doing this. (This is important, but does not help in our understanding. So we are not spending time.)

But who gives us Q? We shall construct ourselves. You can not get something out of nothing. So what is the basic ingredient needed. Here it is. We shall be brief in the details.

Peano Axioms:
There is a set N_0 and a function $s : N_0 \to N_0$ satisfying the following three rules:

(i) $x = y$ if and only if $s(x) = s(y)$.
(ii) there is a unique element θ which is not in the range of s.
This θ, being unique, shall be denoted by 0.

(iii) If $A \subset N_0$, and $0 \in A$, and $x \in A$ implies $s(x) \in A$; then $A = N_0$.

The function s is called successor function.

If you want you can restate axiom (i) as

$$x = y \leftrightarrow s(x) = s(y).$$

There is no need to put 'iff'. The definition of function already tells you that when $x = y$, then you must have $s(x) = s(y)$.

Second axiom means there is no x such that $s(x) = \theta$. if you want you can restate this as

$$(\exists!\theta)(\forall x)\ \neg(s(x) = \theta).$$

The last axiom is called induction axiom, because it allows to use mathematical induction. This is also called well-ordering axiom, because it makes N_0 as well-ordered set; with a natural order to be defined shortly.

It is also possible not to demand uniqueness of θ in (ii). Take any one such θ and name it zero and proceed to axiom (iii). Then it follows that there can not be another such θ. Suppose that $\theta' \neq \theta$, also satisfies (ii) then the set $N_0 - \{\theta'\}$ satisfies conditions of (iii) and hence must equal N_0.

Let us not bother with minor details of no interesting consequence. The above are our axioms. These axioms are called Peano axioms.

You should note something very very interesting about axiom (iii). Usually axioms about a system tell about rules to be satisfied by elements of the system. For example axioms of groups tell you the rules to be satisfied by the elements of the group. Similarly about rings or fields. None of these tell you 'rules that subsets should obey'.

axiom (iii) tells you a rule to be obeyed by subsets of N_0, not by elements of N_0. This is a fine distinction, difficult to appreciate for you, but you should not ignore. The same remark applies to the 'least upper bound axiom' for R. all other axioms tell you rules to be obeyed by real numbers, or existence of special real numbers like zero and one. But the lub axiom is a rule that imposes a rule to be satisfied by subsets of R.

Let us accept that there is such a system (N_0, s). We see later.

Peano arithmetic: from (N_0, s) to $(N_0, +, \cdot, \leq)$:

We shall use the successor map to define addition and multiplication on N_0 as follows. These are defined using induction.

$$x + 0 = x; \qquad x + s(y) = s(x + y).$$

In other words, we first define $x + 0 = x$ for every x. Then we define

$$x + 1 = x + s(0) = s(x + 0) = s(x) = x + 1.$$

and then define $x + 2$ etc. Did we define $x + y$ of every y? Yes. the set of y for which we defined $x + y$ includes 0 and if it includes y then it includes $s(y)$ and hence this is all of N_0.

One can show again several rules like:
(commutativity) $x + y = y + x$ for all x and y;
(associativity) $(x + y) + z = x + (y + z)$ for all x, y, z;
(cancellation law) $x + z = y + z$ implies that $x = y$

We define multiplication

$$x \cdot 0 = 0; \qquad x \cdot s(y) = x \cdot y + x.$$

This satisfies analogues of the above three rules: commutativity; associativity and cancellation law.

Now we can define order:

$$x \leq y \leftrightarrow (\exists z) \;\; x + z = y.$$

You might think this is not right because for every x and y there is always such an z, namely $y-x$. But this is a wrong thought because at this moment we do not know this minus sign; for example there is nothing like $3-4$. At this moment the existential quantifier refers to existence of z in N_0. Since we have nothing else I did not have to say $\exists(z \in N_0)$. We have only N_0 at this stage before us, no more.

Order satisfies several rules:
It is a in linear order;
it has a first element;
it has no last element.

Very interesting, the proof of transitivity uses associativity of multiplication: suppose $x \leq y$ and $y \leq w$ then we have z and u such that $x+z=y$ and $y+u=w$. Now

$$x+(z+u)=(x+z)+u=y+u=w.$$

As I said we are being brief, but you can verify all the rules.

N_0 **to** Z**:**

As already noticed, given x and y there may not always exist u such that $y+u=x$; in other words I may not be able to talk about $x-y$ for some pairs. Of course, the cancellation law tells us that *if there is* one such u then there is *only* one u.

Since we are unable to do subtraction, for some pairs we cook up some ideal elements and put them in our set to make this possible. Basic philosophy is this: The very phenomenon that is not possible is put in as an element and this element represents that phenomenon. But then we should be careful, $3-4$ and $8-9$ should be same.

This is achieved by considering *all possible phenomena*; identify two phenomena if there are reasons for you to believe that they are same. Then hopefully this set would make all phenomena possible. All this is abstract sermon. Let us actually do it.

Let

$$Z_0=\{(a,b): a,b \in N_0\}$$

Thus when I say (a,b) I am thinking of $a-b$, but I can not say so because you will question me: what is this minus sign? Till I introduce this sign I should be careful not to go beyond my vocabulary.

Some people think of $b-a$ when they write (a,b). It makes no difference as long as you think one way or other (and not both) throughout.

We identify some pairs. Say

$$(a,b) \sim (m,n) \leftrightarrow a+m = b+n.$$

Actually I wanted to identify when $a-b=m-n$, I said the same thing without using minus sign, that is, using only the means I have now. It is a routine matter to verify that this is an equivalence relation.

Let the space of equivalence classes be denoted by Z. We have enlarged the set N_0. We can identify N_0 as a subset of Z. Take $n \in N_0$ and let $\varphi(n)$ be the equivalence class containing $(n,0)$ — remember, I have $a-b$ at the back of my mind when I wrote (a,b).

Actually, I can make a nice selection of 'one thing' from each equivalence class and use them rather than thinking of the huge equivalence class all the time. Each equivalence class contains exactly one pair where one of the coordinates is zero.

if $a=b$ then easy to see $(a,b) \sim (0,0)$.

if $a<b$, then there is c such that $a+c=b$, then easy to see $(a,b) \sim (0,c)$.

if $b<a$, then there is c such that $b+c=a$, then easy to see $(a,b) \sim (c,0)$.

These are the only possibilities because we have linear order.

Clearly a pair $(a,0) \sim (0,b)$ implies $a+b=0+0$ so that $a=b=0$.

If $(a,0)=(b,0)$ then $a=b$. If $(0,a)=(0,b)$ then $a=b$.

Thus there is exactly one pair (a,b) in each equivalence class with one coordinate zero. Thus we could think of this element instead of the equivalence class. In other words we could have defined

$$Z = \{(a,b) : a=0 \quad \text{or} \quad b=0\}$$

And we could have though of $(a,0)$ as $a-0=a$ and $(0,a)$ as $0-a=-a$.

Yes you are right, we did complicate life. We could have just said 'add a symbol minus n" for every n, then these symbols along with N_0 is precisely our Z. Such a prescription will serve the present purpose. However it misses the general philosophy, it looks artificial, serves only immediate purpose. The philosophy 'if you can not achieve certain phenomenon, consider the set of phenomena' gives a unity for mathematical constructions.

But you can keep both pictures in mind. you will see first picture looks complicated but 'smooth' to operate; second picture looks smooth but complicated to operate.

On this set Z we define addition multiplication and order:

$$(a,b)+(c,d)=(a+c,b+d).$$

$$(a,b)\cdot(c,d)=(ac+bd,ad+bc)$$

Under addition it becomes a group and multiplication obeys all the three rules we had earlier. We define order as follows.

$$(a,b)\leq(c,d)\leftrightarrow ad+bc\leq ac+bd.$$

This definition makes Z a loset and the order is friendly with addition and multiplication.

In the second picture, these definitions take the following shape. Of course, you can repeat the same earlier formulae.

$$(n,0)+(m,0)=(n+m,0);\quad (0,n)+(0,m)=(0,n+m)$$

To define $(n,o)+(0,m)$ we need to consider two cases. If $n>m$, say, $n=k+m$ then the sum equals $(k,0)$. But if $m>n$ and say $m=k+n$ then the sum equals $(0,k)$. Think about it. Multiplication is as follows.

$$(m,0)\cdot(n,0)=(mn,0);\quad (0,m)\cdot(0,n)=(mn,0)$$

$$(m,0)\cdot(0,n)=(0,mn).$$

Definitely messy. However order takes simpler shape (as it should). Each $(0,n)$ is smaller than each $(m,0)$. Among $(m,0)$ kind it becomes larger as m becomes larger, whereas among $(0,m)$ kind it becomes as m becomes larger.

Of course all of it just provides you the picture you already have in mind

$$\cdot\cdot\,,-3,\ -2,\ -1,\ 0,\ 1,\ 2,\ 3,\ \cdots\cdots.$$

This is how we think of Z from now on. Once constructed, no matter how, Z has existence independent of how it is constructed and by whom.

Z to Q:

The trouble with Z is that while we can add and subtract, we can multiply but not divide. In other words we have additive inverses but not multiplicative inverses. By the way, using known rules, we understand that anything multiplied by zero has to be zero, so when we say multiplicative inverse, we mean only for non-zero things. We know that we can never have multiplicative inverse for zero.

If I could not subtract earlier $3-4$ then we added $(3,4)$ and when we see this 'symbol' we think of $3-4$. In the same way, if I can not divide $3/4$, I add the symbol $(3,4)$ and when I see this I think of $3/4$.

Of course there is a confusing point. When you see $(3,4)$ do you have at your back of the mind $3-4$ or $3/4$? You can not have both. That is why I said earlier how we think of Z. We do not think of $3-4$ now, it is represented by the symbol -1, that is all. There is no pairs. elements of Z are as described at the end of the discussion above. First understand this point.

So since division between some pairs of points of Z is not possible, let us repeat the earlier process. Let

$$Q_0 = \{(a,b) : a, b \in Z, b > 0\}.$$

But as you see I can not think of $3/4$ and $6/8$ as two separate quantities. Define in Q_0 an equivalence relation

$$(a,b) \sim (c,d) \leftrightarrow ad = bc.$$

I wanted to identify if $a/b = c/d$ but I can not say this because I do not have meaning for these quantities. But you see i said the same thing using the permitted vocabulary.

Let Q be the set of equivalence classes. This is our set of rationals. Define

$$(a,b) + (c,d) = (ad + bc, bd), \quad (a,b) \cdot (c,d) = (ac, bd).$$

First observe that bd is nonzero and hence the definition makes sense, the resulting pairs are in our collection. Also observe that these definitions respect the equivalence relation. That is

$$(a,b) \sim (a_1,b_1), (c,d) \sim (c_1,d_1) \Rightarrow (ad + bc, bd) \sim (a_1d - 1 + b_1c_1, b_1d_1).$$

$$(a, b) \sim (a_1, b_1), (c, d) \sim (c_1, d_1) \Rightarrow (ac, bd) \sim (a_1c_1, b_1d_1).$$

This observation, that the operations respect the equivalence relation, help us to define the operations on the space of equivalence classes, that is, on Q.

These operations make Q into a field.

We define order by

$$(a, b) \leq (c, d) \leftrightarrow ad \leq bc$$

Again, this definition respects the equivalence relation and hence defines an order on Q. This makes Q into a loset and the order is compatible with addition and multiplication.

These operations and order extend the operations and order we already have on Z. Of course, for this sentence to make sense you should first be able to see Z inside Q. The equivalence class containing the element $(a, 1)$ is identified with $a \in Z$. If you want to be formal, you can define map φ on Z into Q and say that this map preserves the 'structure'.

In other words Q is an ordered field. It has the following property, called Archimedian property.

$$x \in Q \Rightarrow (\exists \ k \in N_0) \ \ x < k.$$

Every number is smaller than some integer.

Peano again:

The final thing we need to do is to convince ourselves that there is a set N_0 and $s : N_0 \to N_0$ satisfying the axioms of Peano. This comes from set theory and we only indicate without going into the details. We define

$0 = \emptyset$ empty set, that is the set which has nothing.

$s0 = \{\emptyset\}$ set with only one element, namely, the empty set.

$ss0 = \{\emptyset, \{\emptyset\}\}$ set with two elements, namely, the empty set and the set which contains the empty set. Equivalently, it consists of the elements of the previous set (which is only one) and the previous set itself.

$sss0 = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$ set consisting three elements all the elements of the previous set (which are two) and the previous set itself.

In general $s(*)$ consists of all elements of the previous set $*$ along with one more element, namely, the previous set $*$ itself. In symbols you can think of ($*$ may be confusing, think of k and think of $k+1$ instead of sk)

$$k + 1 = k \cup \{k\}$$

before you get confused, see that right side is union of two sets. The first set is k and hence everything which is in k is here. The second set has exactly one thing in it, namely, k itself. Thus this new set has only one extra thing namely the previous set itself.

The set N_0 consists of all these objects we defined and the map is

$$s(k) = k \cup \{k\}$$

This set satisfies all our requirements.

There are two important points in what I said above that make you suspect something is wrong somewhere. Yes, you are justified in not accepting what I said above. Did I already use $0, 1, 2, 3 \cdots$? Can I use them? I finally said N_0 consists of *all these*. What is meant by *all these*? How long should I construct the sets which I was describing above?

This is where set theory enters the picture. Here is a truth.

There is a smallest infinite set ω_0 which has the property

$$x \in \omega_0 \Rightarrow x \subset \omega_0.$$

That is, there is an infinite set ω_0 which satisfies the above condition and if ω is another set satisfying that condition then $\omega_0 \subset \omega$.

We can show that such a set contains all the things listed earlier by us. In other words

$$\emptyset \in \omega_0$$

$$\{\emptyset\} \in \omega_0$$

$$\{\emptyset, \{\emptyset\}\} \in \omega_0$$

$$\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\} \in \omega_0.$$

and so on. I needed some symbols to show you my listing and thus k etc entered in my earlier description. I could have listed my sets one after the other without using the symbol k. But how long can I do? If I used 'etc' how long should you do?

Actually our N_0 is exactly this set ω_0. See how we described it in one sentence without using symbols like k, $k+1$ and without using words like 'so on'. Axioms of set theory allow you to show that there is such a set (and it is unique).

This is called first infinite ordinal. This definition of ordinal numbers is due to von Neumann. Cardinal and ordinal numbers were invented by Cantor. Cardinal numbers tell you counting without any idea how things are placed: one, two, three, etc.

Ordinal numbers tell you counting when things are standing in a line; first, second, third, etc. You can do this for any well ordered set. In any well ordered set there is a least element. That is the first element. Consider its complement, it has a least element, that is the second element of the set and so on. If you have counted first, second, third etc then remove all these. If the resulting set is non-empty take its first element it is ∞-th element of the set. Remove this also. Least element of the remaining set is $(\infty + 1)$-th and so on. What I called ∞ here is precisely ω_0.

You have learnt a little about cardinal numbers, you have learnt a little about well-ordered sets. It is worth while to learn a little about ordinal numbers too, perhaps later sometime if there is some time.

This completes our discussion of construction of R. That is, we have constructed a set and shown operations and order that satisfy all the properties listed last year.

Metric spaces:

We shall proceed to the next topic of our discussion, namely, metric spaces. These are nothing but sets where there is a notion of distance between pairs of points. You can say how far is one point from other.

Once you have this notion, you can understand the concept of closeness, after all two points are close if the distance between them is not much. Once you can feel closeness of points, you can feel if a sequence of points are getting closer and closer to a given point. In other words, convergence of sequences makes sense. Once you can do this, you can talk about continuous functions.

After all continuous functions are just those that preserve convergence. So you see you can build an excellent story, imitate on your set what you have done on R. Of course, you can not add or multiply unless these are possible in your set.

This is not just for the sake of generalisation, it has applications far beyond expectations.

Let X be a nonempty set. For each pair of points x, y in the set we want to have the concept of distance between the points $d(x, y)$. One thing you feel immediately is that $d(x, x) = 0$. Also if x and y are different points then naturally they are separate and must be at a distance no matter how small; thus $d(x, y) > 0$.

One thing experience tells us is that distance is symmetric. the distance from x to y must be same as the distance from y to x. That is, $d(x, y) = d(y, x)$. Of course, I depended on your experience.

It is also quite possible to think that this ned not be true. For example, suppose you try to measure distance by the energy you need to spend from going from one place to other. For points far apart you need to spend more energy in travelling whereas if the points are close you need not spend too much energy. Such a method appears very reasonable. But imagine x is the point at the base of a hill and y is the point on the top of the hill. You will surely agree that it takes more energy to go from x to y than from y to x. Thus it is also conceivable that distance is not symmetric.

However We adapt symmetry.

Finally again from practice, we see that distance is in a sense 'direct distance' whatever it may mean. Thus going from x to z directly should be not more than going from x to z via y. That is $d(x, z) \leq d(x, y) + d(y, z)$. This is called triangle inequality. This is true for any three points you take in the plane. This we knew from elementary geometry.

We shall adapt just these three as the things to be satisfied by distance and see what can be done and what is its use.

Definition: Let X be a (non-empty) set and $d : X \times X \to [0, \infty)$ satisfying the three conditions.

(i) $d(x, y) = 0$ iff $x = y$.

(ii) $d(x, y) = d(y, x)$ for any two points x, y.

(iii) $d(x, z) \leq d(x, y) + d(y, z)$ for any three points x, y, z.

Then we say that d is a distance function on X and we also say (X, d) is a metric space.

If we have a metric space (X, d) and $a \in X$ and $r > 0$ then open ball of radius r with centre a is defined to be the following set: $\{x \in X : d(x, a) < r\}$.

Example 1: X is real number system R and $d(x, y) = |x - y|$.

In this case $B(a, r)$ is the interval $(a - r, a + r)$. Thus open balls are just open intervals.

Example 2: X is the plane R^2 and for $x = (x_1, x_2)$ and $y = (y_1, y_2)$ we put

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$$

As explained earlier this is just Pythagoras theorem.

In this case $B(a, r)$ where $a = (a_1, a_2)$ is the set

$$\{(x_1, x_2) : (x_1 - a_1)^2 + (x_2 - a_2)^2 < r\}$$

which is the usual disc.

Example 3: X is R^n and

$$d(x, y) = \sqrt{\sum_1^n (x_i - y_i)^2}.$$

Example 4: X is again R,

$d(x, y)$ is zero if $x = y$ and one if $x \neq y$.

Here $B(a, r)$ consists of just the point $\{a\}$ as long as $r \leq 1$, but it equals all of R as soon as $r > 1$.

Example 5: X is R^2.

$$d(x, y) = \max\{|x_1 - y_1|, |x_2 - y_2|\}; \quad x = (x_1, x_2) \;\; y = (y_1, y_2).$$

Here ball of radius around $(0, 0)$ consists of the square with vertices $(\pm 1, \pm 1)$.

Example 6: X is R^2.

$$d(x, y) = |x_1 - y_1| + |x_2 - y_2|, \quad x = (x_1, x_2) \;\; y = (y_1, y_2).$$

Here ball of radius around $(0,0)$ consists of the region bounded by the four lines $(\pm x_1) + (\pm x_2) = 1$.

More generally,
Example 7: Fix $p \geq 1$. X is again R^2.

$$d(x,y) = (|x_1 - y_1|^p + |x_2 - y_2|^p)^{1/p}, \quad x = (x_1, x_2) \;\; y = (y_1, y_2).$$

Triangle inequality here is not obvious. We shall prove.

Example 8: X is $C[0,1]$ the space of real valued continuous functions on the unit interval. For f, g

$$d(f,g) = \sup\{|f(x) - g(x)| : 0 \leq x \leq 1\}$$

Since continuous functions on $[0,1]$ are bounded this makes sense. Since for any three functions f, g, h and any point x we have

$$|f(x) - h(x)| \leq |f(x) - g(x)||g(x) - h(x)| \leq d(f,g) + d(g,h).$$

This is true for every x nd so by taking sup over x on left side we get the triangle inequality.

Here given f and $r > 0$ the ball $B(f,r)$ consists of all continuous functions whose graph lies in the band: graph of $f(x) - r$ and graph of $f(x) + r$.

Holder, Minkowski:

Let us first show that the d_p example of last time does indeed satisfy the triangle inequality. In what follows we have two numbers $p, q > 0$ such that

$$\frac{1}{p} + \frac{1}{q} = 1. \qquad (\spadesuit)$$

Of course this already implies $p > 1$ and $q > 1$ and given any $p > 1$ there is exactly one q as above.

We start with the simple observation. Let $a > 0$ and $b > 0$. Then

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

Since ($\spadesuit$) tells us that the right side is convex combination of a^p and b^q we are advised to look for a convex function and use the definition of convexity.

Take numbers x and y such that

$$\exp\{x/p\} = a; \quad \exp\{y/q\} = b$$

so that

$$e^x = a^p; \quad e^y = b^q$$

Now convexity of exponential function completes the proof.

$$ab = \exp\left\{\frac{1}{p}x + \frac{1}{q}y\right\} \leq \frac{1}{p}e^x + \frac{1}{q}e^y = \frac{a^p}{p} + \frac{b^q}{q}.$$

This simple inequality leads **Holder**:

Let $a_1, \cdots, a_n$ and $b_1, \cdots, b_n$ be real numbers. Then

$$|\sum a_i b_i| \leq \left(\sum |a_i|^p\right)^{1/p} \left(\sum |b_i|^q\right)^{1/q}.$$

If any one of the right side quantities is zero then the corresponding a's or b's are zero and hence left side is also zero and so both sides are zero.

So let us assume that the two quantities on right side are non-zero, say, c and d respectively. Take $\alpha_i = |a_i|/c$ and $\beta_i = |b_i|/d$ to see

$$\sum \frac{|a_i b_i|}{cd} \leq \sum \frac{|a_i|^p}{c^p} + \sum \frac{|b_i|^q}{q} = \frac{1}{p} + \frac{1}{q} = 1.$$

gives the stated inequality since $|\sum a_i b_i| \le \sum |a_i b_i|$.

Holder leads to **Minkowski**:

$$(\sum |a_i + b_i|^p)^{1/p} \le (\sum |a_i|^p)^{1/p} + (\sum |b_i|^p)^{1/p}.$$

To prove this, assume that a's and b's are positive.

$$\sum (a_i + b_i)^p = \sum a_i (a_i + b_i)^{p-1} + \sum b_i (a_i + b_i)^{p-1}$$

use Holder note $(p-1)q = p$,

$$\sum (a_i + b_i)^p \le (\sum a_i^p)^{1/p} \left[\sum (a_i + b_i)^p\right]^{1/q}$$

$$+ (\sum b_i^p)^{1/p} \left[\sum (a_i + b_i)^p\right]^{1/q}$$

Bring the common factor of right side to the left side to complete the proof (if that common factor is zero to start with, nothing need to be proved).

If a_i, b_i not necessarily positive replace by their modulus etc.

This leads to the **triangle inequality**:

$$d_p(x, z) \le d_p(x, y) + d_p(y, z).$$

where

$$d_p(\alpha, \beta) = (\sum |\alpha_i - \beta_i|^p)^{1/p}$$

You only need to use Minkowski with $a_i = x_i - y_i$ and $b_i = y_i - z_i$.

limit points:

Next couple of lectures, the agenda is to imitate whatever we did in R and R^n, namely to define convergence, open sets etc. In what follows (X, d) is a metric space. Sequences are sequences in X and subsets are subsets of X.

Definition: A sequence x_n converges to a point x iff $d(x_n, x) \to 0$, equivalently, given $\epsilon > 0$, there is N such that $d(x_n, x) < \epsilon$ for all $n > N$.

A point x is a limit point of the sequence (x_n) if there are terms of the sequence that 'keep coming close' to x. That is, given $\epsilon > 0$ and N, there is $n > N$ such that $d(x_n, x) < \epsilon$. Equivalently, any ball around x contains x_n for infinitely many values of n.

a point x is a limit point of a set if for every $\epsilon > 0$, the set $A \cap B(x, \epsilon)$ is infinite. That is, every ball around x contains infinitely many points of A.

You should be careful not to confuse sequences and sets. For example in R, the sequence

$$1, 2, 3, 1, 2, 3, 1, 2, 3, 1, 2, 3, 1, 2, 3, \cdots\cdots\cdot$$

has limit points 1,2,3. However if you consider the terms of the sequence, it is just $\{1, 2, 3, \}$ and this set has no limit point at all.

We can say x is a limit point of A if every ball around x contains at least one point of A other than x. Of course, if x is a limit point then the definition says that in fact there are infinitely many points of A in any ball around x.

Conversely, given that there is at least one point different from x we can show that there are actually infinitely many points of A as follows. Let $\epsilon > 0$ be given. Pick x_1 in this ball different from x. Note that $x_1 \neq x$ tells $d(x_1, x) > 0$, take the ball of radius $d(x, x_1)/2$ around x and pick one point x_2 in this ball different from x. Pick x_3 different from x in the ball of radius $d(x, x_2)/2$ around x. continue to see infinitely many points of A in the given ball.

We can also say x is a limit point of A iff there is a sequence of distinct points in A converging to x. That is, there is a sequence (x_n), each x_n is in A; $x_n \neq x_m$ for $n \neq m$ and $x_n \to x$. If this happens then the points are distinct tells that each $B(x, \epsilon) \cap A$ is infinite. Conversely, if x is limit point of A, then you can choose inductively for $n \geq 1$ a point $x_n \in B(x, 1/n) \cap A$ different from previous points. This will do.

We can also say x is a limit point of A iff there is a sequence (x_n) of points in A such that $x_n \neq x$ for each n and $x_n \to x$. If x is a limit point, then the previous para gives you a sequence (x_n) of distinct points in A converging to x. Clearly, at most one of them could be x, remove it.

Conversely, if there is a sequence as described above, set $y_1 = x_1$. since all the points of the sequence are different from x, $d(x_1, x) > 0$. Take a point x_{n_1} of the sequence in the ball of radius $d(x, x_1)/2$ around x. Obviously, this is different from x_1, Name it y_2. Now take the ball of radius $d(x, y_2)/2$ and pick x_{n_2} here with $n_2 > n_1$ and continue. This gives a sequence of distinct points from A converging to x.

open sets, closed sets:

A subset U is open if $x \in U$ implies there is $\epsilon > 0$ such that $B(x, \epsilon) \subset U$. that is, whenever there is a point in the set, a ball around that point is contained in U. A subset C is closed if its complement is open.

If you specialize to real line, as already seen earlier, open ball of radius ϵ centred at a is just the interval $(a - \epsilon, a + \epsilon)$. Thus a set is open if: whenever there is a point in the set, a small interval around that point is contained in the set. This is precisely the definition we adapted last year.

Similarly, when $X = R^2$ and $d_2(x, y)$ is the Euclidean distance, then ball of radius $\epsilon > 0$ centered at a point $a = (a_1, a_2)$, is usual geometric disk and thus again this above definition coincides with what we adapted in case of R^2 as well. The above definition coincides with the definition we adapted last year in R^n.

Returning to general metric spaces, we can say C is closed iff it contains all its limit points. To see this let C be closed, need to show no point outside C can be a limit point of C. But C being closed, C^c is open, so if you take a point $x \in C^c$, there is a ball around x contained in C^c and this ball does not contain any point of C so that x can not be a limit point of C.

Conversely, let C contain all it limit points. Let $x \in C^c$. So it is not a limit point of C. So there is a ball around it which does not contain any point of C, except possibly x. But x anyway is not in C. So this ball has no point from C, in other words, every point of C^c has a ball around it contained in C^c. Thus C^c is open and hence C is closed.

Connected spaces:

In a metric space (X, d) the sets X and $\emptyset$ are always open and they are also closed. Are there any other sets which are both open and closed? There may not be.

For example, take $X = R$ and $d(x, y) = |x - y|$, usual metric. Then there is no other set which is both open and closed. Indeed, suppose there is such a set A. Since $A \neq \emptyset$, pick $x \in A$; since $A \neq R$, pick $y \in A^c$. Assume $x < y$. Let $s = \sup\{z \in A : z < y\}$. This sup is sensible because the set is not empty — x is in it; and bounded above — y is a bound. If $s \in A$ then, $y \notin A$ tells $s < y$. Now A is open tells there is a small interval $(s - \epsilon, s + \epsilon) \subset A$ contradicting s is sup of the set. If $s \in A^c$, then A^c open says an interval as earlier is contained in A^c. Now, s being sup of our set nothing above s is in

A and this interval tells us that actually nothing above $s - \epsilon$ is in A again contradicting s is sup of the set.

Similar argument applies if $y < x$.

On the other hand if we take $X = R$ with the metric $d(x, y)$ to be zero or one according as $x = y$ or not, you see that every singleton set is open (ball of radius $1/2$ around that point is contained in it!). This shows that every subset is open. As a consequence every subset is both open and closed.

Definition: A metric space (X, d) is connected if the only subsets which are both open and closed are $\emptyset$ and X. Otherwise, the space is said to be disconnected.

The plane R^2 with usual metric is connected. In fact proceed as above, pick x and y and concentrate on the line joining x and y and arrive at a contradiction. Same argument shows that R^n is connected too. We shall return to connected sets later again.

complete spaces:

One concept we want to imitate is that of a Cauchy sequence. A sequence (x_n) in a metric space (X, d) is Cauchy sequence if given any $\epsilon > 0$, there is an N such that $d(x_n, x_m) < \epsilon$ for every $m, n > N$. A metric space is complete if very Cauchy sequence converges.

Thus a Cauchy sequence is a sequence of points which are coming closer and closer. After some stage distance between any two points is at most one; after a later stage distance between any two points is at most $1/2$ and so on. Thus completeness means that any sequence of points which are coming closer and closer are *actually* coming closer to a point. Every Cauchy sequence is heading somewhere (not falling out of a hole!).

This property of completeness is very important, it helps us to discover points in the space which may not be visible to the naked eye. For example in the real line we found that the sequence of points

$$1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \cdots + \frac{1}{n!}; \quad n = 1, 2, 3, 4 \cdots$$

is a Cauchy sequence and hence converges. it was not any point we knew.

So we named it e. similarly the sequence of points

$$1+\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{n}-\log n; \quad n=1,2,3,4\cdots$$

is a Cauchy sequence, and its limit is a number we did not see earlier and named it γ (Euler's constant).

Now you have a panorama of metric spaces, you can discover points in the spaces, that you are not able to see outright. For example we have the space $C[0,1]$ with sup metric. You can discover that there are elements in this space which are nowhere differentiable; that is, there are nowhere differentiable continuous functions. In fact you will see later that such functions are far more than differentiable functions.

Example: $X = R$, $d(x,y) = |x-y|$ is a complete metric space.
We knew this last year.

Example: $X = (0,1]$, $d(x,y) = |x-y|$. the space is not complete. Indeed $(1/n)$ is a Cauchy sequence that does not converge to any point in X. Let us consider on the same space the following metric

$$d_1(x,y) = |x-y| + \left|\frac{1}{x} - \frac{1}{y}\right|$$

First of all note that, if $x_n \to x$ in this new metric d_1, then in particular, $|x_n - x| \to 0$ so that $x_n \to x$ in the metric d. Conversely, if $x_n \to x$ in the old metric d, then everything being in the space X, we conclude that $(1/x_n) \to (1/x)$ so that $x_n \to x$ in d_1.

Thus the notion of convergence is same under both metrics, this implies (easy to see) that closed sets are same in both and hence open sets are also same in both.

Interestingly enough the space X is complete with metric d_1. In fact if (x_n) is d_1-Cauchy then observe that both

$$|x_n - x_m| \quad \text{and} \quad \left|\frac{1}{x_n} - \frac{1}{x_m}\right|$$

converge to zero as n and m become large. In particular, the first one tells that $x_n \to x$ in R, and the second one tells that this x can not be zero (then $1/x_n$ becomes unbounded). Thus $x \neq 0$ and hence $1/x_n \to 1/x$ and hence

$d_1(x_n, x) \to 0$. In other words every sequence which is cauchy in d_1 does converge in d_1.

Thus the space (X, d_1) is complete where as (X, d) is not; though both have the same notion of convergence and same closed sets and same open sets. What happened is that non-convergent sequences which are Cauchy in d have been destroyed in d_1, they are no longer Cauchy and hence not obliged to converge.

Here is another example.

$$X = (0, 1); \qquad d(x, y) = |x - y|;$$

$$d_1(x, y) = |x - y| + \left|\frac{1}{x} - \frac{1}{y}\right| + \left|\frac{1}{1-x} - \frac{1}{1-y}\right|$$

Both d and d_1 give the same notion of convergence, closed sets etc. But (X, d) is not complete whereas (X, d_1) is complete. Sequences like $(1/n)$ or $\{n/(n+1)\}$ which are Cauchy in d but not converging are no linger Cauchy in d_1 and hence not obliged to converge.

This discussion should tell you that completeness is something that depends heavily on the metric. This is not surprising because, the notion of Cauchy sequence depends on the metric. However, the concept of convergence does not heavily depend on the metric; it depends on the collection of open sets. Of course, at this stage you might say: so what? open sets depend on the metric. In a later course on topology you will see the distinction.

If a metric space is not complete, is there anything that we can do to make it complete. Of course, one thing we can do is to destroy Cauchy sequences which are not converging as happened in the examples above and make the space complete.

Suppose that there are too many Cauchy sequences and we can not destroy all of them. Or imagine that we do not want to change the metric. Can we do anything to make the space complete? In such a case the only possibility is to attach new points to the space and declare them to be limits of Cauchy sequences (which had no limits at present). Remember, the space is not complete due to shortage of points in the set, there are Cauchy sequences but no points to which they converge.

Yes, Cantor discovered a method of completing a metric space. I could give you the general method and then say you can specialise this process to the set of rational numbers to obtain another construction of the real number system. However this will lead to some confusion due to a technical problem. So we shall first execute cantor's idea to construct R.

Cantor's construction of R:

Let us once again pretend that we do not have real numbers. However we do have the set of rationals Q before us. We are constructing R, following Cantor.

Well, why is Q not a model of real number system? It satisfies all the axioms except the least upper bound axiom. There are sets which are bounded above but have no supremum. In Dedikind's construction, we attached points to Q so that such sets have supremum.

Following our earlier observations, we could also have said that Q is not a model of real number system because there are Cauchy sequences which are not converging. We shall now enlarge the space so that such sequences converge. Again remember the general philosophy: if certain phenomena are not taking place, consider all such phenomena.

To make the above statement concrete and bring right perspective, it is worth recalling some of our past actions. If we could not subtract 4 from 3 we considered pair $(3, 4)$ as a point of our space and, in a sense, this pair represented $3 - 4$. Since we felt $3 - 4$ should be same as $7 - 8$ we identified the two pairs $(3, 4)$ and $(7, 8)$. Considered the space of equivalence classes.

If we could not divide 5 by 7 we considered the pair $(5, 7)$ and this represented $5/7$. We felt that this should be same as $10/14$ and so identified the two pairs $(5, 7)$ and $(10, 14)$. We considered the space of equivalence classes.

If the set S of all rational numbers whose square is less than 2 has no supremum, then we considered this set itself as a 'point' and declared this point as supremum of S.

[Let me add a comment in passing which you may ignore. What we actually did was a little different, we considered not just the above set S but actually the set T which consists of all negative rational numbers along with the set of positive rational numbers whose square is less than 2. This set T gives a Dedikind cut but S does not give you a cut. If you want to know why we did this, the answer is simple: Situation was hopeless and we brought some order. If you understood the above three paras, and if you are on imitating spree, then your tendency would be to consider *all* subsets of Q which are bounded above. You will identify two sets if they *appear* to have same sup and consider the collection of equivalence classes. Yes, but luckily enough, it so happens that each equivalence class has *exactly one cut* in it. Thus though I was looking at the equivalence class, I showed you only the cut from that class and *managed* matters. This helped you to consider *one set* rather than *equivalence class of subsets of* Q. Think about it.]

Returning to our present problem, you would have no hesitation considering all Cauchy sequences. Let R_0 be the collection of Cauchy sequences $x = (x_n)$ of rational numbers. We identify

$$x = (x_n) \sim y = (y_n) \leftrightarrow d(x_n, y_n) \to 0.$$

Triangle inequality helps to show that this is an equivalence relation.

Let R be the space of equivalence classes. This is our real number system. equivalence class containing a sequence x is denoted $[x]$.

We define addition by taking $[x]+[y] = [x+y]$. Thus to define sum of two equivalence classes $[x]$ and $[y]$, take one sequence from each class and add (term by term) those two sequences and the class containing this resulting sequence is the sum. This is a good definition. Firstly, the sum of two Cauchy sequences is cauchy. Secondly, it does not depend on which sequence we choose from the equivalence class. Indeed, let $a \in [x]$ and $b \in y]$ then we show $a + b \in [x + y]$ as follows.

$$d(a_n + b_n, x_n + y_n) \leq d(a_n, x_n) + d(b_n, y_n) \to 0.$$

The rules $[x] + [y] = [y] + [x]$ as well as $[x] + ([y] + [z]) = ([x] + [y]) + [z]$ are easy being consequences of similar rules regarding rational numbers.

The zero element is $[\theta]$ where θ is the sequence having zero for all its terms. It is clear $[x] + [\theta] = [x]$. The inverse is also clear, namely, for a given element $[x]$ we take $[-x]$ where $-x$ is the sequence $(-x_n)$.

We shall denote the additive identity by $[0]$.

We define multiplication also in the obvious way. $[x][y] = [xy]$ where $x.y$ is the term wise multiplication of the sequences. To see that this is a good definition, we need to show that xy is a Cauchy sequence and if someone takes different elements $a \in [x]$ and $b \in [y]$ then $ab \in [xy]$.

First recall that a Cauchy sequence is bounded. Indeed, if (x_n) is Cauchy, pick N so that $|x_m - x_n| < 1$ for $m, n \geq N$. In other words for $n \geq N$ we have

$$|x_n| \leq |x_N| + |x_n - x_N| \leq |x_N| + 1.$$

Thus maximum of the finitely many quantities,

$$|x_1|, |x_2|, \cdots |x_{N-1}|, |x_N| + 1$$

will serve asa bound.

To see $(x_n y_n)$ is Cauchy, first pick some bound M for the two Cauchy sequences x and y.

$$|x_n y_n - x_m y_m| \leq |x_n||y_n - y_m| + y_m|x_n - x_m|$$

$$\leq M|y_n - y_m| + M|x_n - x_m|.$$

Given rational $r > 0$ we can choose N such that $|x_m - x_n| < r$ and $|y_m - y_n| < r$ for all $m, n > N$. which can be made small for all large m, n.

Finally, let $(a_n) \sim (x_n)$ and $(b_n) \sim (y_n)$. Shall show $(a_n b_n) \sim (x_n y_n)$. Fix a bound M for all these four Cauchy sequences.

$$|a_n b_n - x_n y_n| \leq |a_n||b_n - y_n| + |y_n||a_n - x_n|$$

$$\leq M|b_n - y_n| + M|a_n - x_n|$$

which can be made small.

The rules $[x][y] = [y][x]$ and $([x][y])[z] = [x]([y][z])$ re easy. The element $[1]$, the equivalence class containing the constant sequence one is the identity element. $[1][x] = [x]$ is clear from definition.

Let $[x] \neq [0]$. We shall show inverse. We start with an observation. since $[x] \neq 0$ we conclude that $|x_n| \not\to 0$. So there is an $r > 0$ such that for infinitely many n we have $|x_n| > r$. since (x_n) is cauchy, there is N such that $|x_n - x_m| < r/2$ for $n, m > N$. If you take any $n > N$ then $|x_n| > r/2$. Indeed, let $n > N$ and pick $m > N$ so that $|x_m| > r$.

$$|x_n| \geq |x_m| - |x_n - x_m| = r - \frac{r}{2} = \frac{r}{2}.$$

Thus for all $n > N$ we have $|x_n| > s$ where $s = r/2 > 0$.

Let (y_n) be the sequence $y_n = 1$ for $n \leq N$ and $y_n = 1/x_n$ for $n > N$. Since $x_n y_n = 1$ for all $n > N$, we see that $[x][y] = [1]$. But we need to convince ourselves that y is a Cauchy sequence. This is immediate because for $m, n > N$

$$|y_m - y_n| = \frac{x_m - x_n|}{|x_m||x_n|} \leq |x_m - x_n|/s^2.$$

which can be made small.

This completes the proof that $[y] \in R$ and is the inverse of $[x]$.

The distributivity $[x]([y] + [z]) = [x][y] + [x][z]$. is immediate.

S o far it is smooth, We now need to define order on R. The simple minded definition $[x] \leq [y]$ if $x_n \leq y_n$ is not meaningful, because this order does not respect the equivalence classes. For example,

$$(2, 1, 1, 1, , 1 \cdots) \quad (0, 1, 1, 1, 1, 1 \cdots)$$

are in the same class. In fact if you change only finitely many terms of a Cauchy sequence then the resulting sequence is equivalent to the original sequence. Thus we should not make the definition depend on all suffixes.

Suppose we modify to say that $[x] < [y]$ if $x_n < y_n$ for all large values of n, then again we run into problems. If $x_n = 1/n$ and $y_n = 2/n$ then $x_n < y_n$ holds for all values of n but they represent the same element $[0]$.

Thus defining order is a little delicate. In stead of defining $\leq$ we start defining $<$. Say $[x] < [y]$ if there is $r > 0$ and N such that $x_n + r \leq y_n$ for all $n > N$. We say $[x] \leq [y]$ iff $[x] = [y]$ or $[x] < [y]$. This definition satisfies our requirements.

Let us start showing that this definition respects the equivalence classes. Let $x \sim a$ and $y \sim b$. Suppose that there is $r > 0$ and N such that $x_n + r \leq y_n$ for all $n > N$. We exhibit $s > 0$ and M such that $a_n + s \leq b_n$ for all $n > M$. Take $s = r/2$. Take $M > N$ so that

$$n > M \Rightarrow |x_n - a_n| < r/4; \quad |y_n - b_n| < r/4.$$

If now $n > M$ then

$$a_n + \frac{r}{2} \leq x_n + \frac{r}{4} + \frac{r}{2} \leq y_n - r + \frac{r}{4} + \frac{r}{2} \leq b_n + \frac{r}{4} - r + \frac{r}{4} + \frac{r}{2}.$$

We need to show that it is a linear order and is friendly with addition and multiplication.

Cantor Construction of Reals:

So far we have:

• R_0, the set of Cauchy sequences of rational numbers.

• R the set of equivalence classes, where two Cauchy sequences x, y are equivalent if $d(x_n, y_n) = |x_n - y_n| \to 0$.

• $[x] + [y] = [x + y]$; zero element is the equivalence class, still denoted by 0, containing the Cauchy sequence: $\{0, 0, 0, 0 \cdots\}$ — all terms of the sequence are zero. of course this class contains other sequences too, for example

$$(a_n = 1/n); \quad (b_n = -1/n); \quad (c_n = 1/n^{19}); \quad (d_n = 1/n!) \cdots\cdots;$$

but NOT the sequence $(\exp\{-n\})$ or $(1/\sqrt{n})$ and so on simply because these are not sequences of rational numbers.

• $[x] \cdot [y] = [xy]$; unit element is the equivalence class containing the sequence: $\{1, 1, 1, \cdots\}$ — all terms equal to one. This class contains other sequences too, for example, you can add any of the above sequences to this sequence.

• R is a field with above operations.

• $[x] < [y]$ if there is a rational $r > 0$ and N such that $x_n + r \leq y_n$ for all $n > N$. $[x] \leq [y]$ if either $[x] = [y]$ or $[x] < [y]$.

We shall show that $\leq$ is a linear order.

(i) $[x] \leq [x]$.

This is clear because $[x] = [x]$ and see the definition of $\leq$.

(ii) If $[x] \leq [y]$ and $[y] \leq [x]$ then $[x] = [y]$.

In the two hypotheses, if equality holds somewhere then there is nothing to be proved. We now show that strict inequality at both places is not possible, this will complete the proof.

If possible $[x] < [y]$ and $[y] < [x]$. Get one N and $r > 0$ and $s > 0$ such that

$$n > N \Rightarrow x_n + r \leq y_n; \quad y_n + s \leq x_n.$$

This is impossible because if we take any one $n > N$ we should have

$$x_n + r + s \leq y_n + s \leq x_n; \quad r + s > 0.$$

We know enough about rationals.

[irrelevant discussion: In the first para, why did he say 'if equality holds in some hypothesis'; can equality hold in one hypothesis and strict inequality hold in the other hypothesis. Such doubts are natural to arise, if you are thinking. But you need not worry. As far as our proof is concerned this question has no consequence. We have proved our claim. It is all that matters. You can, of course, return to your doubt; yes, you are right.]

(iii) $[x] \leq [y]$ and $[y] \leq [z]$ implies $[x] \leq [z]$.

Repeat above argument to show, with obvious notation, $x_n + r + s \leq z_n$ for $n > N$.

(iv) Either $[x] \leq [y]$ or $[y] \leq [x]$ holds.

Suppose that both $[x] < [y]$, $[y] < [x]$ fail. We show that $[x] = [y]$ holds. Start observing that the assumption implies

$$(\forall r > 0)\ (\forall N)\ (\exists n > N)\ x_n + r > y_n; \quad (\spadesuit)$$

$$(\forall r > 0)\ (\forall N)\ (\exists n > N)\ y_n + r > x_n. \quad (\clubsuit)$$

We need to show that $|x_n - y_n| \to 0$. Fix any $r > 0$. Remembering that x and y are Cauchy sequences, fix N so that

$$n, m \geq N \Rightarrow |x_n - x_m| < r/4; \quad |y_n - y_m| < r/4.$$

To complete proof, we shall show $|x_n - y_n| < r$ for $n > N$. First, for this N and $r/4$, using $(\spadesuit)$, $(\clubsuit)$ fix $i > N$ and $j > N$ so that

$$x_i + \frac{r}{4} > y_i; \quad y_j + \frac{r}{4} > x_j;$$

The subtle point is that at this stage we do not know if i and j are same in the inequalities above.

Now take any $n \geq N$. Then

$$x_n \leq x_j + \frac{r}{4} \leq y_j + \frac{r}{4} + \frac{r}{4} \leq y_n + \frac{r}{4} + \frac{r}{4} + \frac{r}{4}$$

and

$$y_n \leq y_i + \frac{r}{4} \leq x_i + \frac{r}{4} + \frac{r}{4} \leq x_n + \frac{r}{4} + \frac{r}{4} + \frac{r}{4}$$

as promised.

(v) For every pair exactly one of the following hold: $[x] < [y]$ or $[x] = [y]$ or $[y] < [x]$.

This is already in the above arguments.

(vi) The order is friendly with addition and multiplication. $[x] < [y]$ implies $[x] + [z] < [y] + [z]$. Also $[0] < [x]$, $[0] < [y]$ imply $[0] < [x][y]$. These follow from definition of order.

We shall now identify Q as a subset of R. Define for $q \in Q$, $\varphi(q)$ to be the equivalence class containing the constant sequence $\{q, q, q, q, q, \cdots\}$. This map is one-to-one, respects (?) addition, multiplication and also order. We simply think of Q as a subset of our R. This needs to be said because, after all, elements of R are *not* rational numbers; they are *not even* Cauchy sequences of rational numbers; they are bags where each bag contains a collection of Cauchy sequences.

We make a useful observation. Given $[x] < [y]$ there is a rational q such that $[x] < q < [y]$. Here is the proof. Fix $r > 0$ and N such that

$$n \geq N \Rightarrow x_n + r \leq y_n$$

If necessary by taking a larger N, we can also assume that

$$n, m \geq N \Rightarrow |x_n - x_m| < \frac{r}{4}.$$

Let

$$q = x_N + \frac{r}{2}.$$

We show this will do. Observe that this is a rational (our sequences are sequences of rational numbers). Let $n > N$.

$$q + \frac{r}{4} = x_N + \frac{r}{2} + \frac{r}{4} \leq x_n + \frac{r}{4} + \frac{r}{2} + \frac{r}{4} \leq y_n$$

showing that $q < [y]$. Also

$$x_n + \frac{r}{4} \leq x_N + \frac{r}{4} + \frac{r}{4} = q$$

showing $[x] < q$.

To complete our construction, we show now that every non-empty subset of R which is bounded above has a supremum. We shall repeat the arguments used in the Dedikind construction. To avoid giving any wrong impression let me add, it is not as though Cantor borrowed the Dedikind construction, they both did around the same time.

Thus let S be a non-empty set bounded above. Take any upper bound $[a]$ and get a rational q such that $[a] < q < [a] + 1$. Similarly if $[x] \in S$, take a rational p such that $p < [x]$. Here we are using the observation made above. And also we identify rational with the constant sequence or more precisely, the class containing the constant sequence.

The upshot of what we did above is to get $p_0 = p$ and $q_0 = q$ so that
(i) $p_0 < q_0$ and
(ii) there exists $[x] \in S$, $p_0 \leq [x]$ and $[x] \leq q$ for all $[x] \in S$.
In words, there are points of S at least as large as p but nothing that exceeds q.

We shall now construct rationals p_n, q_n for $n \geq 1$ such that

(i) $p_n < q_n$. Further, one of these numbers is $(p_{n-1} + q_{n-1})/2$ and the other belongs to $\{p_{n-1}, q_{n-1}\}$.
In other words one of the new numbers is same as the earlier one, and the other is average of the two earlier ones.

(ii) $\exists [x] \in S; p_n \leq [x]$ and $\forall [x] \in S; [x] \leq q_n$.

This is easy. For example consider $r_0 = (p_0 + q_0)/2$. if for all $[x] \in S$ we have $[x] \leq r_0$ then declare $p_1 = p_0$ and $q_1 = r_0$. Otherwise declare $p_1 = r_0$ and $q_1 = q_0$. In general consider the midpoint of the previous two points and proceed.

Clearly, the construction shows the next pair of points are in between the existing pair. Thus

$$p_0 \leq p_1 \leq p_2 \cdots\cdots \leq \cdots\cdots \leq q_2 \leq q_1 \leq q_0.$$

$$q_n - p_n = \frac{q_{n-1} - p_{n-1}}{2}.$$

These in turn show that
(p_n) is a Cauchy sequence and (q_n) is a Cauchy sequence.
$q_n - p_n \to 0$, that is $(p_n) \sim (q_n)$.
As a consequence we can define an element of R by

$$s = [(p_n)] = [(q_n)].$$

We now show that

(i)If $s < [x]$ then $[x] \notin S$; that is, $[x] \in S \Rightarrow [x] \leq s$.
This shows that s is an upper bound of S.

(ii) $[a] < s \Rightarrow \exists [x] \in S; [a] < [x] \leq s$.
This shows that nothing smaller than s will serves as upper bound.

This will then complete the proof.

Proof of (i):

Since $s = [(q_n)] < [x]$, fix $r > 0$ such that $q_n + r \leq x_n$, say for $n \geq N_1$. Since (q_n) is cauchy, we have $|q_m - q_n| \leq r/2$ for $n \geq N_2$. Taking N larger than both N_1 and N_2, we see that for $n > N$

$$q_N + \frac{r}{2} \leq q_n + \frac{r}{2} + \frac{r}{2} \leq x_n$$

In other words $q_N < [x]$ But by construction nothing larger than q_N is in S. Thus $[x] \notin S$.

Proof of (ii):

Use similar argument as above to see that there is N such that $[a] < p_N$ and note that there are points of S at least as large as p_N.

[It is tempting to say that $q_n \to s$; so given $s < [x]$ there is N such that $s \leq q_N < x$. But unfortunately at this moment our vocabulary is limited, we do not know convergence that well. This can however be made rigorous and then used.]

This completes Cantor's construction of R.

I am reminded of a poem.

A centipede was happy - quite;
until a toad, in fun, said
"prey, which leg moves after which"
This raised her doubts to such a pitch
She fell exhausted in the ditch
not knowing how to run.

You should not continue thinking (following Cantor): Aha, I now know real numbers; a real number is a Cauchy sequence of rational numbers. This leads to utter confusion when you later think of sequences of real numbers;

because this would then mean a *sequence of Cauchy sequences* of rational numbers! — not a happy thought.

Or you should not think (following Dedikind): Aha, I know real numbers, real numbers are cuts in Q. This leads to unnecessary confusion when you consider subsets of R. This would then mean *a set of subsets* of Q — not a happy thought.

Then why did we do all this. This gives you practice in handling mathematical objects. This leads to certain confidence. This also familiarizes with mathematical arguments and mathematical proofs. There are at least three other reasons why you should appreciate all this.

Firstly, it is very important for us to know whether there is a system at all satisfying the axioms laid down for R. Once we are sure that such a system exists, we just work with a system satisfying these rules — no matter how such a system is arrived at; either by Dedikind or Cantor method or any other third method. Real numbers attain independent existence irrespective of who constructs or how he/she constructs.

It is just like locating a house. It may be to the right of a hotel or to the left of a building or opposite a shop and so on. Once the house is discovered all these other pointers, hotel, building, shop etc are irrelevant. They might even be confusing.

Second reason is vey important. If you know how to construct a house, you can do very well in the construction business. For example, a customer might want a house with some interesting properties; you can think a little bit with your expertise and make it. Here is one concrete example. I had different example in mind, but after discussion with Uma, I realized this may be more appropriate for you.

I have two vector spaces V and W. Suppose we wanted to multiply vectors in W with vectors in V. We also want again a vector space — let us pretend our vector spaces are over R (it makes no difference). Let us not go into the reasons why such a thing is needed. Can you help? You have the expertise now. You only need to think.

We want to multiply: $v \cdot w$ where $v \in V$ and $w \in W$. Consider such 'things'. But if I write this dot, you would ask me what is the dot. So I say consider (v, w), you accept because you know ordered pairs and the set $V \times W$. So I show you (v, w) but at the back of my mind I have $v \cdot w$. But then this set by itself, is not a linear space. We wanted a linear space.

So I say consider finite linear combinations of the above things, like

$$4\, v_1 \cdot w_1 \;+\; \frac{32}{9}\, v_2 \cdot w_2 \;-\; \sqrt{3}\, v_3 \cdot w_3.$$

or equivalently,

$$4\,(v_1, w_1) + \frac{32}{9}\,(v_2, w_2) - \sqrt{3}\,(v_3, w_3). \quad (\spadesuit)$$

Since 'linear combination of linear combinations' is again a grand linear combination of original things, there is hope that such linear combinations make a linear space. But you would again raise question: what is this plus sign? So I say consider the function f on $V \times W$ defined by

$$f(v_1, w_1) = 4; \quad f(v_2, w_2) = \frac{32}{9}; \quad f(v_3, w_3) = -\sqrt{3}; \quad (\clubsuit)$$

and $f(v, w) = 0$ for other pairs. You do not object to this, you know functions and the above is a legitimate function on $V \times W$. Thus I show you $(\clubsuit)$ but at the back of my mind I have $(\spadesuit)$.

But there is one problem. For example, consider the following. Take $v_1, v_2 \in V$ and $w \in W$. Put $v = v_1 + v_2$. My mind tells me $v \cdot w$ is same as the linear combination $v_1 \cdot w + v_2 \cdot w$. Or to use the 'things I am showing you' I should identify the two functions:
$f(v, w) = 1$ and f is zero for other pairs.
$g(v_1, w) = 1 = g(v_2, w)$ and g is zero for other pairs.
I identify f and g. Thus I consider the collection of functions f on $V \times W$ (which are zero outside a finite set) and define an equivalence relation — well thought out and driven by what we all feel — on this set. This *exactly* meets our demands.

The resulting house goes by the name of tensor product.

The third reason is the importance of understanding symbols without getting confused. We use some symbols for which you would have no objection but while we use the symbol, we have some thing at the back of our mind. Thus, what is at the back of our mind is more important than the symbol we are using. Think about it.

back to open and closed sets:

Before returning to general metric spaces, let us observe something special about open sets in the real line. Let $U \subset R$ be an non-empty open set. For every point $x \in U$ there is an open interval $(x-\epsilon, x+\epsilon) \subset U$. Obviously U is union of all these intervals. However we can be more specific.

Every nonempty open subset of R is a countable disjoint union of non-empty open intervals in a unique way. We shall prove this now.

Let us start clarifying meanings of the terms. Interval means a subset with the following property: if two points are there, everything in between is also there. More precisely, $A \subset R$ is an interval if

$$x < z < y; \quad x, y \in A \Rightarrow z \in A.$$

Suppose A is an interval. Let $a = \inf A$ and $b = \sup A$. if A is not bounded below then we take this inf to be $-\infty$; if A is not bounded above we take the sup to be ∞.

It is easy to show that, for example when the above inf and sup $a, b \in R$ then the interval A must either be $[a, b]$ or $[a, b)$ or $(a, b]$ or (a, b). If it is an open interval, that is, if it is an interval which is an open set then it must be (a, b).

Similarly, when $a \in R$ but $b = \infty$ then the interval A must be either $[a, \infty)$ or (a, ∞). If it is an open interval, then it must be (a, ∞). Similar remark applies when $a = -\infty$ and $b \in R$.

Of course, union of two intervals need not be an interval. We claim that union of two intervals which have a point in common, is again an interval. Indeed let I and J be intervals and z be in both. Let $a < b$ be two points in the union, say $a \in I$ and $b \in J$. If $a < b < z$ then both a and z are in I and hence so is everything in between, in particular, everything in between a and b is in I and hence in $I \cup J$.

If $z < a < b$, then both z and b are in J and hence so is everything in between them. In particular everything in between a and b is in J.

Suppose $a < z < b$. Then everything in between a and z is in I, whereas everything in between z and b is in J. Thus everything in between a and b is in $I \cup J$.

Thus union of two intervals which have a point in common is again an interval.

Returning to our problem, take $x \in U$. By above observation, the union of all intervals I such that $x \in I \subset U$ is itself an interval. Call it I_x. Suppose $a \in R$ and is an end point of I_x, then we claim that $a \notin I_x$. Because if $a \in I_x \subset U$, then there must be an interval $(a-\epsilon, a+\epsilon) \subset U$. But then $I_x \neq I_x \cup (a-\epsilon, a+\epsilon)$ is an interval because a is in both' Also this union is a subset of U because both are so. Also this interval includes x. This contradicts that I_x is the union of all such intervals.

Thus I_x is an open interval.

Actually, I_x is the largest open interval J with the property $x \in J \subset U$. There is nothing to prove here if you look at the definition of I_x.

Thus for every $x \in U$ we have an open interval I_x such that $x \in I_x \subset U$, largest such interval. We now claim that if $x \neq y$ then either $I_x = I_y$ or $I_x \cap I_y = \emptyset$. This is clear because if the intersection is non-empty, then their union is an interval contained in U, includes both x and y.

Let $\mathcal{I}$ be the collection of the *distinct intervals* of the family $\{I_x; x \in U\}$. Since any two intervals in $\mathcal{I}$ are disjoint and all are non-empty we conclude that this family $\mathcal{I}$ must be countable. Take one rational from each to see this.

Thus we can enumerate the family $\mathcal{I}$ as a sequence. Thus

$$U \ = \ I_1 \ \cup \ I_2 \ \cup \ I_3 \ \cup \ \cdots .$$

(finite or infinite) union of disjoint non-empty open intervals.

Suppose that

$$U \ = \ J_1 \ \cup \ J_2 \ \cup \ J_3 \ \cup \ \cdots .$$

union of disjoint non-empty open intervals. We show that the intervals are exactly the same, perhaps enumerated in a different order.

Let $x \in U$. Let $J_k = (a, b)$, say, contains the point x. We argue that J_k is indeed I_x. First observe the following. if a is finite then $a \notin U$. Because, if it were in U, then it must be in one of the other J, but then that J must contain an interval around a, but then such an interval intersects J_k. Remember the intervals J_i are disjoint. Similarly if b is finite then $b \notin U$.

Since $x \in J_k \subset U$ and I_x is maximal such interval we conclude that $J_k \subset I_x$. If I_x is strictly larger than J_k, then an end point of J_k must be finite and must be in I_x. But any finite end point is shown to be not in U. This shows $J_k = I_x$.

This completes the proof.

collection of open sets:

Let (X, d) be a metric space. We defined a subset U to be open if $x \in U$ implies there is an $r > 0$ such that

$$B(x, r) \ = \ \{y : d(x, y) < r\} \subset U.$$

The collection of open sets has the following properties:

(i) As already noted, $\emptyset$ and X are open.

(ii) Union of any collection of open sets is open. This is because, if a point x is in the union it is then in one of the sets which itself already contains a $B(x, r)$.

(iii) Finite intersection of open sets is open. Indeed, let U and V be open. Let $x \in U \cap V$. Then $x \in U$ and $x \in V$ so that there is $r > 0$, and $s > 0$ such that $B(x, r) \subset U$ and $B(x, s) \subset V$. If we take $p = \min\{r, s\}$ then easy to see that $B(x, p) \subset U \cap V$. This being true for every point in the intersection we are done.

Accordingly, we see that $\emptyset$ and X are closed; intersection of any collection of closed sets is a closed set; finite union of closed sets is closed.

We claim that the ball $B(x, r)$ defined above is an open set. So we are justified in calling it open ball. To see this let $x_0 \in B(x, r)$, say

$$d(x, x_0) = \alpha < r; \qquad s = r - \alpha > 0.$$

We claim that $B(x_0, s) \subset B(x, r)$. Indeed

$$y \in B(x_0, s) \Rightarrow d(x, y) \leq d(x, x_0) + d(x_0, y) < \alpha + (r - \alpha) = r.$$

$$\Rightarrow y \in B(x, r).$$

Similarly $C = \{y : B(x, y) \leq r\}$ is a closed set and hence we are justified in calling it closed ball. To see that it is closed, we only need to observe that if $y_n \to y$, and $y_n \in C$ for all n, then

$$d(x, y) = \lim d(x, y_n) \leq r.$$

Thus open ball is indeed an open set and closed ball is a closed set.

interior and closure:

If a set is not open how do we relate it to an open set. we can take the largest open set contained in it. This is called interior.

Let (X, d) be a metric space and $A \subset X$. We define A^o, called interior of A, to be the largest open set contained in A. This makes sense because union of open sets is open. Thus A^o is the union of all open sets contained in A. equivalently, it is the union of all open balls contained in A. A point of A^o is called an interior point of A.

Closure $\overline{A}$ is the smallest closed set that contains A. This is just the intersection of all closed sets that contain A.

We can also say that

$$\overline{A} \;=\; A \cup \{\text{set of limit points of } A\}.$$

Denote the set on left by S. We know that a set is closed iff it includes all its limit points. Thus every closed set that contains A includes all limit points of A. Hence $\overline{A} \supset S$. Conversely, to show that $\overline{A} \subset S$ we show that S is a closed set containing A. Of course $S \supset A$. To show S is closed, let $x \notin S$. Since it is not limit point of A there is a ball $B(x, r)$ which has no point of A other than possibly x. But since $x \notin A$ we conclude that $B(x, r) \subset S^c$ showing S^c is open.

Compact sets:

We shall now imitate to define and discuss notion of compact sets that we studied in R^n.

Let (X, d) be a metric space. A subset $K \subset X$ is said to be compact if the following happens: Every sequence $(x_n) \subset K$ has a *subsequence* that *converges to a point of* K.

The importance of this notion comes from the fact that in R^n compact sets are precisely closed bounded sets; more importantly, every real valued continuous function on a compact set is bounded and attains its bounds.

For general metric spaces the above characterization is too much to expect (though the consequence concerning continuous functions is still correct with exactly the same proof, as we shall see). In fact, 'bounded' does not mean anything because we can always change the metric to another bounded metric without changing convergence.

Also one needs to justify the choice of the word 'compact'. This word gives the impression that the set is not too much 'spread out'; or given lot of material to cover the set we can use only a little of that material to cover the set. Yes, this interpretation is right.

We shall discuss this concept after the midsem.

countable intersection of open sets:

We have discussed the concept of interior and closure. We shall now discuss another way of relating a given set to open sets. This leads to a nice and useful story. We start with simple question.

Let $\mathcal{I}$ be the set of irrational numbers. It is not an open set.

Is it union of open sets? This is silly, union of open sets is open, so we have answered this question. Is this intersection of open sets? This is also silly, every set is intersection of open sets, namely, intersection of all sets $\{x\}^c$ with $x \in A^c$.

We ask, is $\mathcal{I}$ intersection of countable many open sets? This question is better. The answer is, Yes, $\mathcal{I}$ is intersection of $\{x\}^c$ with x running over rationals.

Good, let us repeat with Q the set of rationals. It is not open. Is it intersection of countably many open sets? the answer is not so immediate now. We show it is *not* a countable intersection of open sets.

The above question appears purely set theoretic in nature. However, it is worth recalling that we already needed its answer last year while discussing continuous functions. Let us recall. We constructed a function f on R for which the set of continuity points are precisely the set of irrational numbers. In other words, if x is an irrational number then f is continuous at x; if x is a rational number then f is not continuous at x. naturally, we asked: can we cook up a function f which is continuous at x when x is rational while discontinuous at x when x is not rational.

Returning to our claim, let if possible let

$$Q = U_1 \cap U_2 \cap U_3 \cap \cdots.$$

where each U_i is open. Let $r_1, r_2, r_3, \cdots$ be an enumeration of all rational numbers. We shall manufacture a sequence of intervals $[a_n, b_n]$ such that the following hold.

(i) $r_n \notin [a_n, b_n]$.

(ii) $[a_n, b_n] \subset U_n$.

(iii) $[a_1, b_1] \supset [a_2, b_2] \supset [a_3, b_3] \supset \cdots$.

(iv) $0 < b_n - a_n \leq 1/2^n$.

Conditions (iii) and (iv) and Cantor's intersection theorem tell us that there is a point z common to all these intervals $[a_n, b_n]$. Condition (ii) tells that this point is in all the U_n. Condition (i) tells that this point can not be any rational. This contradicts that Q is the intersection of all the U_n.

We shall now construct such intervals by induction. Start observing that each U_n includes all rationals and is an open set.

Take a non-degenerate open interval contained in U_1. If it includes r_1 take a smaller interval that excludes r_1. Take non-degenerate closed subinterval of this. If it is large, cut it down and make its length at most one. This is $[a_1, b_1]$.

Since $a_1 < b_1$ there is a rational in between and this rational belongs to the open set U_2 as well. Thus get a non-degenerate open subinterval inside $[a_1, b_1]$ which is contained in U_2. If this interval includes r_2 take a smaller subinterval to exclude it. Take a non-degenerate closed subinterval of this. If it is large, cut it down and make length at most 1/2.

Having got intervals for $n = 1, 2, \cdots, k$ we can write down how to get the $(k+1)$-th interval. The reason that such a thing must be written down, instead of saying 'etc etc' or 'do like this' or 'so on' is the following. You need to convince yourself that this procedure can be repeated *for ever.* What would you do if someone says, no you can not continue like this.

I shall write the inductive step just to make sure you know how to do it. Assume we got the first k intervals satisfying the conditions stated above up to k. (Do you realize how important it is to list the conditions, when you make construction inductively, so that they make sense up to k. After all what I needed earlier was $b_n - a_n \to 0$. So I could have written condition (iv) as $b_n - a_n \to 0$. This would be careless because then saying 'condition (iv) holds up to k' does not make any sense. These subtle points you must pay attention to, at least until you clearly grasp what makes sense and what does not! That is why writing proofs is very important.)

Here then is the inductive step. Since $0 < b_k - a_k$ take a rational between a_k and b_k, then it is in $(a_k, b_k) \cap U_{k+1}$. Thus there is a non-degenerate open interval which is contained in $(a_k, b_k) \cap U_{k+1}$. If this includes r_{k+1} take a non-degenerate subinterval which excludes this point. Take a non-degenerate closed subinterval contained in this interval. If its length does not exceed $1/2^{k+1}$ take it; if it is large, cut it down. The resulting interval is $[a_{k+1}, b_{k+1}]$.

This completes the proof.

Let C be the Cantor set. Suppose that we asked: Is $Q \cup C$ a countable intersection of open sets. Then we can not imitate the above proof. Earlier we could show a non-rational point in $\cap U_n$ by avoiding rationals one-by-one. This was possible simply because Q is a countable set.

Let us ask another question which you feel is unrelated to our present discussion. Can you express R as a union of two disjoint non-empty closed sets? That is

$$R = A \cup B; \quad A \neq \emptyset;\ B \neq \emptyset; \quad A, B \text{ closed, disjoint.}$$

This means A is a non-empty proper closed subset of R — connectedness arguments prevent such a thing.

Very good, let us change it to countable union. Can you express R as a countable union of disjoint non-empty closed sets? That is,

$$R = \cup A_n; \quad (\forall n) A_n \neq \emptyset; \quad (\forall n) A_n \text{ closed ;} \quad (\forall n \neq m) A_n \cap A_m = \emptyset.$$

The answer is not immediate. There is an interesting story behind this discussion. The beauty unfolds slowly.

But for now, let us just realize that what we proved above is something that actually proves a better thing.

small sets:

A closed subset $C \subset R$ is said to be small if it does not include a (non-empty) open interval.

Thus for example a singleton set is small. Set of integers is small. The Cantor sets is small. However the interval $[0, 1]$ is not small.

We say that $A \subset R$ is small if its closure $\overline{A}$ is small. That is, there is no open interval (non-empty) contained in $\overline{A}$.

For example every subset of the Cantor set is small, because its closure is contained in C. Set Q of rational numbers is not small, because its closure is all of R. However Q is a countable union of small sets, namely, singleton sets. Here then is a nice theorem whose proof is hidden in the earlier argument.

Theorem: R is not union of countably many small sets.

Recall that we already know that R is not countable, that is it is not countable union of singleton sets. We are now saying better. It can not even be union of countably many sets each 'looking like' Cantor set.

The result we proved earlier can be deduced from this more general theorem. Here is how. Suppose, if possible

$$Q = U_1 \cap U_2 \cap U_3 \cap \cdots\cdots.$$

Then

$$R = \bigcup_{r \in Q} \{r\} \quad \cup \quad \bigcup_n U_n^c.$$

This equality is clear because all rationals are in the first collection of singleton sets; the earlier equality about Q tells that every irrationals is outside some U_n and is hence captured by some U_n^c. This is a countable union. Finally, sets in the first collection are singletons and are hence small. Each U_n^c is already closed and does not contain any rational. Hence U_n^c can not include any open interval (non-empty). So it is also small.

In other words, if you can express Q as above, then R is a countable union of small sets.

So how do we prove this theorem. As I said there is no new idea. if possible let $R = \cup C_n$ countable union of small sets. We assume that the sets C_n are closed, if necessary replace the original sets by their closures. Remember, our definition tells that A is small iff its closure is small.

We shall manufacture a sequence of intervals $[a_n, b_n]$ such that the following hold.

(i) $[a_n, b_n] \cap C_n = \emptyset$.

(ii) $[a_1, b_1] \supset [a_2, b_2] \supset [a_3, b_3] \supset \cdots$.

(iii) $0 < b_n - a_n \leq 1/2^n$.

By (ii) and (iii) and Cantor intersection theorem there is again z common to all these intervals and (i) says that this point is outside all the sets C_n. Thus the union $\cup C_n$ is not all of R. Here is how you construct these intervals.

C_1^c is open. Since C_1 does not contain any interval, you can pick an non-degenerate interval $[a_1, b_1] \subset C_1^c$. If necessary cut it down to satisfy condition of length.

C_2 being small, $(a_1, b_1) \subset C_2$ is false. So $(a_1, b_1) \cap C_2^c \neq \emptyset$ and being open, contains a non-degenerate subinterval. Take closed non-degenerate subinterval of this. Cut down to fulfil length condition to get $[a_2, b_2]$. Next time manufacture $[a_3, b_3] \subset (a_2, b_2) \cap C_3^c$ and proceed.

This completes the proof.

There are several books dealing with metric spaces, for example the book of A. N. Kolmogorov and S. V. Fomin: Introduction to real analysis. You can look at the Hewitt and Stromberg that I mentioned earlier.

Cantor intersection theorem:

In the real line we have shown that a decreasing sequence of closed non-empty intervals with diameter converging to zero have a point in common. We shall now generalize this result to metric spaces. But before doing so, let us understand that the intervals should be as stated in the theorem. A sequence of sets which are decreasing are called nested.

The intervals $\{[n, \infty) : n \geq 1\}$ are closed and nested. But they have no point in common because their diameter is not converging to zero.

The intervals $\{(0,\ 1/n); n \geq 1\}$ have diameter decreasing to zero, they are nested but have no point in common because the sets are not closed.

The intervals $\{[n,\ n+1/n]; n \geq 1\}$ are closed and have diameter decreasing to zero, but they have no point in common because they are not nested.

Theoem: Let (X, d) be a complete metric space, let

$$C_1 \supset C_2 \supset C_3 \supset \cdots$$

be nested sequence of closed setts with diameter converging to zero. Then they have exactly one point in common.

The essential thing is that there is a common point.

If there are two different points $x,\ y$ in common then the diameter of each of the sets must be at least $d(x, y) > 0$ and hence can not go to zero.

If we take one point $x_n \in C_n$, then the sequence $\{x_n\}$ is Cauchy. Indeed $\{x_n : n \geq N\} \subset C_N$ so that as soon as the diameter of C_N is smaller than ϵ, we can conclude that $d(x_m, x_n) < \epsilon$ for all $m, n \geq N$. By completeness of X this sequence must converge to a point x. Since each C_N is closed and x is also limit of the sequence $(x_n : n \geq N)$ which is contained in C_N we see that $x \in C_N$. This is true for every N.

Thus the intersection is exactly one point.

Actually the above property characterises completeness. More precisely, let (X, d) be a metric space. Suppose every nested sequence of closed sets with diameter converging to zero have a common point. Then the space is complete. This is seen as follows. Let (x_n) be a Cauchy sequence. Set $C_n = \overline{\{x_i : i \geq n\}}$. Here the overline denotes closure. This is a nested sequence of closed sets. The Cauchy property of the given sequence shows that their diameter is converging to zero. Thus there is a point common to all, say, x. We claim $x_n \to x$.

Baire's theorem:

Let (X, d) be a metric space. A closed set $C \subset X$ is small if it does not contain any non-empty open set. In other words, it has no interior points. More generally, we say that a set $A \subset X$ is small if its closure does not contain any non-empty open set. Here then is an extremely powerful theorem.

Theorem (Baire Category theorem):
A complete metric space can not be written as a countable union of small sets.

In other words, if $C_1, C_2, \cdots$ is a sequence of closed sets each of which has no interior point, then there is a point of X which is not in any one of these sets. Of course, there will be, not one but plenty of, points outside all the sets C_n, as we will see.

Proof of this theorem is very simple, but the theorem is itself very powerful. It will help you to see objects which you can not see with ordinary eyes! Let us see two applications before proving this theorem. You will learn later many many applications.

nowhere differentiable functions:

This application is due to Banach. Consider the set $X = C[0, 1]$, space of real valued continuous functions on the interval $[0, 1]$. We show that there are plenty of functions which do not have finite derivative at any point whatso-ever. In fact, the set of such functions is much larger than differentiable functions.

We know that X is a complete space. For integers $n, m > 1$ consider the following set.

$$E_{n,m} =$$

$$\left\{ x \in C[0,1] : \left(\exists t; 0 \leq t \leq 1 - \frac{1}{n}\right) \left(\forall h; 0 < h < \frac{1}{n}\right) \left|\frac{x(t+h) - x(t)}{h}\right| \leq m \right\}.$$

Observe that if a function is differentiable at some point $t < 1$, then it is in one of the above sets for some $n, m > 1$. Indeed let x be a function having a finite derivative at a point $t_0 < 1$. Get n so that $t_0 < 1 - 1/n$. Consider the fhe function $\varphi(h) = [x(t_0 + h) - x(t_0)]/h$ on $[0, 1/n]$. Of course at $h = 0$, the value is its limit which is the finite right derivative of x at t_0. This is continuous and hence bounded, say, by m. Then clearly $x \in E_{n,m}$ for this n, m.

The set $E_{n,m}$ is closed. To see this first recall that if $x_n \to x$, then x_n converges uniformly to x. If $t_n \to t$ then $x_n(t_n) \to x(t)$. This is because, given $\epsilon > 0$, we can fix N so that $d(x_N, x) < \epsilon/2$ for $n \geq N$. Since x is continuous, we fix N_1 so that $|x(t_n) - x(t)| < \epsilon/2$ for $n \geq N_1$. if now $n \geq N \vee N_1$, then

$$|x_n(t_n) - x(t)| \leq |x_n(t_n) - x(t_n)| + |x(t_n) - x(t)| < \epsilon.$$

Now to see that $E_{n,m}$ is closed, take $x_i \in E_{n,m}$ with $x_i \to x$. get t_i for x_i. If necessary take subsequence and assume that $t_i \to t^* \in [0, 1/n]$. Now take any h with $0 < h < 1/n$. Then the facts $t_i \to t^*$ and $t_i + h \to t^* + h$ combined with the observation of the para above show that x also satisfies the required inequality for difference quotients.

Finally we show that the set $E_{n,m}$ does not contain any non-empty open set. This will then show that each $E_{n,m}$ is a small set. We start with an observation.

Let P be a polynomial and $\epsilon > 0$, then we claim that there is a function x such that $d(x, P) < \epsilon$ and $x \notin E_{n,m}$.

How does this help us? If $E_{n,m}$ contains a non-empty open set, then by Weierstrass theorem this open set must contain a polynomial and hence a ball around this polynomial is contained in the open set. In other words, if you assume that interior of $E_{n,m}$ is non-empty, then there is a polynomial P and $\epsilon > 0$ such that $B(P, \epsilon) \subset E_{n,m}$. But the observation refutes precisely such a statement, it tells that there is a x in this ball which is not in $E_{n,m}$.

To prove the observation stated above, fix a bound c for the derivative of P on $[0, 1]$. Consider the following function z on $[0, 1]$; it is made up of straight line segments; it starts with $z(0) = 0$; increases with slope $s = c + 2m$ till it reaches $\epsilon/2$; then decreases with slope $-s$ till it reaches $-\epsilon/2$; then increases with slope s till it reaches $\epsilon/2$ and then decreases etc; all this continues till you reach $t = 1$ and then stops. Convince yourself that you do reach $t = 1$ after finite number of these ups and downs. You then see that this defines a continuous function on $[0, 1]$.

Let $x = P + z$. Then clearly $d(P, x) = \sup |z(t)| < \epsilon$. Now take any $t < 1$. Take any $h > 0$ so that $t + h$ is also before the next corner of z. Thus you have plenty of $h > 0$ at your disposal. Then

$$\frac{x(t+h) - x(t)}{h} = \frac{P(t+h) - P(t)}{h} + \frac{z(t+h) - z(t)}{h}$$

By mean value theorem the first term on right side is between $-c$ and $+c$. By construction the second term is either $c + 2m$ or $-c - 2m$. Keep in mind that $t, t + h$ are in the same line segment of z. As a result the difference quotient for x is at least $2m > m$.

This shows that each $E_{n,m}$ is small. Similarly, the following sets are also small.

$$F_{n,m} =$$

$$\left\{ x \in C[0,1] : \left(\exists t; \frac{1}{n} \leq t \leq 1\right) \left(\forall h; -\frac{1}{n} < h < 0\right) \left|\frac{x(t+h) - x(t)}{h}\right| \leq m \right\}.$$

Thus we have countably many small sets $E_{n,m}$ and $F_{n,m}$ for $n, m \geq 1$. If a function has a derivative at any point of $[0, 1]$ then it must be in one of these sets. actually we can be more precise as follows. If a function has a finite right derivative at any point of $[0, 1)$ then it must be in one of the $E_{n,m}$. If a function has a finite left derivative at any point of $(0, 1]$ then it must be in one of the $F_{n,m}$.

Baire's theorem tells that there are functions outside all these sets. If we take such a function then it can not have a finite right derivative or finite left derivative at any point what-so-ever in $[0, 1]$. In particular it is not differentiable at any point.

Just to make you understand the right and left derivatives, let us consider the function $x(t) = |t - 1/2|$. Then this is not differentiable at the point $t_0 = 1/2$. Convince yourself of this. However this has left derivative equal to -1; and right derivative equal to $+1$ at the point $t_0 = 1/2$. Thus the functions whose existence we asserted above can not even be like this. Think about it.

Before going to next application of Baire's theorem, let us remember a little history about the hero of this application: Banach.

There was a time when hell descended on earth; in the form of second world war. Poland suffered very heavily. As far as Maths is concerned here is a brief view.

Some could emigrate early on: Alfred Tarski (logician), Antony Zygmund (analyst), Jerzy Neyman (statistician), Samuel Eilenberg (topologist), Stanislaw Ulam (set theory, computation etc) and several others.

Some stayed on and survived the war, either by going underground or showing that they have pure blood, whatever it may mean: W. Sieprinski, K. Kuratowski, H. Steinhaus and others.

Some committed suicide: F. Hausdorff.

Some were put to death in camps: S. Saks, J. Marcinkiewicz, J. Schauder, A. Rajchman, A. Lindenbaum and many many many others.

Some were saved by Director of lab (in Lwow, a city in Poland) that makes Typhus vaccine. Since Germans needed it, they allowed him to choose volunteers. This vaccine needs lice. Growing lice is done by carefully packing them and attaching to the calf or thigh of human so that they suck the blood. These volunteers are called lice-feeders. Our hero Banach was one such. He survived the war but died soon after due to failed health (and lung cancer).

R as union of closed sets:

Our second application of Baire's theorem is to show the following which answers a question that we raised earlier.

Theorem: R can not be expressed as union of of infinitely many nonempty disjoint closed sets.

We already knew that there is no subset of R which is open and closed except $\emptyset$ and R itself. Thus we can not express R as union of two disjoint nonempty closed sets. This in turn implies that we can not express $R = \cup_1^k C_i$ where C_i are nonempty disjoint closed sets and $k > 1$. If this could be done then, you can simply take $A = C_1$ and $B = \cup_2^k C_i$ to see R is union of two nonempty disjoint closed sets.

Thus R can not be expressed as a finite union of more than one nonempty disjoint closed sets. The theorem says you can not express R even as countably infinite union of nonempty closed sets.

Let, if possible

$$R = C_1 \cup C_2 \cup C_3 \cup \cdots\cdots.$$

Let the interior of C_i be denoted by J_i. Denote

$$U = J_1 \cup J_2 \cup J_3 \cup \cdots\cdots.$$

Since none of the C_i can be open we see each J_i is a proper subset of C_i, possibly empty. This U is an open set and its complement, denoted by H is therefore a non-empty closed set. Thus H is a complete metric space in its own right, metric is same $d(x, y) = |x - y|$ for $x, y \in U$.

$$H = (C_1 \cap H) \cup (C_2 \cap H) \cup (C_3 \cap H) \cup \cdots\cdots$$

$$= (C_1 - J_1) \cup (C_2 - J_2) \cup (C_3 - J_3) \cup \cdots\cdots.$$

The plan of the proof now is the following. We shall now forget R for a minute and concentrate on the complete metric space H. We show that the sets on the right side above are small in H. This contradicts Baire's theorem for the complete metric space H.

Let us start with an observation. If an interval (a, b) contains a point, say x, of $C_1 \cap H$, then it contains points from other $C_i \cap H$ $(i \neq 1)$ as well. This is easy. If the entire interval (a, b) is contained in C_1 then it would have been removed as part of J_1, So this interval must have points from, say, C_5. Let us say $y \in C_5 \cap (a, b)$. Either $y < x$ or $y > x$ because C_i are disjoint. Suppose $y < x$; similar argument applies in the other case. Consider all points in $[y, x]$ which are in C_5 and take its sup, name it z. This is sensible because the set we are considering includes y and is hence non-empty; moreover it is bounded by x. Also C_5 being closed we conclude that $z \in C_5$. In particular $z < x$ and points in between these are not in C_5. in other words there is no interval around z contained in C_5 which means $z \in C_5 - J_5$ and thus $z \in C_5 \cap H$ and $z \in (a, b)$ as claimed.

To complete executing our plan, first notice that each $C_i \cap H$ is a non-empty closed subset of H. That it is non-empty is already noted earlier. It is closed in H because whenever a sequence of points from here converge to a limit then the convergence is 'usual convergence' and hence the limit is in both the closed sets C_i as well as H.

finally suppose $C_i \cap H$ contains a non-empty set open in H. That is, there is a point $x \in C_i \cap H$ and $\epsilon > 0$ such that $\{y \in H : d(x, y) < \epsilon\} \subset C_i \cap H$. In other words $(x - \epsilon, x + \epsilon) \cap H \subset C_i \cap H$. But this is not possible as observed above. This completes the proof that each $C_i \cap H$ is a small set in the complete metric space H.

And completes proof of the theorem.

Just to impress upon you the phrase 'small in H', let us consider R and Z the set of integers. Each $\{n\}$ is small in R. However none of these singleton sets are small in Z. In fact each of these are both closed and open in Z.

Thus smallness depends on the background. In the background of R, each $\{n\}$ is small. In the background of Z, each $\{n\}$ is not small. You should understand this point.

Proof of Baire:

This is exactly same as the one for real line, there is no new idea. The execution is made possible by Cantor intersection theorem for complete metric spaces.

Before you get the wrong impression that Baire just imitated the real line proof, let me say the following. Even for the real line it is due to Baire. It is not that the theorem existed for R and he extended it to metric spaces. We discussed real line case first only to understand the argument in a familiar territory, so that the general case would pose no problem.

So let (X, d) be a complete metric space and let $C_1, C_2, C_3, \cdots$ be small closed sets. We exhibit a point of X which is not in any of the C_i.

In what follows a ball means ball of positive radius. Every open ball contains a closed ball with same center (say, any strictly smaller radius) and every closed ball contains an open ball with same centre (same radius). In what follows we ask you to choose an open ball inside a closed ball or we ask you to take a closed ball inside an open ball. Then you should select with the same center as mentioned above. This is our agreement.

Take any open ball B_1 of your choice. We promise to get open balls

$(B_i : i \geq 1)$, closed balls $(F_i : i \geq 1)$ such that the following holds.

$$B_1 \supset F_1 \supset B_2 \supset F_2 \supset B_3 \supset F_3 \supset \cdots \supset B_{n-1} \supset F_{n-1} \supset \cdots \quad (*)$$

$$F_i \subset C_i^c; \quad i = 1, 2, 3, 4, \cdots \qquad (**)$$

$$\text{diameter}(F_i) \leq 1/2^i. \qquad (***)$$

Let us see what happens then. Condition $(*)$ says that the sets F_i are nested closed sets; condition $(***)$ with Cantor intersection theorem then gives a point x common to all F_n; condition $(**)$ says that this point is not in any of the sets C_i. as promised. Actually condition $(*)$ tells that this point is in the open set B_i you gave.

Here is how we construct the sets. Since C_1 is small surely $B_1 \cap C_1^c \neq \emptyset$ and is open. So take a closed ball

$$F_1 \subset B_1 \cap C_1^c.$$

If you have possibility of choosing large ball, restrain, choose ball of radius at most 1/2.

Take open ball $B_2 \subset F_1$ as per our agreement above. Since C_2 is small $B_2 \cap C_2^c \neq \emptyset$. So select closed ball

$$F_2 \subset B_2 \cap C_2^c.$$

Again make sure diameter of F_2 is at most $1/2^2$. Take open ball $B_3 \subset F_2$ as per our agreement. Then select a closed ball

$$F_3 \subset B_3 \cap C_3^c.$$

Make sure its diameter is at most $1/2^3$.

Here is the inductive step. Suppose we got the balls (B_i) and (F_i) for $i = 1, 2, \cdots, n-1$ as satisfying the three conditions. here is how we construct B_n and F_n. Of course B_n is the open ball contained in F_n as per our agreement earlier. Since C_n is small take closed ball

$$F_n \subset B_n \cap C_n^c.$$

If you have a possibility of choosing large F_n, cut it down to have diameter smaller than $1/2^n$.

This complete the construction and proof of the theorem.

I would like to stress once again two points which I made earlier. First is this. After getting the first three sets you could say, 'continue like this'. But when you write a proof, you must make sure that such a continuation 'for ever' is indeed possible. So it is important for you to show this.

Secondly, you see, what we want is just that diameter$(F_n) \to 0$ to get a point common to all of them; it does not matter how it converges to zero, it does not need to be smaller than $1/2^n$. Thus in stating my conditions suppose I carelessly stated $(***)$ as: diameter $(F_n) \to 0$. You will *not be able to construct* sets by induction simply because it does not make sense to say that we have constructed sets up to n satisfying the three conditions.

Thus whenever you need to make an *unending construction* you must be able to write the conditions in a way that they make sense inductively; an you should be able to explain that having done the construction up to a stage it can be continued to the next stage.

Please pay attention and do think about it.

completion of a metric space:

Having seen how important it is to have a complete space, the natural question is the following. If the metric space is not complete, is there anything we can do complete it?

Why is the space not complete. There are Cauchy sequences which are not converging. So either we should make sure such sequences are not Cauchy or provide a point (of convergence) for each such sequence. The first alternative works if we can provide suitable metric without changing the notion of convergence. This is possible if the space is already an open subset of a complete metric space. This is possible even if the space is a set which is countable intersection of open sets in a complete space. It stops there and does not work for any metric space.

Besides, in the procedure described above, there are two problems. firstly, I said if your space is an open subset of a complete space you can change the metric. But how do you recognize that there is indeed a bigger space which is indeed complete and our space is indeed an open subset of it? Second point is that one may not like to change the distance. After all, if a particular metric is natural then changing it, just to achieve some other desired property, makes the new metric artificial and devoid of meaning. such a thing should be avoided.

So we look for the second alternative. This is what Cantor did. We mentioned this issue when we described Cantor's construction of real numbers. Exactly the same procedure works, not only for the set of rational numbers, but for any space.

This is what we do now.

So let (X, d) be a metric space. Consider the space X^1 of all cauchy sequences in X. Thus each element of X^1 is a Cauchy sequence (x_n).

Define $(x_n) \sim (y_n)$ if $d(x_n, y_n) \to 0$. This is an equivalence relation on X^1. Denote the space of equivalence classes by X^*. Thus elements of X^* are bags. Each bag contains Cauchy sequences which are equivalent. If there is one Cauchy sequence (x_n) in a bag then every Cauchy sequence (y_n) equivalent to (x_n) is also in that bag — and nothing else is there. Elments of X^* are denoted $[x]$.

Let us observe that if (x_n) and (y_n) are Cauchy sequences in X then the limit $\lim d(x_n, y_n)$ exists. Indeed, to show this, it is enough to show that the sequence of real numbers $\{d(x_n, y_n)\}$ is a Cauchy sequence of real numbers. But this is easy

$$|d(x_n, y_n) - d(x_m, y_m)| \leq d(x_n, x_m) + d(y_n, y_m).$$

and hence can be made as small as we please for all large values of m and n. You only need to see that both terms on the right side can be made small.

Further if $(a_n) \sim (x_n)$ and $(b_n) \sim (y_n)$ then $\lim d(a_n, b_n) = \lim d(x_n, y_n)$. This is again because

$$|d(a_n, b_n) - d(x_n, y_n)| \leq d(a_n, x_n) + d(b_n, y_n) \to 0.$$

Let us define for $[x]$ and $[y]$ in X^*

$$d^*([x], [y]) = \lim d(x_n, y_n).$$

The above analysis shows that this limit exists and is well-defined on X^*.

We show that d^* is a metric on X^*. Clearly $d^*([x], [x]) = 0$. Also $d^*([x], [y]) = 0$ implies, by definition of the equivalence relation, that $[x] = [y]$. Symmetry, $d^*([x], [y]) = d^*([y], [x])$ is also clear. the triangle inequality of d leads to the triangle inequality for d^*.

We shall now identify X as a subset of X^*. For $p \in X$, let $\varphi(p)$ be the constant sequence $x_n = p$ for all n, more precisely, $\varphi(p)$ is the bag containing this sequence. Then this map is one-one. This preserves distance too. That is $d(p,q) = d^*([p],[q])$.

We shall show that X is a dense subset of X^*; (X^*, d^*) is a complete metric space. Thus X is *enlarged* to a set X^* and the metric is also *extended* to d^* to make the space complete.

By showing that X is dense, we are saying that for every point of $z \in X^*$ there is a Cauchy sequence $(x_n) \subset X$ which converges to it. In other words, in this process of completion we have not added unnecessary points, every new point z we aded, being limit of a Cauchy sequence (x_n) in X.

This leads to the feeling that the completion is unique. Yes, this is so. We have not yet defined what is completion. We shall do and then prove that completion is indeed unique.

There are some books that I have consulted from time to time for nice problems apart from the books I already mentioned. I am giving below. But you should feel free (and also make it a habit) to consult any book from the library. It is not enough to be able to understand what I say; it is very important to be able to understand others too.

You should develop the habit of reading material and internalizing it. That is, *do not* classify it as easy or difficult (you are not here to judge, though you can do it); do not reproduce it word-to-word (this is not a memorizing contest, though you can do it); *but* understand and think about it till you are able to explain to others in your own words.

An introduction to complex analysis and geometry
John P D'Angelo.

p-adic analysis compared with real
Katok Svetlana.

A primer of real functions
Ralph P Boas.

completion continued:

We have a metric space (X, d). We considered the space X^1 of Cauchy sequences (x_n). Defined an equivalence relation $(x_n) \sim (y_n)$ if they appear to be converging to the same point; more precisely, if $d(x_n, y_n) \to 0$. X^* is the space of equivalence classes or bags as we called them. Define $d^*([x], [y]) = \lim d(x_n, y_n)$. We showed that this limit exists and does not depend on the sequences you have taken from the two bags.

Defined the map $\varphi : X \to X^*$ as follows. For $p \in X$, $\varphi(p)$ is the bag containing the constant sequence all of whose terms equal p. We showed that this is distance preserving map.

Recall that, for a metric space (Z, ρ), a subset $D \subset Z$ is dense in Z if every non-empty open set contains a point of D. equivalently, given $z \in Z$ and $\epsilon > 0$, there is $p \in D$ such that $\rho(z, p) < \epsilon$..

We now show that range $\varphi(X)$ is dense in X^*. To do this, take $[x] \in X^*$ say the bag containing the Cauchy sequence (x_n). Let $\epsilon > 0$. We shall show $p \in X$ such that $d^*(\varphi(p), [x]) < \epsilon$. Choose N such that $n, m \geq N$ implies $d(x_n, x_m) \leq \epsilon/2$. Take $p = x_N$. We show that this does. Since $\varphi(p)$ is the constant sequence $\{p, p, p \cdots\}$ we see, by calculation,

$$d^*(\varphi(p), [x]) = \lim d(x_N, x_n) \leq \epsilon/2 < \epsilon.$$

We now show that X^* is a complete metric space. But before we do this let us make an observation which will avoid some notational confusion later.

Let (S, d) be a metric space and $T \subset S$ be a dense subset.

Assume the following: (x_n) is a Cauchy sequence; each x_n belongs to T then there is a point of $x \in S$ such that $x_n \to x$. Then the metric space S is complete.

Equivalently, every Cauchy sequence in S converges.

In other words, if you know Cauchy sequences with terms coming from T converge (in S, we are not saying that they converge to points in T) then every Cauchy sequence in S converges.

Proof is simple and as follows. let (x_n) be a Cauchy sequence in S. for each n, using denseness of T pick $y_n \in T$ so that $d(x_n, y_n) < 1/2^n$. Now we see that (y_n) are all points in T. The plan is to show that (y_n) is a Cauchy sequence and get its limit guaranteed by hypothesis, and show that original sequence (x_n) also converges to this limit. Towards showing (y_n) is Cauchy, fix $\epsilon > 0$. Choose N so that $m, n \geq N$ implies $d(x_n, x_m) < \epsilon/2$ and also $\sum\limits_{m \geq N} 2^{-m} < \epsilon/2$. if now, $m, n \geq N$

$$d(y_m, y_n) \leq d(y_m, x_m) + d(x_m, x_n) + d(x_n, y_n) \leq \epsilon/2 + \epsilon/2.$$

So let $y_n \to y$. Towards showing that $x_n \to y$, you only need to note that $d(x_n, y_n) \leq 1/2^n \to 0$. Or, explicitly, given $\epsilon > 0$, choose N so that $|y_n - y| < \epsilon/2$ for $n \geq N$ and also $1/2^N < \epsilon/2$. Then for $n \geq N$,

$$d(x_n, y) \leq d(x_n, y_n) + d(y_n, y) \leq \epsilon/2 + \epsilon/2.$$

This completes the observation. Let us now return to showing that X^* is complete. From the observation, it suffices to show that Cauchy sequences whose points come from the dense set $\varphi(X)$ converge. Accordingly, take a Cauchy sequence $\varphi(p_n)$ where $p_n \in X$ for each n. Since φ is distance preserving map, we conclude that $x = (p_n)$ is a Cauchy sequence in X and is an element of X^1, the space of Cauchy sequences. We show $\varphi(p_n) \to [x]$. Take $\epsilon > 0$. Choose N so that $d(p_n, p_m) < \epsilon/2$ for $n, m \geq N$. We show now $d^*(\varphi(p_i), [x]) < \epsilon$. for $i \geq N$. Keep in mind that $\varphi(p_i)$ is the constant sequence

$$p_i, \; p_i, \; p_i, \; p_i, \cdots$$

and x is the sequence

$$p_1, \; p_2, \; p_3, \; p_4, \cdots.$$

Use definition of d^* to see

$$d^*(\varphi(p_i), [x]) = \lim_n d(p_i, p_n) \leq \epsilon/2.$$

The last inequality is from the fact that $i \geq N$ and as soon as $n \geq N$ we know $d(p_i, p_n) \leq \epsilon/2$.

definition of Completion:

Let (X, d) be a metric space. By completion of (X, d) we mean a metric space (Z, ρ) and a map $\varphi : X \to Z$ such that the following hold.
(i) Z is a complete metric space.
(ii) φ is distance preserving map.

(iii) Range $\varphi(X)$ is dense in Z.

Condition (ii) tells you that $\rho(\varphi(x), \varphi(y)) = d(x, y)$. In particular if $x \neq y$ then $\varphi(x) \neq \varphi(y)$. Thus φ is a bijection from X to $\varphi(X)$. Since the distance is also preserved, when you see $\varphi(X)$ it looks exactly like X. In other words you see a replica of X in this new space.

To put it differently, if you rename the point $\varphi(x)$ as x, then you see X in Z. Thus many times we regard $X \subset Z$. Just like, while constructing real number system starting from rationals, we considered either cuts or Cauchy sequences; but no matter what, we regarded Q as a subset of the R we constructed.

Such a view leads to the nice feeling that the extra points, that is points of $Z - X$, are the new ones needed to show as limits of Cauchy sequences. Thus the new space is not as abstract as it appears, but original set of pints with new things thrown in where ever necessary.

Condition (iii) tells you that you have not added unnecessary points. More precisely, if $z \in Z - \varphi(X)$, then denseness of $\varphi(X)$ tells you that there is a sequence $(x_n) \subset X$ such that $\varphi(x_n) \to z$. But then $\varphi(x_n)$ is a Cauchy sequence (recall any convergent sequence is a Cauchy sequence, in any metric space). But φ being distance preserving, we conclude (x_n) itself is a Cauchy sequence in X. It does not converge in X [if it did, say to x then $\varphi(x_n)$ converges to $\varphi(x)$ a point of $\varphi(X)$ contradicting that it converges to a point of $Z - \varphi(X)$.] Thus the point z is essential to show as limit of the Cauchy sequence (x_n).

Of course condition (i) tells you that your new space is complete. Thus we 'embedded' X in a complete space without bringing in un-necessary points in the process.

Thus what we have shown is that every metric space has a completion. (X^*, d^*) satisfies all the three conditions.

We shall now show that a completion is unique. More precisely, if there are two completions (Z_1, ρ_1, φ_1) and (Z_2, ρ_2, φ_2) then there is distance preserving bijection between them that keeps X fixed. This means there is $f : Z_1 \to Z_2$ which is distance preserving and $f(\varphi_1(x)) = \varphi_2(x)$ for all $x \in X$. Remember $x \in X$ looks like $\varphi_1(x)$ in Z_1 whereas it looks like $\varphi_2(x)$ in Z_2.

Sometimes statement of a claim itself includes how to start its proof.

Thus let there be two completions with notation as above. We define f as follows. For $x \in X$ we define $f(\varphi_1(x)) = \varphi_2(x)$. Now take any $z \in Z_1$. Since $\varphi_1(X)$ is dense in Z_1 take a sequence $\varphi_1(x_n) \to z$ in Z_1. In particular, $\{\varphi_1(x_n)\}$ is a Cauchy sequence in Z_1. Since φ_1 is distance preserving we conclude that (x_n) is Cauchy in X. But now φ_2 is distance preserving tells us that $\{\varphi_2(x_n)\}$ is Cauchy in Z_2. since Z_2 is complete the limit $\lim \varphi_2(x_n)$ exists in Z_2 and this limit is defined as $f(z)$.

This is a good definition because if some one takes a different sequence $\varphi_1(y_n) \to z$ then we see

$$\rho_2(\varphi_2(x_n), \varphi_2(y_n)) = d(x_n, y_n) = \rho_1(\varphi_1(x_n), \varphi_1(y_n)) \to 0$$

so that $\varphi_2(y_n)$ also converges o the same limit as $\varphi_2(x_n)$.

This completes the definition of f on Z_1 to Z_2. It is distance preserving because for two points $z, w \in Z_1$ take $\varphi_1(x_n) \to z$ and $\varphi_1(y_n) \to w$ and see

$$\begin{aligned} \rho_2(f(z), f(w)) &= \lim \rho_2(\varphi_2(x_n), \varphi_2(y_n)) = \lim d(x_n, y_n) \\ &= \lim \rho_1(\varphi_1(x_n), \varphi_1(y_n)) = \rho(z, w). \end{aligned}$$

It follows that f is one-one too. Indeed if $z \neq w$, then

$$\rho_2(f(z), f(w)) = \rho_1(z, w) \neq 0$$

showing that $f(z) \neq f(w)$.

Shall show that f is onto Z_2. Indeed, if $w \in Z_2$, then using $\varphi_2(X)$ is dense, take $\varphi_2(x_n) \to w$. Repeating the earlier argument conclude $\{\varphi_2(x_n)\}$ and hence $\{x_n\}$ and hence $\{\varphi_1(x_n)\}$ are Cauchy (in their respective spaces) Using completeness of Z_1 get limit z of this last sequence and argue $f(z) = w$.

This completes the proof of the fact that every metric space has a completion and it is unique in the sense (loosely speaking) between any two completions there is a distance preserving bijection that is identity on X, which is inside both. Of douse, more precise statement is as mentioned earlier.

This process is due to Cantor. What we have done earlier to construct real numbers starting from rationals is precisely this completion; no more and no less.

This analysis shows some general facts. here is one.

Suppose (Z_1, ρ_1) and (Z_2, ρ_2) are two complete metric spaces. suppose D is a dense subset of Z_1 and f is a distance preserving map defined on D into Z. Then the map can be extended to a distance preserving map on Z_1 into Z_2. Proof is already contained in what we have done above. Take any $z \in Z_1$; use D is dense, get $x_n \in D$ such that $x_n \to z$, conclude (x_n) is Cauchy, conclude $f(x_n)$ is Cauchy, use Z_2 is complete, get its limit and declare that as $f(z)$. One shows that this is well defined; does not depend on the sequence (x_n) you have taken from D.

Also this is distance preserving just as above. Such an extension is unique too, because any distance preserving map has the property that when $x_n \to x$, then $f(x_n) \to f(x)$. This shows that for z, its value $f(z)$ must equal to what we defined.

of course this may not be onto. One can show that it is onto iff $f(D)$ is dense in Z_2.

Here is another fact that comes out of the analysis above. Let X^1 be a set and $d^1 : X \times X \to [0, \infty)$ satisfying the following:
$d^1(x, x) = 0$.
$d^1(x, y) = d^1(y, x)$.
$d^1(x, z) \leq d^1(x, y) + d^1(y, z)$.

Thus d^1 is nearly a metric. It falls short of being metric only because it may not satisfy: $d^1(x, y) = 0$ implies $x = y$. If this condition is also satisfied then this is indeed a metric.

Let us start with (X^1, d^1) as above. then define $x \sim y$ if $d^1(x, y) = 0$. This is an equivalence relation. Indeed, $x \sim x$ because $d^1(x, x) = 0$. If $x \sim y$ then $y \sim x$ by symmetry of d^1. Triangle inequality shows that $x \sim y$ and $y \sim z$ implies $x \sim z$.

Let us consider the space of equivalence classes, denote, X. Define $d([x], [y]) = d^1(x, y)$. This is well defined. Indeed, if $u \in [x]$ and $v \in [y]$, then

$$d^1(u, v) \leq d^1(u, x) + d^1(x, y) + d^1(y, v) = 0 + d^1(x, y) + 0 = d^1(x, y).$$

and

$$d^1(x,y) \leq d^1(x,u) + d^1(u,v) + d^1(v,y) = 0 + d^1(u,v) + 0 = d^1(u,v).$$

Thus $d^1(x,y) = d^1(u,v)$.

It is not difficult to verify that (X,d) is a metric space. Thus d satisfies all the three rules above and the missing rule too.

If you carefully see the completion process this is precisely what we did. The space X^1 is the space of Cauchy sequences and
$d^1((x_n),(y_n)) = \lim d(x_n,y_n)$.
The above equivalence relation is precisely what we used to make bags of Cauchy sequences and the new metric was d^*.

In other words the process of completion due to Cantor not only outlined how to complete a metric space, producing a construction of real numbers; it has also thrown out certain general techniques like the above two. [Of course, historically, matters are different. Cantor did for real line construction. But what he did was so powerful that it applied for metric spaces when they were discovered].

Some times, completion has a concrete representation. For example consider $(0,1)$ with usual $d(x,y) = |x-y|$. The space is not complete. To complete it you should consider space of cauchy sequences. But in this case completion is just $[0,1]$. This is obvious pif you consider the map φ to be the map $\varphi(x) = x$ on $(0,1)$. Here $X = (0,1)$, $Z = [0,1]$. We know Z is complete, φ preserves distance and its range $\varphi(X)$ is dense in Z.

But then what happens if you repeated the construction? Well you see apart from the bags containing constant sequences $\{p,p,p\cdots\}$ for $p \in (0,1)$; you will get only two new bags. They are: bag containing the sequence $\{1/2, 1/3, 1/4, 1/5, \cdots\}$ and bag containing $\{1/2, 2/3, 3/4, 4/5, \cdots\}$. If you name these two bags as zero and one you have $[0,1]$.

category, pseudo metric, isometry:

I have resisted the temptation of mentioning technical words. This is because many times students use such a technical word but unfortunately can not explain what it means. You should know that technical word is only an agreement to use a compact word instead of long expression. The most important thing is to know and understand what the word stands for.

The d^1 we described above which nearly satisfies the axioms of a metric is called a pseudo-metric.

Thus **pseudo metric** on a set X is a function d defined on $X \times X$ which satisfies axioms of metric except possibly: $d(x, y) = 0$ implies $x = y$. Thus a metric is a pseudo-metric, but a pseudo metric need not be metric. But what we described above produces a metric space from a pseudo-metric space.

For example, you can consider R^3 and define

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + (x_3 - y_3)^2}; \quad x = (x_1, x_2, x_3), y = (y_1, y_2, y_3).$$

This is a pseudo metric but not a metric.

Distance preserving maps are called isometries. Thus an **isometry** from a metric space (Z_1, d_1) to a metric space (Z_2, d_2) is a map T satisfying $d_1(x, y) = d_2(Tx, Ty)$.

As mentioned at the beginning of our excursion into metric spaces, the plan is to execute convergence, continuity, and imitate some of our calculus concepts in a more general setting. We discussed sequences and convergence. A function from a metric space X to a metric space Y is said to be **continuous** if the following holds: whenever $x_n \to x$ then $f(x_n) \to f(x)$.

Several routine results from calculus can be imitated. For example a function is continuous iff inverse image of open sets are open. Also, the (ϵ, δ) definition holds. More precisely, define f, as above from X to Y is **continuous at a point** a if the following holds: $x_n \to a$ implies $f(x_n) \to f(a)$.

Then one can show f is continuous at a iff given $\epsilon > 0$, there is a $\delta > 0$ such that the following holds: $d_Y(f(x), f(a)) < \epsilon$ whenever $d_X(x, a) < \delta$. Also one can show that f is continuous iff it is continuous at every point. Proofs are exactly same as in the usual case discussed in Calculus.

Returning to isometries, you can immediately see that an isometry is a continuous map. It is more than continuous. It satisfies a stronger property, namely, preserves distances.

What we called small sets are called nowhere dense sets. Thus **nowhere dense set** is a set A such that closure of A has no interior point. That is, the only open set contained in $\overline{A}$ is the empty set.

Here is the reason for this word. if you take Q the set of rationals then it is dense in R. It was referred to as everywhere dense. Suppose I take D to be the set of all integers union set of rationals in $(0, 1)$. Then it is clear that D is not dense in R, for example, the interval $(14, 15)$ does not have any point of D. But you can see that this set D is dense in a part of R, namely, in the open set $(0, 1)$ (or $[0, 1]$ does not matter).

A small set is not dense in any open set $U \neq \emptyset$. That is why it was called nowhere dense. A set is **first category** if it is union of countably many nowhere dense sets. A set is of **second category** if it is not of first category.

Thus we can state **Baire's theorem** as: Avery complete metric space is of second category.

Try to understand these technical words, then you can follow literature. Nobody uses the word small sets, distance preserving maps etc.

Banach's contraction mapping principle:

We need to discuss compactness and continuous maps. But let us discuss one more important theorem in Complete metric spaces. This is due to Banach.

Let X be a metric space. A map $T : X \to X$ is called a contraction if there is a number c such that $0 \leq c < 1$ and for all points $d(Tx, Ty) \leq cd(x, y)$. Thus application T reduces distances in a strong sense. It is not simply distances are reduced, they are reduced by an assured proportion c.

Of course, contraction is a continuous map. Indeed if $x_n \to x$ then

$$d(Tx_n, Tx) \leq cd(x_n, x) \to 0.$$

Theorem (Banach fixed point theorem, Banach contraction mapping principle).

If T is a contraction of a complete metric space X to itself then there is unique point x^* such $Tx^* = x^*$.

Such a point $Ta = a$ is called a fixed point, because T fixes it. That fixed point is unique is not surprising. If a and b are fixed points then

$$d(a, b) = d(Ta, Tb) \leq cd(a, b); \quad i.e. \quad d(a, b) = 0.$$

Existence o fixed point is also easy, we show that whatever point x you start with, the sequence $x, Tx, T^2x, T^3x, \cdots$ converges to x^*.

This theorem has several applications and several generalizations. We are not concerned with any generalization. There are several ways contraction maps arise and are useful.

suppose that f is a C^1 map of $[a, b]$ to itself. Since derivative of f is continuous and the interval is closed bounded, there is a number c such that $|f'| \leq c$. Suppose that (by chance) $c < 1$. Then the map f is a contraction map. Indeed by mean value theorem, if you take $x \neq y$ then

$$\frac{|f(x) - f(y)|}{|x - y|} = |f'(\theta)| \leq c.$$

In other words, $|f(x) - f(y)| \leq c|x - y|$.

Or equivalently, we have $d(f(x), f(y)) \leq cd(x, y)$.

Banach Contraction mapping principle:

Let (X, d) be a complete metric space. suppose $T : X \to X$ is a contraction map., that is, there is a number c; $0 \leq c < 1$ such that $d(Tx, Ty) \leq cd(x, y)$ for all points x, y. Thus the distance between images is 'smaller' than the distance between original points. In other words T decreases distance. The main point is that distance reduction is by a fixed proportion; not simply that it is reduced.

The theorem says that there is a unique fixed point, that is, a point x^* such that $Tx^* = x^*$. We saw that there can not be two fixed points. To show that there is one, let us start with any point x. Put

$$x_0 = x; \quad x_1 = Tx_0; \quad x_n = Tx_{n-1} \quad n \geq 1.$$

We shall show that (x_n) is a Cauchy sequence.

Then by completeness there is a point x^* such that $x_n \to x^*$. Observe that if $a_n \to a$ then $Ta_n \to Ta$ because

$$d(Ta_n, Ta) \leq cd(a_n, a) \to 0.$$

Thus $Tx_n \to Tx^*$. But $Tx_n = x_{n+1} \to x^*$. Thus $Tx^* = x^*$.

We now show that (x_n) is Cauchy. For $n \geq m$

$$d(x_n, x_m) \ \leq \ cd(x_{n-1}, x_{m-1}) \ \leq \ c^2 d(x_{n-2}, x_{m-2})$$

$$\leq \ \cdots \ \leq \ c^m d(x_0, x_{n-m})$$

$$\leq \ c^m \{d(x_0, x_1) + d(x_1, x_2) + \cdots + d(x_{n-m-1}, x_{n-m})$$

$$\leq \ c^m d(x_0, x_1)\{1 + c + c^2 + c^3 + \cdots + c^{n-m}\}$$

$$\leq c^m d(x_0, x_1) \frac{1}{1 - c}.$$

Let $\epsilon > 0$ be given. Since $c < 1$ choose N so that $c^N d(x_0, x_1)/(1-c) < \epsilon$. Let now $n, m \geq N$. without loss of generality, let us say $n \geq m$. The estimate above shows that $d(x_n, x_m) < \epsilon$. This completes proof.

This is a very useful tool and a powerful technique. We shall discuss some examples.

First let us note that every continuous map of an interval $[a, b]$ to itself has a fixed point; whether it is a contraction or not. This is because if f is the map, then $f(x) - x$ is a continuous function. Since values of f are in $[a, b]$, we see $f(a) - a \geq a - a = 0$ and $f(b) - b \leq b - b = 0$. The intermediate value theorem completes the proof.

However the above argument does not tell that fixed point is unique. In fact it may not be, as for example, the identity map $f(x) = x$ shows. The argument does not tell you how to get a fixed point. If f is a C^1 function (continuously differentiable) then the derivative is bounded. If $|f'| \leq c < 1$ then f is a contraction and so if you start from any point in this interval and keep applying f successively, you will be heading to the fixed point.

Not only that, you get an idea of how close you are to the fixed point. Indeed the estimate obtained above tells you (letting $n \to \infty$) that

$$d(x_m, x^*) \leq c^m \, d(x_0, x_1)/(1 - c).$$

contraction on R:

Consider the function

$$f(x) = \frac{1}{2} \sin(\cos x) + 239$$

Then you see $|f'| \leq 1/2$ and hence it is a contraction on R. Thus there is a unique fixed point. Of course, this is nothing exciting because you see that the values of the function lie in the interval $239 \pm 1/2$. So you can regard f as a map of this interval to itself. But as explained earlier, you have a method to obtain the fixed point.

contractions on R^n:

When is a linear map of R^n a contraction? More generally, let us consider an affine map $Tx = Ax + b$ where $A = (a_{ij})$ is an $n \times n$ matrix and b is a vector in R^n. Obviously whether a map is a contraction or not depends on the metric and here we have several metrics on R^n and under each of them it is complete — all the metrics are equivalent. But let us consider, for illustration, only two of those.

Let us consider the Euclidean distance

$$d(x,y) = \sqrt{\sum (x_i - y_i)^2}.$$

Thus

$$d(Tx,Ty) = \sqrt{\sum_i \left\{ \sum_j a_{ij}(x_j - y_j) \right\}^2} \leq \sqrt{\sum_i \left\{ \sum_j a_{ij}^2 \quad \sum_j (x_j - y_j)^2 \right\}}.$$

$$= \sqrt{\sum_{i,j} a_{ij}^2}\, d(x,y).$$

Thus if $\sum\limits_{i,j} a_{ij}^2 < 1$ then the map T is a contraction (in the Euclidean metric) and hence there is a fixed point.

Let us take the distance

$$d(x,y) = \sum_i |x_i - y_i|.$$

Then denoting the column sum $\sum\limits_i |a_{ij}| = c_j$ and $c = \max c_j$ we have

$$d(Tx,Ty) = \sum_i \left| \sum_j a_{ij}(x_j - y_j) \right| \leq \sum_i \sum_j |a_{ij}|\ |x_j - y_j|$$

$$= \sum_j c_j |x_j - y_j| \leq c d(x,y).$$

Thus if $\max\limits_j \sum\limits_i |a_{ij} < 1$ then T is a contraction in the d_1 metric and hence has a fixed point.

Let us take the distance

$$d(x,y) = \max |x_i - y_i|$$

Then denoting the row sum $r_i = \sum\limits_j |a_{ij}|$ and $r = \max r_i$

$$d(Tx,Ty) = \max_i \left| \sum_j a_{ij}(x_j - y_j) \right| \leq \max_i \{ \sum_j |a_{ij}| d(x,y) \} \leq r d(x,y)$$

Thus if $\max\limits_i \sum\limits_j |a_{ij}| < 1$ then T is a contraction in the d_∞ metric and hence has a fixed points.

Thus the main point is that if any one of the above conditions holds then there is a unique fixed point. Of course you can consider the d_p metrics too.

inverse function theorem:

Another application of the contraction mapping theorem is for proving the inverse function theorem. We proved it last year. We have an open set $\Omega \subset R^n$ and a C^1 function f defined on Ω to R^n. We are given a point $x_0 \in \Omega$ and are told that the derivative matrix f' at x_0 is non-singular. Then the conclusion is that there is an open set U with $x_0 \in U \subset \Omega$ and an open set $V \subset R^n$ such that f is one-one on U onto V and the inverse map on V onto U is again C^1. Of course there is an expected formula for the derivative of the inverse map.

You can look up that proof. After showing open sets U and V we had a small estimate which is reminiscent of contraction. We used it to conclude that for a point $y \in V$ there can not be two points $x_1, x_2 \in U$ with $f(x_1) = f(x_2) = y$. However, to show that there is a point x at all, we used a hands on calculation involving solution of linear equations. We could have used fixed point theorem. This was done in your calculus III course by Balaji. So we shall not repeat. Actually, you can look up this specific point in our notes, just to get a glimpse of how we grazed (and not used) the fixed point theorem.

differential equations:

We shall now discuss another important application. This application is for solving differential equations. In high school you discussed theory of equations. We have number, do not know what it is, but we know that it satisfies $x^2 - 5x + 6 = 0$. We need to find the number. In this case there is explicit formula, you can 'factorize' this polynomial. Sometimes there was no explicit formula but still you were able to deduce existence of solution using some other rules (Descarte's rule of signs or intermediate value theorem or whatever).

In the study of differential equations, you have a function; you do not know what it is; but you know that the function and its derivatives obey some relation. You need to find the function. Just as in the case of theory of equations it is in general difficult to find explicit formula for the function. Sometimes you can solve explicitly and get a formula for the function. But such a luxury is very rare. We should be satisfied if we know that there is a function at all satisfying the differential equation. We should be more than happy if we know how many such functions are there. If we succeed, we shall try to understand the solution more.

If some one asked you to solve $x'(t) = \sin t$ then it is very easy because the function is $x(t) = -\cos t + c$ for some number c. This was simple because of two reasons. firstly there is no appearance of the function on the right side. Secondly we could integrate the function. You can integrate whenever right side does not involve the function x (or its derivatives etc). Whether you can get a formula or not depends on the right side. For example $x'(t) = \exp\{-t^2\}$ will not allow you to come up with formula. But in these cases you are sure that the integral of the right side is a solution.

Serious problems arise if the right side also involves the function x. For example solve: $x'(t) = \sin x(t)$. We do not even know if there is a solution at all (yes, there is, you will see). There is no need for us to define what is a differential equation and what is meant by a solution. This is reserved for a later course. We consider a specific problem, so you will have no trouble following.

Differential equations arise in several contexts. More important are they in physics. Imagine that in the space in this room at every point there is a force. Of course, when I say force you will ask; on what is it acting. No, it is not yet acting — it has the potential to act. For example if there is a particle at the point P the force at that point will push it in a particular direction. suppose it is pushed to Q, then the force at that point Q will push it to R and the force at the point R will push it to etc and so on.

There is just one subtle point which is very important. I simply said the force at P will act on my particle and push it to Q. Actually when the particle is on its way to Q it will pass through intermediate points, but there are forces at these points and they do their bit as well to push. Thus the particle may actually not reach Q and it may be pushed to Q'. Pause and think about the situation our particle is in. since there are forces at every point, our particle is pushed *at every time instant*. In other words it is continuously pushed around. Life for our particle is not as discrete as I made it out in the earlier para.

So what is the problem. Well, I now place a particle at this point P. Tell me how it travels. Tell me the trajectory or path the particle takes. so you should tell me at every time instant t the position of the particle. We shall not consider in this generality and in three dimensions.

Another situation where these arise is in geometry. Basically, I want to draw a curve and I have been instructed 'how the curve should curve'. More precisely, at every point of the plane, R^2, a vector is given. I was given a starting point P. I should draw my curve starting at the given point P and at any point on the curve, the tangent is as prescribed at the point, remember there are placed vectors at every pint of R^2 and if your curve passes through a point then tangent to your curve at that point should be as suggested.

Now let us make matters precise. We are given an open set $U \subset R^2$ and a point $(t_0, x_0) \in U$. We are given a function $F : U \to R$. We are asked to locate an interval $(t_0 - \delta, t_0 + \delta)$ and a differentiable real valued function x on this interval satisfying two conditions: $x(t_0) = x_0$ (initial condition) and for every t in this interval $x'(t) = F(t, x(t))$ (differential equation).

You see the change of attitude. Our task is 'local'. In a small interval around the given point t_0 we should solve the problem. The idea is that once you do this you can 'continue' from where you 'reached'. Try to make sense of it. It does not concern us now because we are not going to carry it out and find out what is the largest interval on which the solution can be defined. We shall show locally and stop. Observe that $(t, x(t))$ is a point on the curve and we are wanting at that point of the curve the derivative should be same as value of F at that point.

We solve this problem assuming certain conditions. Here is the precise theorem.

Theorem (Picard):
Given
(i) Open set $U \subset R^2$ and a point $(t_0, x_0) \in U$.
(ii) $F : U \to R$ which is continuous and there is a number M such that $|F(t, x) - F(t, y)| \le M|x - y|$ for all points (t, x) and (t, y) in U.
Then *there exists*
an interval $(t_0 - \delta, t_0 + \delta)$ and a real valued differentiable function x on this interval such that its graph lies in U and at every point t in this interval $x'(t) = F(t, x(t))$.

Proof is very simple. fix a closed ball B with centre (t_0, x_0) such that $B \subset U$. Let K be an bound for F on the compact set B. Choose $\delta > 0$ so that
(i) $M\delta < 1$ and

(ii) $[t_0-\delta, t_0+\delta] \times [x_0 - K\delta, x_0 + K\delta] \subset B$.

Before we produce the function, let us make an observation that motivates the later considerations.

Solving

$x'(t) = F(t, x(t))$ and $x(t_0) = x_0$ $\quad (\spadesuit)$

is same as solving

$x(t) = x_0 + \int_{t_0}^{t} F(s, x(s))ds.$ $\quad (\clubsuit)$

Indeed, suppose ($\spadesuit$) holds. Then x being differentiable, must be continuous and F being continuous $F(t, x(t))$ is continuous. In other words x' is continuous. Then rules of integration tell us that

$$x(t) - x(t_0) = \int_{t_0}^{t} x'(t)dt = \int_{t_0}^{t} F(s, x(s))ds$$

which is ($\clubsuit$).

Conversely suppose ($\clubsuit$) holds. Then the fundamental theorem of integration gives ($\spadesuit$).

Consider the space X consisting of functions $x \in C[t_0-\delta, t_0+\delta]$ satisfying the following two conditions.

(i) $x(t_0) = x_0$ and

(ii) values of x lie in the interval $[x_0 - K\delta, x_0 + K\delta]$.

In other words the space of continuous functions on the said interval whose value at t_0 is as required and $|x(t) - x_0| \leq K\delta$ for every t in this interval.

The space $C[t_0-\delta, t_0+\delta]$ is a complete metric space — exactly like $C[0, 1]$ is complete. Its subset X we are considering is a closed subset. If a sequence of continuous functions satisfying the two conditions converge to a function in the space (so uniform convergence) then the limit function also satisfies the two conditions. Hence X is a complete metric space.

Let us define a map T of X to itself by

$$Tx(t) = x_0 + \int_{t_0}^{t} F(s, x(s)ds.$$

The above analysis suggests that if we find a fixed point of this map then it satisfies our requirements. Thus the proof is completed by showing firstly, that the map T we defined takes X to itself and secondly, it is a contraction.

Since F is continuous and indefinite integral is continuous we see that Tx is continuous function. Actually the indefinite integral is continuous even if integrand is not continuous, but we never proved such a theorem. We however proved that for a continuous function the indefinite integral is continuous. This is all what we need to see Tx is a continuous function.

Also from definition of Tx we see that its value at t_0 is indeed x_0. Further, since x takes values in the interval $[x_0 - K\delta, x_0 + K\delta]$ the graph of x lies in our rectangle $[t_0 - \delta, t_0 + \delta] \times [x_0 - K\delta, x_0 + K\delta]$ and hence the value of x is bounded by K. Thus

$$|Tx(t) - x_0| = \left| \int_{t_0}^{t} F(s, x(s)ds \right| \leq K|t - t_0| \leq K\delta$$

Thus $Tx \in X$ for each $x \in X$.

Finally take any $x, y \in X$ and $t \in [t_0 - \delta, t_0 + \delta]$. Let us assume $t > t_0$; similar argument applies if $t < t_0$.

$$|Tx(t) - Ty(t)| = \left| \int_{t_0}^{t} \{ \ F(s, x(s)) - F(s, y(s)) \ \} ds \right|$$

$$\leq \int_{t_0}^{t} |F(s, x(s) - F(s, y(s)| ds \leq \int_{t_0}^{t} M|x(s) - y(s)| ds$$

$$\leq \int_{t_0}^{t} Md(x, y) \ ds \leq M\delta d(x, y).$$

This shows $d(Tx, Ty) \leq (M\delta)d(x, y)$ and since $M\delta < 1$ we have proved that T is indeed a contraction.

This complete the proof.

Recall that a function g of one variable is said to be Lipschitz if there is a number M such that $|g(x) - g(y)| \leq M|x - y|$ for all x, y in its domain. The number is called the Lipschitz constant. Thus what we demanded, apart from continuity of F, is that it be Lipschitz in the second variable x for each fixed value for the first variable t. Further the Lipschitz constant does not depend on t, thus the same number M works for all t. This condition is expressed by saying that F is *Lipschitz in the second variable, uniformly in the first variable.*

You might be wondering what has such a theorem got to do with physics, because, the equations you come across there involve second derivatives. Remember, the simplest situation is $F = ma$ or $mx''(t) = F$. Here $x(t)$ is position of particle at time t and $x'(t)$ will give you velocity and $x''(t)$ gives the acceleration. so to know where the particle is you should solve a more complicated equation that we considered. I shall explain the trick.

Suppose you want to solve $x''(t) = -x(t)$. Of course, you know the solutions already. Solving for one function which involves second derivative is transformed to a problem of solving for two functions but involves only first derivative as follows. Solve for two functions x and y with

$$x'(t) = y(t); \qquad y'(t) = -x(t)$$

If you could solve the first problem and get x, take $y = x'$ to see you have a solution for the later problem. Conversely, if you have solution x, y for the second problem then x solves the first problem.

The method for solving for two functions is similar to the above, involves no new ideas but we shall not get into.

integral equations and iterations:

We shall show two other interesting flowers from the fixed point theory garden. These are a little advanced and we need tools like scissors or blade to pluck these flowers; bare hands will not do as in the earlier examples. So we only see and be happy.

We discussed conditions for an affine map to be contraction. More precisely we have an $n \times n$ matrix and an n vector b. We considered the following map on R^n.

$$Tx_i = \sum_j A_{ij}x_j + b_i; \quad i = 1, 2, \cdots, n.$$

Instead of denoting vectors by (x_i) let us denote as $\{x(t) : 1 \le t \le n\}$ and accordingly we denote the matrix by $A(s,t)$. Thus the map takes the form

$$Tx(s) = \sum_t A(s,t)x(t) + b(s); \quad 1 \le s \le n.$$

Its appearance suggests a natural interpretation. Think of x as a function!

More precisely, consider $C[0,1]$. Suppose you are given a function b in this space. suppose we are given a continuous function A on $[0,1] \times [0,1]$. Consider the problem of finding a function x on $[0,1]$ such that

$$x(s) = \int_0^1 A(s,t)x(t)dt; \quad 0 \le s \le 1.$$

Thus if you define Tx as the function on the right side above, then the problem is to find a fixed point of $Tx = x$.

The natural stage for this problem is not the set of continuous functions on $[0, 1]$ but functions x on $[0, 1]$ such that $x^2(t)$ is integrable. This can be solved using Banach fixed point theorem. Such problems are called integral equations, as the appearance itself suggests.

Here is another problem. What is the Cantor set? You are familiar with it. Start with interval $[0, 1]$ and delete middle one-third intervals repeatedly. Here is another way of looking at it.

Consider the two functions on the interval $[0, 1]$ into itself.

$$T_1(x) = \frac{1}{3}x; \qquad T_2(x) = \frac{2}{3} + \frac{1}{3}x.$$

Start with the 'seed' the two points $\{0, 1\}$. If you keep on applying the above maps what do you get?

$$K_0 = \{0;\ 1\}$$

$$K_1 = \{0;\ 1/3;\ 2/3;\ 1\}.$$

$$K_2 = \{0;\ 1/9;\ 2/9;\ 3/9;\ 6/9;\ 7/9;\ 8/9;\ 1\}.$$

What is happening to these sets? They are converging to the Cantor set! You will wonder how that can happen because each of these sets is finite and the limit should, at the best, be countable. Also you will wonder what this has got to do with contraction. Also, this appears like a complicated way of explaining Cantor set that we know so well! Is it of any use at all. These issues are what we explain now.

Consider the space X which consists of all non-empty closed subsets (equivalently, compact subsets) of $[0, 1]$. For example the set $\{0, 1\}$ or any finite set or Cantor set or the interval $[1/3, 1/2]$ etc are all elements of this set X. The set $\{1/n : n = 1, 2, 3, \cdots\}$ is not an element of this set but the set $\{0, 1/n; n = 1, 2, 3, \cdots\}$ is an element of this set.

Hausdorff defined a nice metric on the set X which makes it a complete metric space (actually compact metric). This is a nice metric: Two sets which appear to your eye close are close in this metric. We shall not get into precise definition.

Define the following map T on this space.

$$TK = T_1(K) \cup T_2(K) = \left\{ \frac{1}{3}x;\ \frac{2}{3} + \frac{1}{3}x; \quad x \in K \right\}.$$

It so happens that T is a contraction of this space X. If you start with the set K_0 described above then the successive iterations lead you to K_1, K_2 and so on. It is possible to prove that the sequence $T^n K_0$ converges to C, the Cantor set in the Hausdorff distance. Thus you get Cantor set as a result of iteration of a contraction.

So what is the use. I need to send you a picture of Cantor set. Scanning and sending it takes too much space. I can feed the map T and the seed and instruct iteration. Then the computer can iterate a large number of times and plot the resulting set. This will be an excellent approximation of the Cantor set. Eventhough not all points of the Cantor set are plotted, enough are plotted and your eye believes it is seeing Cantor set!

Of course, you might ask whether the construction process, namely, starting with an interval and deleting middle one-thirds could as well be followed. This process can be iterated a large number of times and the resulting set can be plotted. No matter how many times you iterate you will plot intervals and the picture will not reveal the true nature of Cantor set, after all, Cantor set does not contain any non-trivial interval.

Of course, I illustrated using a trivial set that you are familiar, but most important situation is when you want to describe beautiful designs. There is a beautiful theory behind manufacturing impressive designs.

Compact spaces:

We have been discussing completeness and its consequences. Let us pass to another topic which is equally important.

A metric space (X, d) is compact if the following is true: Given any collection of open sets whose union is X, we can pick finitely many of those open sets whose union equals X. In other words, given a collection of open sets that cover X, there are finitely many of those which also cover X. The first order of business is to relate ti to what we know in R^n.

Theorem 1: X is compact implies every sequence has a limit point.

Suppose a sequence (x_n) has no limit point. Then given any point x there is an open ball B_x such that $x \in B_x$ and which does not contain infinitely many terms of the sequence; that is, there is a stage after which no term of the sequence is in this ball. Consider such a ball for each point. These balls cover X. Pick finitely many of these balls which cover X. But then, there is a stage after which no term of the sequence is in any of these finitely many balls. But those points are in X! This contradiction proves the statement.

Theorem 2: X is compact implies every sequence has a convergent subsequence.

We know that there is a subsequence converging to limit point of the sequence. so this follows form the previous theorem.

Theorem 3: X is compact implies that every Cauchy sequence converges.

If a Cauchy sequence (x_n) has a subsequence that converges to a point x, then the sequence itself converges to x. This is easy and can be seen as follows. Let $n_1 < n_2 < n_3 < \cdots$ be such that

$$x_{n_1}, x_{n_2}, x_{n_3}, \cdots \to x.$$

We show $x_n \to x$. Let $\epsilon > 0$ be given. Fix k so that $d(x_{n_i}, x) < \epsilon/2$ for $i \geq k$. fix $N > n_k$ such that $d(x_n, x_m) < \epsilon/2$ for $m, n \geq N$. Let us take any m larger than N. Pick an n_i larger than N. Then

$$d(x_m, x) \leq d(x_m, x_{n_i}) + d(x_{n_i}, x) \leq \epsilon.$$

To prove the theorem, note that given any Cauchy sequence, the previous theorem says that there is a subsequence which converges. But then as seen above the sequence itself converges.

Theorem 4: X is compact implies the following. For any given $\epsilon > 0$, finitely many ϵ balls cover X. That is, there are finitely many balls whose radius is ϵ and their union equals X.

Here you can interpret balls as open balls or as closed balls, the statement is true.

Given $\epsilon > 0$ take open ball of radius ϵ around each point $x \in X$. Clearly all these open balls cover X. Since open ball is open set, compactness implies finitely many of these balls cover X.

Closed Balls with same centres and same radius ϵ too cover X.

Theorem 5: If X is compact then the following two conditions hold.
(i) X is complete.
(ii) For any given $\epsilon > 0$, finitely many ϵ balls cover X.
Conversely if (i) and (ii) hold then the metric space is compact.

If X is compact, then by Theorem 3 every Cauchy sequence converges and hence the space is complete. Second part is just the previous theorem.

Conversely let us assume that (i) and (ii) hold. We show X is compact. Fix any collection $\mathcal{U}$ of open sets whose union is X. We repeat exactly the same proof that we did in R^n.

Assume that no finite sub collection covers X. Take finitely many closed balls of radius one which cover X. If each one of these can be covered by finitely many sets from $\mathcal{U}$ then surely X can also be covered. Since this is not the case, fix one such ball B_1 which can not be covered by finitely many sets from $\mathcal{U}$.

Take finitely many closed balls of radius $1/2$ which cover X. Take their intersection with the above B_1. If each one of these can be covered by finitely many sets from $\mathcal{U}$ then surely B_1 can also be covered. Since this is not the case, fix one ball B_2 of radius $1/2$ so that $B_1 \cap B_2$ can not be covered by finitely many sets from $\mathcal{U}$.

Thus by induction, we can get a sequence of balls (B_i) such that
For each i, B_i is a closed ball of radius $1/i$ and
for each n, $B_1 \cap B_2 \cap \cdots \cap B_n$ can not be covered by finitely many sets from $\mathcal{U}$.

By completeness and Cantor intersection theorem you get a point $x \in \cap_{i \geq 1} B_i$. Since $\mathcal{U}$ covers X, pick $U \in \mathcal{U}$ with $x \in U$. Pick $\epsilon > 0$ with $B(x, \epsilon) \subset U$. Pick N so that $1/N < \epsilon/4$. Then we claim $B_N \subset U$. To see this first observe that distance between any two points of B_N is at most $2/N$ (go via centre). Also $x \in B_N$. Thus for any $y \in B_N$ we have $d(x, y) \leq 2/N \leq \epsilon/2$. In other words $y \in B(x, \epsilon) \subset U$.

Thus $B_1 \cap B_2 \cap \cdots \cap B_N$ is covered by just one set from $\mathcal{U}$. But our construction says this is not possible. This completes the proof that finitely many sets from $\mathcal{U}$ indeed cover X.

Note that R is complete but not compact. Thus condition (i) alone is not enough to deduce compactness. The metric space $X = (0,1)$ satisfies condition (ii) but is not compact. Thus condition (ii) alone is not enough to deduce compactness.

Theorem 6: X is compact iff every sequence has a convergent subsequence.

If X is compact then theorem 2 already shows that every sequence has a convergent subsequence.

Conversely, assume that every sequence has a convergent subsequence. We show X is compact. We verify the two conditions of the previous theorem.

To show completeness, take any Cauchy sequence, then hypothesis tells us that there is a convergent subsequence but then the sequence itself converges as observed earlier.

Let $\epsilon > 0$ be given. We need to show that finitely many ϵ balls cover the space. suppose it is false. Take any point x_1. Since $B_1 = B(x_1, \epsilon)$ is not all of X, pick $x_2 \notin B_1$ and let $B_2 = B(x_2, \epsilon)$. since $B_1 \cup B_2$ is not all of X, pick $x_3 \notin B_1 \cup B_2$. proceeding in this way we pick a sequence of points

$$x_1, x_2, x_3, \cdots; \quad x_n \notin \bigcup_{i \leq n-1} B(x_i, \epsilon).$$

Clearly distance between any two points of the sequence is at least ϵ. This sequence has no convergent subsequence. If it has, say converging to p then $B(p, \epsilon/4)$ should contain at least two terms of the sequence, which would mean distance between those two points is at most $\epsilon/2$.

This completes the proof.

We shall now discuss compactness of subsets. Let (X, d) be a metric space. Let $Y \subset X$. Recall that we can regard Y itself as a metric space; metric being the restriction of d to $Y \times Y$.

Theorem 7: The following two statements are equivalent.

(i) Given any family of subsets of Y which are open in Y there are finitely many whose union equals Y.

(ii) Given any family of open subsets of X whose union includes Y, there are finitely many of those whose union includes Y.

The importance of this theorem is the following. When you re talking about compact subsets of a pace X, any of the above two statements can be used as definition. The statement (i) forgets the background and says that the metric space (Y, d) is compact. On the other hand statement (ii) does not forget the background and states in terms of the open sets of X, does not refer to open subsets of Y at all.

To prove the theorem we only need to observe the following;
(*) A set $U \subset Y$ is open in Y iff there is a set $V \subset X$ which is open in X such that $U = V \cap Y$.

Assume, for a moment, truth of the above statement. We prove the theorem as follows. Let (i) hold. Let (V_α) be a family of sets open in X whose union includes Y. Then $U_\alpha = V_\alpha \cap Y$ are open in Y and cover Y and finitely many U_α cover Y and then the corresponding V_α cover Y.

Conversely let (ii) hold and (U_α) be a collection of sets open in Y whose union is Y. For each α pick a set V_α which is open in X and $U_\alpha = V_\alpha \cap Y$. these (V_α) cover Y and so finitely many of them cover Y and the corresponding U_α cover Y.

Proof of (*) is routine. Suppose V is open in X and let $U = V \cap Y$. Need to show U is open in Y. Let $y \in U$. So $y \in V$. Since V is open in X there is a ball in X, say $B_X(y, r) \subset V$. The ball in Y of radius r, that is $B_Y(y, r)$ is nothing but $B_X(y, r) \cap Y$ and is hence contained in U. This shows U is open.

Conversely, let U be an open set in Y. We need to exhibit a set V open in X such that $U = V \cap Y$. For each $y \in U$ we get a ball in Y, say, $B_Y(y, r) \subset U$. This r depends on the point y, we did not show it in the notation. Let $B_X(y, r)$, the ball in X with the same centre be denoted by A_y. Thus remember $A_y \cap Y = B_Y(y, r)$. Union of all these balls A^y be denoted by V. Then V is open in X and $V \cap Y = U$.
This completes the proof.

You should carefully understand the above proof. It is actually trivial, but you can say so only if you understood it.

continuity:

So far we have been taking about sets — closed, open, compact, connected, first category, second category etc.

We shall now study functions between metric spaces.

Let (X, d) and (Y, ρ) be two metric spaces and $f : X \to Y$. We say that f is continuous if it preserves convergence. That is, $x_n \to x$ in X implies that $f(x_n) \to f(x)$ in Y.

Remember our main idea in taking up metric spaces is the feeling that if we have a concept of 'how close things are' we can do at least a part of calculus that deals with convergence and continuous functions.

After all $x_n \to a$ meant that x_n is getting closer and closer to a. Further, our idea of continuous function on R to R is that: when x is close to a then $f(x)$ should be close to $f(a)$. This was made precise by saying that whenever $x_n \to a$ then $f(x_n) \to f(a)$.

We shall now explore continuous functions and their properties.

Continuous functions:

Let (X, d) and (Y, ρ) be metric spaces and $f : X \to Y$ be a function and $a \in X$. We say that f is continuous at a if the following is true: $x_n \to a$ in X implies $f(x_n) \to f(a)$ in Y. We say f is continuous function if it is continuous at every point $a \in X$.

All the results that we did in calculus concerning functions on R are all true with nearly the same proofs.

Theorem 1: f is continuous at a iff the following holds: for every $\epsilon > 0$ there is a $\delta > 0$ such that $\rho(f(x), f(a)) < \epsilon$ whenever $d(x, a) < \delta$.

Suppose the condition holds. To show f is continuous at a take any sequence $x_n \to a$. We need to show that $f(x_n) \to f(a)$. So fix $\epsilon > 0$. We shall show a stage after which $\rho(f(x_n), f(a)) < \epsilon$. With this $\epsilon > 0$ find $\delta > 0$ as assured. Since $x_n \to a$ we have $d(x_n, a) < \delta$ after some stage and clearly then $\rho(f(x_n), f(a)) < \epsilon$.

To prove the converse, assume that the condition fails. Fix an $\epsilon > 0$ for which we can not find $\delta > 0$ satisfying the condition. Since $\delta = 1/n$ does not satisfy the requirement we can pick x_n so that $d(x_n, a) < 1/n$ and yet $d(f(x_n), f(a)) \geq \epsilon$. Clearly $x_n \to a$ but $\rho(f(x_n), f(a)) \geq \epsilon$ for all n. So f is not continuous at a.

Theorem 2: f is continuous at a iff the following holds: For any set $V \subset Y$ which includes a ball around $f(a)$ the inverse image $f^{-1}(V)$ includes a ball around a.

suppose the condition holds. Then we verify the statement of the previous theorem. So let $\epsilon > 0$ be given for which we need to find $\delta > 0$. Take V to be the ball of radius ϵ around $f(a)$ and use condition to get a ball $B(a, \delta) \subset f^{-1}(V)$. This δ will do.

To prove the converse, suppose that the statement of the previous theorem holds. We show that condition of the present theorem holds. Take V as stated, suppose it includes ball of radius $\epsilon > 0$. get a $\delta > 0$ as in the previous theorem. Observe that $f^{-1}(V)$ includes $B(a, \delta)$.

Theorem 3: f is continuous iff for every open set $V \subset Y$, the set $f^{-1}(V)$

is open in X.

This is immediate from previous theorem. Indeed, suppose f is continuous. Let V be open in Y. To show $f^{-1}(V)$ is open in X take any point $a \in f^{-1}(V)$. Then $f(a) \in V$ and V being open, contains a ball around $f(a)$ so by previous theorem $f^{-1}(V)$ includes a ball around a. Thus $f^{-1}(V)$ includes a ball around each of its points. that is, it is open.

to prove the converse, suppose that condition of the theorem holds. To show f is continuous at say $a \in X$, take $x_n \to a$. Need to show $f(x_n) \to f(a)$. Take open ball of radius $\epsilon > 0$ with centre $f(a)$. Inverse image of this ball is open and also contains a and so by from hypothesis, includes a ball around a, say, $B(a, \delta)$. Since $x_n \to a$ we conclude that $x_n \in B(a, \delta)$ after some stage. But then after that stage $f(x_n)$ is in the ϵ ball around $f(a)$. Thus $f(x_n) \to f(a)$.

Theorem 4: f is continuous iff for every closed set $C \subset Y$, the set $f^{-1}(C)$ is closed in X.

f is continuous and C^c is open implies $f^{-1}(C^c)$ is open and hence $f^{-1}(C)$ is closed. Conversely, Given condition implies that for any open set $f^{-1}(U^c)$ is closed and hence $f^{-1}(U)$ is open.

Theorem 5: (i) If f and g are real valued continuous functions on X, then so are $f + g$, fg and $32f$.

(ii) Fix a point $z \in X$. The function $f(x) = d(z, x)$ is a continuous function. More generally, if $A \subset X$ is a non-empty set then $f(x) = d(x, A)$ is a continuous function. Recall $d(x, A) = \inf\{d(x, z) : z \in A\}$.

(iii) In particular, given a closed set C, there is a non-negative continuous function which takes the value exactly for points in the closed set. Equivalently, given an open set U there is a non-negative continuous function which takes strictly positive values exactly on U.

First part follows because convergence in R respects addition, multiplication etc. Actually the same is true for functions taking values in R^n.

Second part has already been noted earlier as a consequence of triangle inequality.

Third part follows because $d(x, C) = 0$ iff $x \in C$; when C is a closed set. If $C = \emptyset$ take the constant function 1.

Theorem 6: If $f : X \to Y$ is continuous and X is compact, then $f(X)$ is compact.

If you have a collection $\mathcal{U}$ of open sets in Y covering $f(X)$ then the family of open sets $\{f^{-1}(U) : U \in \mathcal{U}\}$ covers X and take a finite family that covers X, say, $\{f^{-1}(U_i) : 1 \leq i \leq k\}$ then the finite family $\{U_i : 1 \leq i \leq k\}$ covers $f(X)$. (You should verify this and not simply reproduce this sentence).

In particular, since we know that compact sets are precisely closed and bounded sets we conclude as a special case of the above theorem: if f is a continuous function defined on a closed bounded subset of R, then the range of the function is again closed and bounded. In particular, it attains its bounds too, as they are part of the range.

Theorem 7: If $f : X \to Y$ is continuous and X is connected, then $f(X)$ is connected.

Observe that $f : X \to f(X)$ is a continuous function. If $f(X)$ is not connected, then there is a non-empty proper subset $A \subset f(X)$ which is both closed and open. Clearly its inverse image is both closed and open and non-empty and proper subset of X.

In particular, since we know that the only connected subsets of R are intervals we conclude the following: Range of a real valued continuous function defined on an a connected metric space is again an interval. In particular range of a continuous function defined on an interval in R is again an interval. as a result the function has intermediate value property, that is, if $f(x) = a$ and $f(y) = b$ and $a < c < b$ then there is a point z such that $f(z) = c$ and $x < z < y$.

Homeomorphism:

A function $f : X \to Y$ is called a homeomorphism if it is one-one, onto, continuous and inverse map is also contious from Y to X.

Remember, homeomorphism need not preserve the metric. for example the function $f(x) = 1/x$ is a homeomorphism of $(0, 1)$ onto $(1, \infty)$ that does not preserve the distance.

Then why is it called homeomorphism? Generally when you say isomorphism or homomorphism etc they preserve the structure that we have. Homomorphisms preserve 'forward' and isomorphisms preserve both ways (for this to make sense isomorphism has to be bijections and etc).

Metric on a space has provided us collection of open sets, which is not visible in the notation. This structure is preserved by a homeomorphism. That is, a homeomorphism is a bijection f with the property: $U \subset X$ is open in X iff $f(U) \subset Y$ is open in Y. Equivalently, the collection of closed sets is preserved. Thus homeomorphism is a bijection f such that the following holds: $C \subset X$ is closed in X iff $f(C)$ is closed in Y. Equivalently convergence is preserved. Thus homeomorphism is a bijection f with the property: $x_n \to x$ in X iff $f(x_n) \to f(x)$ in Y.

Thus, if two spaces are homeomorphic then by renaming, one looks like the other. Renaming x as $f(x)$ the space X looks like Y. But please do remember, this appearance is only for convergence; not for actual distances.

Homeomoprhism is an equivalence relation among metric spaces. Aha, the collection of metric spaces is NOT a set! (Do not waste time on this point.) What we mean is the following. If you take any specific collection of metric spaces, then among the collection homeomorphism is an equivalence relation. We can state the same property without talking about collections at all as follows.

X is homeomorphic to X (identity map).

If X is homeomorphic to Y then Y is homeomorphic to X (take inverse map).

If X is homeomorphic to Y and Y is homeomorphic to Z, then X is homeomorphic yo Z (take composition of maps).

Here are some examples:

The space $[0,1]$ is not homeomorphic to any of the spaces $[0,1)$; $(0,1]$; $(0,1)$. Because the former is compact and the later are not compact.

The space $(0,1)$ is not homeomorphic to $[0,1)$; $(0,1]$. There is one end point in the later spaces so that even after removing that point the space is still connected. But in the space $(0,1)$ removal of any one point makes it disconnected. More precisely, if $f : [0,1) \to (0,1)$ is homeo, then f restricted to $(0,1)$ in the domain is a homeomorphism onto $(0,1) - \{f(0)\}$. The domain is connected but the range is not.

Of course the spaces $[0,1)$ and $(0,1]$ are homeomorphic, for example $f(x) = 1 - x$ is one such. This is not the only map which exhibits homeomorphism between these two spaces.

Are any of the above spaces homeomorphic to the unit circle
$S^1 = \{(x, y) : x^2 + y^2 = 1\}$.

No. This is not homeomorphic to $[0, 1]$ because there are two points in the later, namely the end points, so that the space remains connected even after removing those. However removal of any two distinct points makes S^1 disconnected. Of course S^1 is compact but $(0, 1)$ and $[0, 1)$ are not compact.

Of course given bounded interval (a, b) the map $f(x) = a + x(b - a)$ sets up a homeomorphism from the interval $(0, 1)$ to (a, b). The interval $(0, 1)$ is homeomorphic to $(1, \infty)$ by the map $g(x) = 1/x$. The interval $(0, \infty)$ is homeomorphic to $(-\infty, 0)$ via the map $h(x) = -x$. Also $(0, \infty)$ is homeomorphic to (a, ∞) using $\varphi(x) = a + x$. Now by using compositions we can easily deduce the following: Any two (non-empty) open intervals in R are homeomorphic, bounded or unbounded.

We can not say the same about closed intervals. Any two closed bounded intervals (non-degenerate) are homeomorphic. Obviously, a closed bounded interval, which is compact, can not be homeomorphic to unbounded closed interval. Also even among unbounded closed intervals, you can show that $[1, \infty)$ is not homeomorphic to $(-\infty, \infty) = R$. you can do this by arguing that removal of a point still keeps the first set connected where as removal of any one point makes R disconnected. You can also argue as follows. First showing that a homeo must be either strictly increasing or strictly decreasing and thus image of any homeo can only go in one direction from the image of the end point.

Consider the set

$$S = \{(x, \sin(1/x)) : 0 < x \leq 1\} \cup \{(0, y) : -1 \leq y \leq 1\}$$

This is not homeomorphic to any of the above spaces. This is a compact space. This is also connected. The earlier ones are all path-connected but this is not. We shall understand this now.

S being a closed bounded subset of R^2 it is compact. To show that S is connected, you can, for example, say the interval $(0, 1]$ is compact and its continuous image $x \mapsto (x, \sin(1/x))$ is compact and hence its closure, namely S, is connected.

Here we used the following fact. Suppose that we have a subset $Y \subset X$. Suppose that Y is connected. Let Y^* be the closure of Y in the space X.

Then Y^* is also connected. In fact, if $A^* \subset Y^*$ is a proper non-empty set which is both closed and open in Y^*, then $A = A^* \cap Y$ is both open and closed in Y. But then Y being connected we should have either $A = \emptyset$ or $A = Y$. If $A = Y$ then the $A^* \supset Y$. But A^* is closed in Y^* so it must be all of Y^*. Similarly, if $A = \emptyset$ then $A^* = \emptyset$

Of course, you can show S is connected with bare hands as follows. Let $A \subset S$ be non-empty subset which is both closed and open. Suppose it has a point $(a, \sin(1/a))$ then the entire curve should be in A. In fact, if some $(b, \sin(1/b))$ is not in A then consider all t between a and b such that $(t, \sin(1/t)) \in A$ and arrive at a usual (?) contradiction by considering the sup or inf of this set of points t. If the entire curve is in A then A being closed, it must include the other points of S from the Y-axis too. Note that each point $(0, y)$ in S is limit of a sequence of points on the curve.

However S is not path connected. For example there is no path joining $P = (1, \sin 1)$ to $Q = (0, 0)$. Prove this statement, if needed read earlier para again. Remember path in S means continuous function φ on $[0, 1]$ with values in S. It is path joining two points P and Q means that $\varphi(0) = P$ and $\varphi(1) = Q$.

You can ask whether R and R^2 are homeomorphic. You can use the argument involving connectedness and deleting point to argue that the answer is no. The same holds for R and any R^n for $n > 1$. It is also true that R^m and R^n are not homeomorphic if $m \neq n$, but the proof is not so simple.

Homeomorphism is an important concept in understanding spaces. For example if you have understood a group, then you have understood all groups isomorphic to it. If you have understood a vector space, then you have understood all vector spaces that are isomorphic to it. similarly, if you have understood a metric space, then you have understood all spaces homeomorphic to it; in some sense. Remember homeomorphism may not preserve distance. Thus understanding here means convergence and continuous functions.

separable spaces:

Recall that a metric space is separable if there is a countable set D such that every non-empty open sets contains a point of D. Such a set D is called dense set.

Theorem: A compact metric space is separable.

Let (X, d) be a compact metric space. For each n finitely many balls of radius $1/n$ cover the space, take such finitely many balls (your choice) and let F_n be their centres. Note that given any point $x \in X$, there is a point $p \in F_n$ such that $d(x, p) < 1/n$.

Let $D = \cup F_n$. Each F_n being finite, we conclude that D is countable. Given any point $x \in X$, there is a point $p_n \in F_n$ such that $d(x.p_n) < 1/n$. In other words $p_n \to x$. thus every ball around x contains points from D. This wing true of every point x we see that every non-empty open ball contains points from D. Thus D is dense, showing that a compact metric space is separable.

Consider R, real line and the collection $\mathcal{B}$ of open intervals with rational end points. Then we have the following properties: (i) the family $\mathcal{B}$ is countable and (ii) every open set is union of some of these intervals from $\mathcal{B}$.

Indeed the fact that there are only countable many pairs of rational tells us that $\mathcal{B}$ is countable. Further given any open set V and a point $x \in V$, there is an interval $(x - \epsilon, x + \epsilon) \subset V$; taking rationals a and b with $x - \epsilon < a < x$ and $x < b < x + \epsilon$ and setting $J = (a, b)$ we see $J \in \mathcal{U}$ and $x \in J \subset V$. In other words V is union of all intervals in $\mathcal{B}$ which are contained in V.

We now show that similar result is true in metric spaces.

Theorem: Let (X, d) be a separable metric space and D be any countable dense set. Let

$$\mathcal{B} = \{B(p, 1/n) : p \in D; n \geq 1.\}$$

then the family $\mathcal{B}$ is countable and every open set is union of some sets from $\mathcal{B}$.

In other words, using open balls of radius $1/n$ ($n \geq 1$) around points of D we can describe all open sets.

The fact that D is countable tells that $\mathcal{B}$ is a countable collection. Take any ball $B(x, r)$. Pick n such that $(2/n) < r$. D being dense pick $p \in D$ with $d(x, p) < 1/n$. Observe that if $y \in B(p, 1/n)$ then

$$d(x, y) \leq d(x, p) + d(p, y) \leq 2/n < r.$$

In other words $x \in B(p, 1/n) \subset B(x, r)$. as a result, if you take any open set V and any $x \in V$ then use the fact that $B(x, r) \subset V$ for some $r > 0$ to conclude that there is an $B(p, 1/n)$ such that

$$x \in B(p, 1/n) \subset V.$$

In roher words V is union of all balls from $\mathcal{B}$ which are contained in V.

compactness revisited:

Before discussing metric spaces, one motivation was that, if we have concept of convergence then we can discover more and more points in the space. Let us fully recall what it means.

In R we are comfortable with rational numbers because they are described simply as ratio of integers. We somehow knew that the sequence

$$1,\ 1+1,\ 1+1+\frac{1}{2!},\ 1+1+\frac{1}{2!}+\frac{1}{3!},\ \cdots$$

converges and we named its limit as e and discovered that it was not something known to us earlier. Similarly, the sequence

$$\frac{1}{n}+\frac{1}{n+1}+\frac{1}{n+2}+\frac{1}{n+3}+\cdots+\frac{1}{2n}; \quad n=1,2,3,\cdots$$

converges. We also know that

$$1+\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{n}-\log n; \quad n=1,2,3,\cdots$$

also converges. Both the sequences described above have the same limit. We named it γ. We do not yet know whether we have discovered a new (non-rational) number.

In the same way in a metric space we know some points and using the notion of convergence we can discover many other points. If the space is compact, then manufacturing a sequence will lead to a subsequence which converges and hence a point is discovered.

For example, in the space $C[0,1]$ we know that polynomials are dense (remember, unless stated otherwise, the metric is sup metric). Suppose I some how produce a convergent sequence of polynomials, then I can talk about limit of this sequence, thus obtaining a legitimate element of $C[0,1]$. Hopefully, I can discover properties of this new function by looking at the known sequence of polynomials which are converging to it.

In fact, exponential function, sine and cosine functions are obtained as limits of polynomials. Remember our discussion on power series.

But how do I know whether some sequence is converging or not? One possibility is that the space is complete and your sequence is Cauchy sequence. But to show something is Cauchy involves estimating distance between elements of the sequence. Another easier way is to show that the sequence lies in a compact set. Of course, this does not imply that the sequence converges, in general it does not. But there are convergent subsequences.

So it is important to know what sets are compact and what are not. For example in R or R^n compact sets are precisely closed and bounded sets. The characterization of compact metric spaces is too general and we would now specialise to one particular example.

Arzela-Ascoli:

Let us consider the space $C[0, 1]$. How do you know if a given subset of this space is compact? Answer to this question is provided by a theorem that was proved (independently) by two Italian mathematicians. Arzela and Ascoli. This is what we discuss now.

But first observe the following. For each $n = 2, 3, 4, \cdots$ define a continuous function x_n on $[0, 1]$ as follows. $x_n(t)$ starts at zero when $t = 0$ and increases with slope $2n$ till it reaches 1 at $t = 1/2n$ and then decreases the same way reaching zero at $t = 1/n$ and it stays at that value from then on. Let C be the set consisting these functions $\{x_n; n \geq 2\}$.

All these functions are non-negative and are bounded by one. Thus this is a bounded set. Remember a set is bounded if there is a number M such that distance between any two points of the set si at most M. Also this is a closed set. Indeed, if there is any limit point, equivalently, if there is a subsequence that converges to some x then it must also converge pointwise too and hence $x \equiv 0$. But all these functions are far from zero, indeed, $d(x_n, 0) = 1$ for each n. So there is no limit point.

Thus this set is closed and bounded but not compact. We need more than 'bounded'. This is what we discover now.

Home Assignment:

Given a compact set K contained in an open set U, **show** an $r > 0$ such that

$$\bigcup_{x \in K} B(x, r) \subset U.$$

Solution : Suppose not. then for every n there are points $x_n \in K$ and $y_n \in U^c$ such that $d(x_n, y_n) < 1/n$. Since K is compact there is a subsequence of (x_n) which converges to a point, say x of K. Let us take (x_n) to be the subsequence itself. If you do not like then use (x_{n_k}) below instead of (x_n). But then

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

In other words, $y_n \to x$ and U^c being closed we conslude that $x \in U^c$. But $x \in K$ and $K \subset U$. This contradiction proves the result.

OR: For each $x \in K$ pick $r_x > 0$ so that $B(x, 2r_x) \subset U$. Now consider the balls

$$\{B(x, r_x) : x \in K\}$$

covers K and hence finitely many of them cover K. Say the balls around $x_1, x_2, \cdots x_k$. Take

$$r = \min\{r_{x_1}, r_{x_2}, \cdots, r_{x_k}\}$$

We say this will do. To argue our claim, take $x \in K$ and take $y \in B(x, r)$. We need to show $y \in U$. But y is in one of the selected finitely many balls, say, ball around x_i. Thus

$$d(x, x_i) < r_{x_i}$$

But then

$$d(y, x_i) \leq d(y, x) + d(x, x_i) \leq r + r_{x_i} \leq 2r_{x_i}$$

and we know ball of radius $2r_{x_i}$ around x_i is contained in U. Thus $y \in U$.

OR: Consider the function

$$f(x) = d(x, U^c).$$

We know it is a continuous function. Since K is compact it has a minimum on K and the minimum is attained. Note that since $K \subset U$ we see that

$f(x) > 0$ for each point $x \in U$. Thus it is non-zero at points of K. Thus this minimum of f on K must be $r > 0$. This will do.

K is compact and C is closed and $K \cap C = \emptyset$ then **show** $d(K, C) > 0$.

If $U = C^c$ then U is open; hypothesis implies that $K \subset U$ and previous exercises completes the proof. Then $d(K, C) \geq r > 0$ where r is as obtained above.

Is the above statement true if both sets are given to be closed.
Consider the two sets contained in R:

$$A = \{1, \quad 2, \quad 3, \quad 4, \quad 5 \cdots\}$$

$$B = \{2 + \frac{1}{2}, \quad 3 + \frac{1}{3}, \quad 4 + \frac{1}{4}, \quad 5 + \frac{1}{5} \cdots\}$$

Of course, you could think of high school pictures too. You must have drawn hyperbola many times and also learnt the word *asymptote*. Here are subsets of R^2.

$$A = \{(x, y) : x > 0, y > 0, xy = 1.\}$$

$$B = \{(x, y) : y = 0\}$$

C_n is decreasing non-empty closed sets. **show** $\cap C_n$ is non-empty.
Solution: Take a point $x_n \in C_n$. Since all these are in C_1, a compact set it has a convergent subsequence, say, converging to x. Since the sequence is contained in C_1 we see $x \in C_1$. After the first term all terms of the subsequence are in C_2 and hence $x \in C_2$. In general after the k-th term all terms of the sequence and subsequence are in C_k and hence $x \in C_k$.

OR: If the intersection is empty, then every point of C_1 is outside some C_n. In other words

$$C_2^c, \quad C_3^c, \quad C_4^c, \cdots$$

cover C_1 and hence finitely many of them cover. Since C_i are decreasing, the complements are increasing. Thus one of these covers C_1. Say $C_j^c \supset C_1$, or $C_j \subset C_1^c$. But $C_j \subset C_1$. The only possibility is that $C_j = \emptyset$, contradiction.

Is this true if all the sets are closed, non-empty decreasing.
Take $[n, \infty)$ subsets of R.

is this true if all the sets are closed non-empty decreasing but one is compact. Yes, because if C_k is compact, consider the sets only after the k-th stage. Observe that closed subset of a compact set is compact. Thus all sets after the k-th stage are compact.

How do we show that a closed subset C of a compact space K is compact? Several ways. If you are given a collection of open sets which cover C, include C^c and say now K is covered; take a finite sub cover and delete C^c from this (if it is here). The remaining finitely many cover C. Alternatively, take a sequence from C, since it is also a sequence from K take a converging subsequence, but C being closed the limit of the subsequence must be in C already.

In a compact space a family of closed sets is given. Every finitely many of them have a point in common. **show** all the sets have a point in common.

If not, the grand intersection is empty; equivalently their complements (which are open) cover the space; get finitely many of them which cover and get contradiction (?).

If a metric space satisfies the above condition then **show** the space is compact. This is precisely definition of compactness if you take complements. Think.

If every bounded real valued continuous function attains its supremum, then **show** every such function attains its infimum too.

Take bounded real f. Take $g(x) = -f(x)$. then g is also bounded continuous function, use hypothesis about sup for g and observe that gives inf for f.

When every real bounded continuous function attains its bounds then **show** the space is compact.

First we show that the space is complete. Denote the space by X. Take a Cauchy sequence (x_n) Consider the function $f(x) = \lim d(x, x_n)$. The limit exists because the sequence of real numbers $\{d(x, x_n)\}$ is Cauchy. This is because

$$|d(x, x_n) - d(x, x_m)| \leq d(x_n, x_m) \to 0 \quad m, n \to \infty.$$

f is continuous because $|f(x) - f(y)| \leq d(x, y)$. Unfortunately f may not be bounded. So take $g(x) = \min\{f(x), 1\}$. Thus g is bounded, continuous, non-negative. Its infimum is zero. Given $0 < \epsilon < 1$ (get N so that for $d(x_m, x_n) < \epsilon$ for all $n, m > N$. Observe $f(x_N) = \lim d(x_N, x_n) \leq \epsilon$ and hence $g(x) = f(x) \leq \epsilon$. Thus g being non-negative we conclude zero is its infimum. This is attained at say x_0. Clearly then $\lim d(x, x_n) = 0$ or $x_n \to x$.

fix $\epsilon > 0$. Shall show we can cover X by finitely many balls of radius ϵ. if not you can get a sequence $\{x_n\}$ such that distance between any two points is at least ϵ. Let

$$A_n = \{x_n, \quad x_{n+1}, \quad x_{n+2}, \quad \cdots\}.$$

$$f_n(x) = d(x, A_n); \quad g_n(x) = \min\{f(x), 1/2^n\}; \quad n = 1, 2, 3, \cdots.$$

Then each f_n and g_n are continuous functions, so is $g(x) = \sum g_n(x)$ (why?). Note $g(x) > 0$ at all points. Since all A_n are closed subsets we see that g_1 itself is positive at all points outside A_1. And $g_k(x_n) > 0$ for each $k > n$. Thus g is positive at points of A_1 too. Since $g_k(x_n) = 0$ for $k = 1, 2, \cdots, n$ we see

$$g(x_n) \leq \frac{1}{2^{n+1}} + \frac{1}{2^{n+2}} + \cdots = \frac{1}{2^n}.$$

In other words infimum of g is zero but is not attained at any point.

OR: Instead of separately proving two parts as above you can start by saying suppose there is a sequence without any convergent subsequence. repeat the second step above.

If every real continuous function is bounded, **show** that the space is compact.

Repeat same argument above, take $1/g$.

If you carefully observe, we have repeated our old proofs on the real line. You can first complete the space and also argue as I explain din class. Think about it.

back to Arzela-Ascoli:

Consider $C[0, 1]$ with sup metric. Let K be a compact subset. Let us see some of its properties.

(i) K is closed.

We knew it, but let us recall. If a is a limit point of K then every open ball $B(a, 1/n)$ and hence closed ball $\overline{B}(a, 1/n)$ contains points of K. Since $a \notin K$, complements of these closed balls cover K and can not have finitely many of them covering K.

Or, you can also say take any sequence in K, then it contains a subsequence converging to a point of K, so if the sequence itself converges to a then $a \in K$. Thus K is closed.

Since $[0, 1]$ is a compact subset of R, every continuous function on it is bounded. That is, given $x \in C[0, 1]$ there is a number M such that $|x(t)| \leq M$ for all t. Even if you take finitely many continuous functions then there is one bound for all of them. In general, if you take infinitely many functions then there may not be a common bound for all these functions. But if you have a compact family then there is a common bound.

(ii) There is a number M such that $|x(t)| \leq M$ for all $x \in K$ and all $0 \leq t \leq 1$.

Indeed we know that the function

$$\varphi(x) = d(x, 0) = \max |x(t)|$$

is a continuous function on $C[0, 1]$, in particular, on K. Thus K being compact this function must be bounded on K, that is, there is a number M such that $\varphi(x) \leq M$ for all $x \in K$. This is what we wanted.

Since $[0, 1]$ is a compact subset of R, every continuous function on it is uniformly continuous. In other words, let $x \in C[0, 1]$. Given $\epsilon > 0$, there is a number $\delta > 0$ such that $|x(t) - x(s)| < \epsilon$ for all t, s with $|t - s| < \delta$. Even if you take finitely many continuous functions then there is one $\delta > 0$ which works for all of them. In general, if you take infinitely many functions then there may not be a common $\delta > 0$ which works for all them. But if you have a compact family then there is one $\delta > 0$ that works for all of them..

(iii) Given $\epsilon > 0$, there is a $\delta > 0$ such that $|x(t) - x(s)| < \epsilon$ whenever $x \in K$ and $|s - t| < \delta$. (Of course, through out it is assumed that s and t are in $[0, 1]$).

Proof is simple. Fix $\epsilon > 0$. Remember, K being compact, there are finitely many balls of radius $\epsilon/3$ that cover K; say

$$B(x_i, \epsilon/3); \quad 1 \leq i \leq n.$$

Fix $\delta > 0$ such that

$$|s - t| < \delta \Rightarrow (\forall i;\ 1 \leq i \leq n)|x_i(s) - x_i(t)| < \epsilon/3.$$

This does. For, given now any $x \in K$ and any s, t with $|s-t| < \delta$; let us pick i so that $x \in B(x_i, \epsilon/3)$. remember, this means $|x(u) - x_i(u)| < \epsilon/3$ for all u. Thus

$$|x(s) - x(t)| \leq |x(s) - x_i(s)| + |x_i(s) - x_i(t)| + |x_i(t) - x(t)| \leq 3 \times \epsilon/3 = \epsilon.$$

Arzela-Ascoli Theorem Let $K \subset C[0,1]$. Then K is compact iff the followng three conditions hold.

(i) K is closed.

(ii) There is a number M such that $|x(t)| \leq M$ for all $x \in K$ and all $0 \leq t \leq 1$.

(iii) Given $\epsilon > 0$, there is a $\delta > 0$ such that $|x(t) - x(s)| < \epsilon$ whenever $x \in K$ and $|s-t| < \delta$. (Of course, through out it is assumed that s and t are in $[0,1]$).

Moreover, if conditions (ii) and (iii) hold for a set K, then they hold for its closure and hence its closure is compact.

Last sentence is obvious. Indeed if a sequence of functions converges in the space, then they converge point-wise too. Thus the same bound works and the same δ works.

We have already observed that if K is compact, then the three conditions hold. Let us now assume that K satisfies the three conditions and show that it is compact. So take a sequence (x_n) in K. We need to exhibit a subsequence converging to a point of K. Of course if we just exhibit a converging subsequence then it must converge to a point of K because of condition (i).

Since the space $C[0,1]$ is complete, it is enough to exhibit a subsequence which is Cauchy. This is achieved in two steps.

(A) there is a subsequence

$$x_{n_1}, \quad x_{n_2}, \quad x_{n_3}, \cdots\cdots$$

such that for every rational number $r \in [0,1]$ the sequence of numbers

$$x_{n_1}(r), \quad x_{n_2}(r), \quad x_{n_3}(r), \cdots\cdots$$

converges.

(B) the sequence in (A) is Cauchy.

Let us prove (B) first. To avoid ugly notation, let us rename $y_1 = x_{n_1}$, $y_2 = x_{n_2}$, and so on. Thus (y_n) is a sequence of elements in K and for every rational number r, the sequence of numbers $\{y_n(r)\}$ converges. Let us show (y_n) is Cauchy.

Fix $\epsilon > 0$. Use condition (iii) with $\epsilon/3$ and get a $\delta > 0$. Take partition

$$0 = r_0 < r_1 < r_2 < \cdots < r_k = 1$$

where each r_i is rational and $r_{i+1} - r_i < \delta$. Such a partition can easily be got. For example, you can, by taking smaller δ if necessary, assume your δ is rational. Take now multiples of $\delta/2$ and stop when you go out of $[0, 1]$ taking the last point to be one.

Now choose N such that

$$n, m \geq N \Rightarrow (\forall i \; 1 \leq i \leq k) \; |y_n(r_i) - y_m(r_i)| < \epsilon/3.$$

This is possible because each of the sequences $(y_n(r_i) : n \geq 1)$ is convergent and hence Cauchy. We claim this N will do.

To see this, fix any $n, m > N$ and any $t \in [0, 1]$. we need to show that $|y_n(t) - y_m(t)| < \epsilon$. Let us say, $r_i \leq t \leq r_{i+1}$

$$|y_n(t) - y_m(t)| \leq |y_n(t) - y_n(r_i)| + |y_n(r_i) - y_m(r_i)| + |y_m(r_i) - y_m(t)|$$

$$\leq 3 \times \epsilon/3 = \epsilon.$$

The first and last terms are smaller than $\epsilon/3$ by choice of δ and the middle term by choice of N.

This proves (B)

Let us now prove (A). This is routine 'diagonal argument' we have done once long ago. Let us do it again. Again to avoid ugly notation let us reformulate our problem as follows.

(AA) For each $i = 1, 2, 3, \cdots$ we have a bounded sequence of reals, $\{x^i(n) : n \geq 1\}$. The claim is that there is one subsequence of integers $\{n_1 < n_2 < n_3 < \cdots\}$ so that for each i the subsequence $\{x^i(n_1), x^i(n_2), x^i(n_3), \cdots\}$ converges.

How does this show (A)? We are given a sequence of functions (x_n) from K. So for each rational r we can evaluate our functions at this point to obtain a sequence of numbers $\{x_1(r), x_2(r), x_3(r), \cdots\}$. This sequence is bounded by condition (ii) of the theorem. Thus we have countably many sequences as in the above para. the only difference is that in the earlier para I have one sequence for each i, here we have one sequence for each r. But it does not matter because the set of rationals is countable and we can enumerate it as one sequence. Thus (A) is a consequence of (AA)

We shall now proceed to prove (AA). The idea is simple but execution needs vocabulary. Here is the idea. List your sequences as follows.

$$\begin{gathered} x^1(1),\ x^1(2),\ x^1(3),\ x^1(4),\ \cdots \\ x^2(1),\ x^2(2),\ x^2(3),\ x^2(4),\ \cdots \\ x^3(1),\ x^3(2),\ x^3(3),\ x^3(4),\ \cdots \\ \vdots \end{gathered}$$

Since the first row is bounded, select a subsequence which converges. In other words, put cross marks on some terms so that if you read the sequence along the cross marks then it converges. Now read the second row only along the cross marks. Since it is bounded, there is a subsequence which converges. In other words if you make some cross marks into double crosses, then the second row read along the double crosses converges. Read the third row only along the double crosses and make some of these into triple crosses so that when you read third row along the triple crosses it converges. Continue. The required subsequence is the first crossed place, second double crossed place, third triple crossed place etc. This does. Too much english!

Here is the execution. Choose

$$n_{11} < n_{12} < n_{13} < \cdots \tag{1}$$

so that the sequence

$$x^1(n_{11}), \quad x^1(n_{12}), \quad x^1(n_{13}), \quad \cdots$$

converges. Possible because we started with a bounded sequence. Now consider the sequence of numbers

$$x^2(n_{11}), \quad x^2(n_{12}), \quad x^2(n_{13}), \quad \cdots$$

and take a subsequence that converges. That is,

$$n_{21} < n_{22} < n_{23} < \cdots \tag{2}$$

from among (1) so that

$$x^2(n_{21}), \quad x^2(n_{22}), \quad x^2(n_{23}), \quad \cdots$$

converges. Now consider

$$x^3(n_{21}), \quad x^3(n_{22}), \quad x^3(n_{23}), \quad \cdots$$

and take a subsequence that converges. That is

$$n_{31} < n_{32} < n_{33} < \cdots \tag{3}$$

from among (2) so that

$$x^3(n_{31}), \quad x^3(n_{32}), \quad x^3(n_{33}), \quad \cdots.$$

In general choose

$$n_{k1} < n_{k2} < n_{k3} < \cdots \tag{k}$$

from among

$$n_{k-1,1} < n_{k-1,2} < n_{k-1,3} < \cdots \tag{$k-1$}$$

so that

$$x^k(n_{k1}), \quad x^k(n_{k2}), \quad x^k(n_{k3}), \cdots$$

converges. Put

$$n_1 = n_{11}, \quad n_2 = n_{22}, \quad n_3 = n_{33}, \quad \cdots \tag{$*$}.$$

We claim that for each i,

$$x^i(n_1), \quad x^i(n_2), \quad x^i(n_3), \quad \cdots$$

converges. We first argue that $n_k > n_{k-1}$. This is because the k-th selection is a subset of $(k-1)$-th selection. In other words, at the worst we might have

$$n_{k1} = n_{k-1,1}, \quad n_{k2} = n_{k-1,2}, \quad \cdots n_{k,k-1} = n_{k-1,k-1}$$

showing $n_k = n_{kk} > n_{k-1}$. Thus we have a strictly increasing sequence of numbers (n_k).

Since (*) is a subsequence of (1) we see that the first sequence converges along this subsequence. Since (*) is a subsequence of (2) except possibly for the first term, we see second sequence converges along the subsequence.

Since (*) is a subsequence of (3) except possibly for the first two terms we see that the third sequence converges along this subsequence. In general (*) is a subsequence of (k) except possibly for the first ($k-1$) terms so that the k-th sequence converges along this subsequence.

This completes proof of (AA) and thus of (A) and thus of the theorem

There are just two points that need too be mentioned.

Firstly, I used a phrase like 'the first sequence converges along this subsequence'. we have never precisely defined what it means. I hope it is clear, but here it is. suppose we have a sequence $a = (a_n)$ of real numbers and a strictly increasing sequence of integers $S = (m_i)$. We say that the sequence a converges along the subsequence S, if the sequence $\{a_{m_1}, \ \ a_{m_2}, \ \ a_{m_3}, \ \ \cdots\}$ converges.

Second point is the following. we did some construction by induction. I cautioned once that when ever you do such a construction, you should precisely write down what you are going to do and then carry out. Carry out first step and assuming that you constructed up to $(k-1)$-th step, explain how you would do next step. do not say 'like this' 'like that' 'so on' etc. I may also add that subtle point is that when you *before hand* list properties, they should be interpretable inductively. if you are not clear you should refer too an earlier notes where this was discussed.

I carried out an inductive construction without listing before hand what I am proposing to do. That was done in order not to interrupt the thought process set in motion by the motivation. Here is the claim. We construct for each k a strictly increasing sequence

$$S_k = \{n_{k1}, \quad n_{k2}, \quad n_{k3}, \quad \cdots\}$$

of natural numbers so that (i) elements of S_k are among S_{k-1} and (ii) the k-th sequence (x^k) converges along S_k.

Vocabulary:

I withheld some technical names so that you can familiarize and appreciate the concept and meaning before the term is introduced. But it is time to learn the terms so that you can follow literature (and be able to communicate with others).

Recall that small sets are nowhere dense sets. A set A is nowhere dense if interior of its closure is empty, that is, the only open set contained in $\overline{A}$ is the empty set, that is, every non-empty open set contains points not in $\overline{A}$. Countable union of nowhere dense sets is called a set of the first category. A set which is not of first category is called set of second category. A set whose complement is of first category is called a *residual set*.

For example in the real line R the set Q of all rationals is of first category; the set A of all irrational numbers is a residual set; the interval $(0,1)$ is a set of second category. The interval $(0,1)$ is not residual because its complement is not of first category.

Thus residual sets are not only large sets, their complement is first category. On the other hand a set and its complement may both be second category. Of course, if you have a complete space then a set and its complement can not both be of first category because then whole space will be of first category, violating Baire.

In discussing compactness we came across the concept: Given any $\epsilon > 0$, finitely many balls of radius ϵ cover the space. When a space has this property, it is called *totally bounded*.

Recall that a set is bounded if there is a number M such that $d(x,y) \leq M$ for all $x, y \in M$. equivalently, there is an M such that $d(x_0, y) \leq M$ for all $y \in M$ where x_0 is some fixed point. if this last statement holds for one x_0 then it holds for all points x_0.

A totally bounded set is bounded. In fact take $\epsilon = 1$ and say k balls of radius one with centres at $\{x_i : 1 \leq i \leq k\}$ cover the set then for any points

x, y

$$d(x, y) \leq \max\{d(x_i, x_j) : 1 \leq i, j \leq k\} + 2.$$

In fact, take two points x, y, say $x \in B(x_i, 1)$ and $y \in B(x_j, 1)$, then

$$d(x, y) \leq d(x, x_i) + d(x_i, x_j) + d(x_j, y) \leq 2 + d(x_i, x_j).$$

You should not mistakenly think that the distances are bounded by some number like $2k + 2$ or some thing.

For example if you take $X = \{0, 1000\}$ with distance $|x - y|$ then two balls of radius one cover. But the diameter is far far large.

However a bounded set need not be totally bounded. We knew that even on the real line the metric $\min\{|x - y|, 1\}$ is bounded but clearly the space is not totally bounded; for no $\epsilon < 1$ you can cover by finitely many balls of that radius.

Thus we can state criterion for compactness as: completeness plus totally bounded.

In a metric space X a set A is said to be an ϵ-*net* if every point of the space within ϵ of some point of A; that is,

$$x \in X \Rightarrow (\exists\, p \in A) \;\; d(x, p) < \epsilon.$$

Equivalently, $B(A, \epsilon) = X$.

For example in R the set of integers is an 1-net, also 3/4-net, but not a 1/2-net. An ϵ-net need not be finite. Of course, in a totally bounded set, you can get finite ϵ-net for every $\epsilon > 0$. In fact, this can be taken as the definition of totally bounded.

We considered $A \subset C[0, 1]$ in Arzela-Ascoli theorem. It has the property that there is one bound for all functions in A. We then say that the set A is uniformly bounded or equi-bounded. Thus a subset A is *equi-bounded or uniformly bounded* if there is a number M so that $|f(x)| \leq M$ for all $f \in A$ and all x.

Similarly we had the property that given $\epsilon > 0$ there is one $\delta > 0$ such that $|f(s) - f(t)| < \epsilon$ for each $f \in A$ and each s, t with $|s - t| < \delta$ This property is called *equi-uniformly continuous* (or uniformly uniformly continuous, but this sounds awkward).

Thus we can state AA theorem as; A subset is compact iff it is closed, equi-bounded and equi-uniformly continuous.

We said that if a set satisfies conditions (ii) and (iii) then it may not be compact but its closure is compact. Such sets are called *pre-compact*. More precisely a set $A \subset X$ is called pre-compact if its closure is compact.

Thus a subset of $C[0,1]$ is pre-compact iff it is equi-bounded and equi-uniformly continuous. A subset of R is pre-compact iff it is bounded (in the usual metric).

Arzela-Ascoli:

No new idea is needed to see that the AA theorem holds for any compact metric space. This is what we explain now.

Let (X, d) be a compact metric space. Thus every real valued continuous function on X is bounded. Let $C(X)$ be the space of real continuous functions on X with sup metric

$$d(f, g) = \sup_x |f(x) - g(x)|.$$

Firstly this is a metric on the space. The space is complete. Proof is exactly as for $C[0,1]$. Take $\{f_n\}$ Cauchy (remember sup metric); then for each x, the sequence of numbers $\{f_n(x)\}$ is Cauchy; say $f(x)$ is its limit; argue f_n converges to f uniformly; conclude f is continuous and $d(f_n, f) \to 0$.

Arzela-Ascoli Theorem: A subset $K \subset C(X)$ is compact iff the following three conditions hold.

(i) K is closed

(ii) There is M such that $|f(x)| \leq M$ for all $f \in K$, all $x \in X$.

(iii) Given $\epsilon > 0$ there is $\delta > 0$ such that $|f(x) - f(y)| < \epsilon$ for $f \in K$ and $x, y \in X$ with $d(x, y) < \delta$.

Proof is as earlier. suppose K is compact. Since compact sets are closed (i) holds. Since $\Psi(f) = d(f, 0)$ is a continuous function on K it is bounded giving (ii). Given $\epsilon > 0$ choose finitely many balls of radius $\epsilon/3$ covering K; using uniform continuity take $\epsilon/3$ and get δ for these finitely many centers and argue that the same δ works for all $f \in K$; showing (iii).

Let now K satisfy the conditions. To show it is compact take a sequence $\{f_n\}$ in K. Need to exhibit a subsequence converging to a point of K. Remember that a compact space is separable; take a countable dense set D in X; use hypothesis (ii) and diagonal argument to get a subsequence $\{f_{n_k}\}$ such that for every point $p \in D$, the sequence of numbers $\{f_{n_k}(p)\}$ converges; use condition (iii) to argue $\{f_{n_k}\}$ converges, use condition (i) to say that this limit is in K — showing K is compact.

again as earlier if a set satisfies conditions (ii) and (iii) then its closure also satisfies those conditions and hence the closure is compact.

Compact metric spaces arise often in practice. For example the space of infinite sequences of $+1$ and -1 is a very natural phase space for a spin system consisting of very very large number of particles. For each configuration you can associate energy or magnetization and so on, very useful functions. (Actually you do so on finite products and then take limits. but let us not bother).

Stone-Weierstrass:

We shall now discuss probably the most complicated theorem of our course. This generalizes the Weierstrass theorem, you learnt last year, which says the following. Given any real valued continuous function f on the interval $[0,1]$, and $\epsilon > 0$, there is a polynomial P such that $|P(x) - f(x)| < \epsilon$ for each $x \in [0,1]$. With our present notation we can restate this as follows: given $f \in C[0,1]$ and $\epsilon > 0$, there is a polynomial P such that $d(f,P) < \epsilon$. Or simply put, the set of polynomials is dense in $C[0,1]$.

Given any polynomial P, and $\epsilon > 0$, there is a polynomial Q with rational coefficients such that $d(P,Q) < \epsilon$. Since the set of polynomials with rational coefficients is countable, we can also conclude that the space $C[0,1]$ is separable. Thus it is a separable complete metric space or a Polish space.

The theorem that we are going to discuss is a generalisation of the above. Let X be a compact metric space and $C(X)$ be the space of real valued continuous functions on X with sup metric.

Theorem:
Let $D \subset C(X)$ where X is a compact metric space. If D satisfies he

following three conditions then it is dense in $C(X)$.

(i) D is an algebra. That is, sum, product, constant multiple of functions in D is again in D.

(ii) Constant functions are in D.

(iii) D separates points. That is, given two points $p \neq q$ in X, there is an $f \in D$ with $f(p) \neq f(q)$.

Of course we could have stated (ii) to simply say that the constant function $x \equiv 1$ is in D. Then by (i) every constant function would also be in D.

Note that if condition (iii) fails then D can not be dense. This is obvious because if every function in D takes the same value at the two points p and q then so do their limits and hence it can not be dense. For example the function $f(x) = d(p, x)$ will not be in the closure of D.

That D is an algebra simply means that it is a vector space and closed under multiplication.

This theorem gives new information even in the case of $[0, 1]$. For example, if you take any integer $k \geq 1$ then the set of polynomials

$$a_0 + a_1 x^k + a_2 x^{2k} + a_3 x^{3k} + \cdots + a_n x^{nk}; \quad n \geq 1, a_i \in R$$

is dense in $C[0, 1]$. That is, polynomials in powers of x^k are dense. For example, you can consider powers of $x^{1000000}$ and their linear combinations.

Or you can take polynomials in $\sin x$. They are dense too. That is, the set of functions

$$a_0 + a_1 \sin x + a_2 (\sin x)^2 + a_3 (\sin x)^3 + \cdots + a_n (\sin x)^n : \quad n \geq 1; a_i \in R$$

is dense. Of course these you can obtain from a clever change of variable from usual Weierstrass theorem.

Proof of the theorem:

Let $\overline{D}$ be the closure of D. We understand the set enough to conclude that it must be all of $C(X)$.

1^o. The set $\overline{D}$ is again an algebra separating points and containing constants.

Since D already has the last two properties any larger set, in particular, closure has those properties. We only need to show that it is an algebra. But it is clear from the fact that if $f_n \to f$ and $g_n \to g$ then $f_n + g_n \to f + g$ and $cf_n \to cf$ and $f_n g_n \to fg$. remember all these are uniform convergences.

2^0. D and hence $\overline{D}$ strongly separates points: given $p \neq q \in X$ and numbers a, b there is $f \in D$ such that $f(p) = a$ and $f(q) = b$.

Indeed take φ which takes different values at p and q and define

$$f(x) = \frac{\varphi(x) - \varphi(q)}{\varphi(p) - \varphi(q)} a + \frac{\varphi(x) - \varphi(p)}{\varphi(q) - \varphi(p)} b$$

Even though φ appears at several places, only numbers calculated from it appear mostly. The constant function $\varphi(q)$ is in D and hence so is $\varphi - \varphi(q)$ and so is the ratio appearing in the first term above and so is the first term itself. Thus it is a function in D. So is the second term. so is their sum. It takes the required values at the given points.

Thus we have control on values we want the function to take at two given points. We can still get such a function from D.

3^o. $f \in \overline{D}$ implies $|f| \in \overline{D}$.

Indeed, take $\epsilon > 0$. Let the range of f be contained in the interval $[c, d]$. By usual Weierstrass theorem get a polynomial $P(t)$ such that for all t in this interval $||t| - P(t)| < \epsilon$. In other words the function $|t|$ is approximated by a polynomial on this interval. It is clear that $||f(x)| - P(f(x))| < \epsilon$, for all $x \in X$. In other words the distance between the two functions $|f|$ and $P(f)$ is at most ϵ. Remember that if p is the polynomial,

$$P(t) = a_0 + a_1 t + a_2 t^2 + \cdots + a_k t^k$$

then $P(f)$ is the function

$$a_0 + a_1 f + a_2 f^2 + \cdots + a_k f^k.$$

This is in $\overline{D}$ because f is in it and it is an algebra. Thus there are elements close to $|f|$ which are in $\overline{D}$. The later being closed, we conclude that $|f|$ itself is in it.

4^o.

$$f, g \in \overline{D} \Rightarrow f \vee g, f \wedge g \in \overline{D}.$$

Recall

$$f \vee g(x) = \max\{f(x), g(x)\}; \quad f \wedge g = \min\{f(x), g(x)\}.$$

The claim is immediate from the previous observation and the fact

$$f \vee g = \frac{f + g + |f - g|}{2}; \quad f \wedge g = \frac{f + g - |f - g|}{2}.$$

5^o. Fix $f \in C(X)$, fix a point $p \in X$ and fix a number $\epsilon > 0$. Then there is a function $h \in \overline{D}$ such that

$$h(p) = f(p); \quad h(x) < f(x) + \epsilon$$

Of course this function depends on p and should have been denoted by h_p but to lighten the notation we did not do that. Thus we can get a function in our set which is close to a given f. Well, not really close; but does not go too far above f; it may however go too far below f.

Proof is simple. For any $q \neq p$ fix a function $\varphi_q \in \overline{D}$ which takes same values as f at both p and q. This is possible by previous observation. Let

$$U_q = \{x : \varphi(x) < f(x) + \epsilon\}.$$

Then for each q we have an open set conning the point q and of course all these sets contain p as well thus they cover X. Pick finitely many, say

$$U_{q_1};\ U_{q_2};\ U_{q_3};\ \cdots U_{q_k}.$$

Consider the function

$$h = \varphi_{q_1}\ \wedge\ \varphi_{q_2}\ \wedge\ \cdots \varphi_{q_k}.$$

Since each $\varphi_q(p) = f(p)$ we see $h(p) = f(p)$. Given any point $x \in X$ it is in some U_{q_i} and so $\varphi_{q_i}(x) < f(x) + \epsilon$ and so $h(x) < f(x) + \epsilon$. Since each $\varphi_q \in \overline{D}$ a previous observation tells us that $h \in \overline{D}$.

6^o. Fix $f \in C(X)$ and $\epsilon > 0$ the there is a $g \in \overline{D}$ such that

$$f(x) - \epsilon < g(x) < f(x) + \epsilon; \quad \forall\, x \in X.$$

In other words $d(f, g) < \epsilon$. Since this is true for every $\epsilon > 0$ and since $\overline{D}$ is closed we conclude that $f \in \overline{D}$. Since $f \in C(X)$ is arbitrary we conclude that $\overline{D} = C(X)$. In other words D is dense in $C(X)$; proving the theorem.

This is simple. For each $p \in X$, use earlier observation to get h, now let us denote it by h_p so that

$$h_p \in \overline{D}; \quad h_p(p) = f(p); \quad h_p(x) < f(x) + \epsilon\ \ \forall\, x \in X.$$

Set

$$V_p = \{x : h_p(x) > f(x) - \epsilon.\}$$

Then $p \in V_p$. In other words the family of all V_p cover the space X and take a finite sub family that covers, say,

$$V_{p_1};\ V_{p_2};\ \cdots; V_{p_k}.$$

Set

$$g = h_{p_1} \vee h_{p_2} \vee \cdots \vee h_{p_k}.$$

Given any point x, it is in some V_{p_i} and so $h_{p_i}(x) > f(x) - \epsilon$ and hence $g(x) > f(x) - \epsilon$. Also given $x \in X$ each $h_p(x) < f(x) + \epsilon$ and hence so is the finite max $g(x)$. By earlier observation, we see $g \in \overline{D}$.

This completes the proof.

The 'type' of argument to prove a little of what is needed in 5^o and then improve it in 6^o is called boot-strapping.

We have already seen implications of this theorem for continuous functions on the interval $[0, 1]$. We can use this theorem to show that for a compact metric space X, the space $C(X)$ is a separable metric space. Thus it is a Polish space (complete separable metric space).

We can also show that trigonometric polynomials are dense in the space of continuous functions f on $[0, 2\pi]$ with $f(0) = f(2\pi)$.

Space Filling Curves or Peano Curves:

Cantor startled the community by showing that R^2 has same number of points as R. But of course, he could exhibit only some identification and it was unclear whether one could get a continuous function.

Peano startled the community by showing that there is a continuous function from the unit interval $[0, 1]$ onto the unit square $[0, 1] \times [0, 1]$. After all, a continuous function from the interval to the plane should 'trace a curve' which should be 'one dimensional'. But that mental picture was proved to be wrong. This curve of Peano fills up all of the square. That is why such functions from then on are called Peano curves or space filling curves.

Soon after words Hilbert gave another geometric construction which is more fertile, in the sense, it inspired others and also gave a general method of construction. It also helped to evaluate some interesting properties such maps can possess.

One knew properties such maps can not possess. For example such a map can not be one-one. If f is a continuous map of a compact metric space X onto another compact metric Y space which is one-to one; then its inverse is already continuous and hence such a map is a homeomorphism. In fact, if V is open in X, then V^c is closed and hence compact (because X is so) and hence so is $f(V^c)$ and hence it is closed in Y and hence $f(V) = [f(V^c)]^c$ is open. This shows that f^{-1} is continuous.

In our case $X = [0,1]$ and $Y = [0,1] \times [0,1]$ and we knew that they can not be homeomorphic. It was also known that such maps can not be differentiable.

At that time when they were found, these were curious objects. But now-a-days they are finding applications in computer science. After all, by talking about point t on the line you can mean point $f(t)$ in the square. In other words, you can stay on the line and talk about two dimensional things. This was found useful in data organisation. I do not know much of it.

However, there are some interesting algorithmic questions where these maps are found useful — for example, for the travelling salesman problem (known to be 'hard' in some sense). This problem has a data of finitely many points in the unit square and asks for a tour of these points which is economical; shortest possible. There is no idea how to make a beginning of a tour. If you have a Peano curve you can make a beginning (good or bad). Take those points given to you, take 'some' points $\{t_i\}$ such that your data is $\{f(t_i)\}$ and tour in the increasing order of the points t_i. In other words the problem is linearized, an order has been brought.

I shall describe the Hilbert method but do not carry out because it needs some vocabulary. Our interest is only to make it clear to ourselves that such curves exist. We are not interested in either studying their properties or in using them. So we follow a simpler method later and completely work out. But you should know Hilbert method.

Divide the interval $[0,1]$ into four parts

$$I_1 = [0, 1/4];\ I_2 = [1/4, 1/2];\ I_3 = [1/2, 3/4];\ I_4 = [3/4, 1].$$

Divide the square also into four parts

$$R_1 = [0, 1/2] \times [0, 1/2]; \quad R_2 = [0, 1/2] \times [1/2, 1]$$

$$R_3 = [1/2, 1] \times [1/2, 1]; \quad R_4 = [1/2, 1] \times [0, 1/2].$$

Draw a curve joining by straight lines, the mid point of R_1 to mid point of R_2 to R_3 to R_4. Set up a map f that takes I_j to the part of your curve in R_j. This is our first function f_1.

Now divide each I_j into I_{jk} for $k = 1, 2, 3, 4$, and similarly the squares. assign the four parts of the line I_j to the four parts of the square R_1 and so on and draw a curve made up of straight lines. This is our second function f_2. You need to be careful and give an algorithm on how to do it. This can be done.

To do this you need either picture or vocabulary. But the point is that these functions, so obtained, converge to a function which is continuous and its image is all of the square. We shall not pursue this process.

We exhibit our curve (due to Lebesgue) after two observations.

Remember the construction of the Cantor set. We define

$$I_0 = [0, 1/3]; \qquad I_2 = [2/3, 1]$$

$$I_{00} = [0, 1/3^2];\ I_{02} = [2/3^2, 3/3^2];\ I_{20} = 6/3^2, 7/3^2];\ I_{22} = [8/3^2, 9/3^2].$$

and so on.

$$C_0 = [0, 1];\ C_1 = I_0 \cup I_2;\ C_2 = \cup\{I_{ij}; i, j = 0, 2\}$$

$$C = \cap C_n$$

is the Cantor set.

Observation 1: There is a continuous function f defined on the Cantor set onto the unit interval $[0, 1]$.

This is simple. we know that a point x is in the Cantor set iff it has triadic expansion involving only the digits zero and two. That is,

$$x = \frac{x_1}{3} + \frac{x_2}{3^2} + \frac{x_3}{3^3} + \cdots; \quad x_i = 0, 2.$$

equivalently
$$x = \frac{2x_1}{3} + \frac{2x_2}{3^2} + \frac{2x_3}{3^3} + \cdots; \quad x_i = 0, 1.$$
Keep in mind that such an expansion is unique. It is worth remembering that the point $1/3$ is in C and it has the eligible ternary expansion $.0222222\cdots$. Also remember given any sequence of zeros and ones, the series $\sum (2x_i)/3^i$ converges and the number so defined is a point in $[0,1]$ and is in the Cantor set.

Here then is the function
$$f(x) = \sum \frac{x_i}{2^i}; \qquad x = \sum \frac{2x_i}{3^i}.$$
This is good definition because, given a point $x \in C$ it has a unique eligible ternary expansion (that is involving zeros and twos).

Clearly given any number $y \in [0,1]$, it has a binary expansion $\sum y_i/2^i$ with each y_i zero or one and then the point $x = \sum (2y_i)/3^i$ is in C and $f(x) = y$ to show that f is onto $[0,1]$.

Finally continuity of f is seen as follows. let $x_n \to x$, all in C. If $x \in I_0$ then first digits of the x_n after some stage must be zero. Otherwise $x \in I_2$ and after some stage the first digit of the x_n must be two.

Similarly depending to which of the four intervals I_{ij} the point x belongs, we see that the second digit of x_n must be same as that of x after some stage. In other words, given any k, $f(x_n)$ will have the same first k digits as that of $f(x)$ after some stage. This shows continuity.

Observation 2: there is a continuous one-one map g on C onto $C \times C$.

Of course any such map is a homeomorphism, as noted earlier. We are thinking $C \times C \subset [0,1] \times [0,1]$. If $x \in C$ and
$$x = \frac{x_1}{3} + \frac{x_2}{3^2} + \frac{x_3}{3^3} + \cdots; \quad x_i = 0, 2.$$
we put
$$g_1(x) = \frac{x_1}{3} + \frac{x_3}{3^2} + \frac{x_5}{3^3} \cdots; \qquad g_2(x) = \frac{x_2}{3} + \frac{x_4}{3^2} + \frac{x_6}{3^3} + \cdots;$$
$$g(x) = (g_1(x), g_2(x)).$$

Since each x_i is zero or two, we see that both $g_1(x)$ and $g_2(x)$ are in C. We have just separated the even and odd digits. This is one-one. Remember an eligible expansion is unique for points in C. This is also onto $C \times C$ because given two points in C we can interlace the digits of these two points to get a point of C whose image is the given pair.

That it is continuous is immediate because the digits are unchanged. (of course, this is to be interpreted carefully, there is a change in the 'place' of the digit).

Observation 3: Given a continuous function h on C to $S = [0,1] \times [0,1]$ we can extend it to a continuous map of $I = [0,1]$ to S.

The only issue is to define the function on the points of the deleted intervals. Let $x \in (a,b)$ a deleted interval. Thus $a, b \in C$. Hence we know $h(a)$ and $h(b)$. Extend linearly in between. More precisely put

$$h((1-\theta)a + \theta b) = (1-\theta)f(a) + \theta f(b); \qquad 0 \leq \theta \leq 1$$

Since unit square is convex, h so defined still takes values in S. As the point θ goes from 0 to 1; the point $t = (1-\theta)a + \theta b$ travels from a to b and the h values travel on the straight line from $h(a)$ to $h(b)$. We are stll using the same symbol h for this extended function. Thus now h is defined on all of I and takes values in S.

We need to show that h so defined is continuous at all points. Of course, it is continuous at all points outside the Cantor set. Indeed, let $x \in I_n$, one of the deleted open intervals and $x_i \to x$. Then after some stage each $x_i \in I_n$ and so the definition given above takes over to show continuity.

Let $x \in C$. Let $x_i \to x$. We need to show that $h(x_i) \to h(x)$. Given $\epsilon > 0$ plot (a', b') around x so that for every point $y \in C \cap (a', b')$ we have $d(f(x), f(y)) < \epsilon$. This is possible because f is given to be continuous on C. Let us take $a, b \in C$, $a' < a < x < b < b'$. This is possible because every point of C is a limit point of C.

We now claim that for every point y in this interval (a,b), whether y is in C or not, we have the inequality $d(h(x), h(y)) < \epsilon$. Since this last inequality holds for points of $y \in C$ in this interval, we only need to show that it holds for points not in C. But then any such point is in a deleted interval (a_n, b_n).

Since $a, b \in C$ we conclude that

$$d(h(a_n), h(x)) \le \epsilon; \quad d(h(b_n), h(x)) \le \epsilon.$$

In other words, both $h(a_n)$ and $h(b_n)$ are in the ϵ-ball around $h(x)$. Remember balls are convex. Since $h(y)$ is on the line joining $h(a_n)$ and $h(b_n)$, we see that $d(h(y), h(x)) \le \epsilon$. This shows continuity.

We shall now exhibit space filling curve.

Take $f : C \to [0, 1]$ as in observation (1).
Take $g(x) = (g_1(x), g_2(x)) : C \to C \times C$ as in observation (2).
Define $h : C \to S$ by

$$h(x) = (h_1(x), h_2(x)); \quad h_1(x) = f(g_1(x));\ h_2(x) = f(g_2(x)).$$

The map h is a continuous function on C — it is composition of continuous functions. Given $(a, b) \in S$; pick $s, t \in C$ with $f(s) = a$ and $f(t) = b$ and pick $x \in C$ with $g(x) = (s, t)$. Then verify $h(x) = (a, b)$. This shows that h is onto S. Now use observation (3) to extend the map as a continuous map from all of I to S. Of course, it still remains onto S.
This does.

Cantor sets:

We discussed once that theoretically one could construct similar to the Cantor set, by cutting into more number of pieces. For example, at each stage you divide an interval into five subintervals of equal length; instead of three. You pick the first, third and fifth; instead of the first and last. The procedure remains same. Such sets are also called Cantor sets. Here is the precise definition.

A compact subset of R is called *perfect* if every point is a limit point of the set. For example the usual Cantor set, or the interval $[0, 1]$ are such sets. The set

$$\{0,\ 1,\ 1/2,\ 1/3,\ \cdots\}$$

is closed but not perfect.

The set Q of all rationals is not closed, but every point of it is a limit point of the set.

A perfect nowhere dense set is called *a Cantor set*. This means our set should be

closed,

bounded,

every point must be a limit point,

should not contain any (non-empty) open interval.

Thus usual Cantor set is an example of a Cantor set, it is called *the Cantor set*. The construction described above will also lead to a Cantor set. Mathematics, like music, can be improvised with your ideas. For example should we cut the interval into fixed number of pieces? First step cut into three pieces and leave even one, that is, the middle one. You select the other two intervals. Next, cut each selected part into five pieces and leave even ones, that is, the second and fourth part in each of the first level intervals. You have six intervals. Next cut each selected interval into seven parts etc.

It is all your pick! If you want you can divide into p_n subintervals at the n-th stage. Here p_n is the $(n+1)$-th prime.

You would probably think that you are getting many many very different sets. But this is only illusion as far as sets are concerned.

Here is a very interesting theorem: Such sets are all the same! More precisely, let P and Q be two such sets. Then there is a homeomorphism h of R to itself which sends P to Q. This last phrase means $h(P) = Q$. This is proved exactly like a similar theorem we proved quite a while ago, namely, given two countable dense sets in R, there is a homeo that sends one to the other. We shall not carry out the proof.

You should not come to the erroneous conclusion that all such sets are same. No. We only said that they *look same*. But different sets have *different qualities and expertise* from analysis/arithmetic point of view. We shall not pursue this line of thought. We shall leave aside our long excursion to metric spaces and return to real line once again.

what next:

We completed our discussion of metric spaces. We have done some basic results and applications. There are two possibilities for the remaining period — we can discuss either power series with complex coefficients or Fourier series.

We discussed real power series $\sum a_n x^n$ and discussed radius or interval of convergence in our first course. we found a number $R \geq 0$ such that the power series converges (actually, converges absolutely) for every real number x with $|x| < R$; does not converge for any real number x with $|x| > R$; and for $x = \pm R$, it may or may not converge depending on the particular series.

The same analysis can be done even if you have power series $\sum a_n z^n$ with complex coefficients. There is a real number $R \geq 0$ such that whenever you take a complex number z with $|z| < R$ the series converges (converges absolutely). Whenever you take a complex number z with $z| > R$ the series does not converge. Exactly the same proof that we did for real series goes through.

However we shall not continue this line of thought. Mainly because, introducing the theory with the statement that the same proof as the real case goes through, is perhaps the worst way of introducing the splendour of complex analysis. You can tell your complex analysis teacher to carry out the details concerning power series.

Of course, we did discuss complex valued functions of complex variable and their derivatives. This was when we were discussing functions from R^2 to R^2. We derived the Cauchy-Riemann equations; some of its interesting consequences; related the complex derivative to the derivative matrix when you regard the function as a map of R^2 to R^2 instead of from C to C. We shall discuss some relevant aspects of complex analysis as we go along.

We shall now discuss Fourier series. We would first review what we know about C^n, the complex n dimensional space.

complex n-dimensional vector space:

You studied the n-dimensional complex vector space and inner product. Let us recall. C^n is the set of all n tuples $z = (z_1, z_2, \cdots, z_n)$ of complex numbers. We define inner product

$$\langle z, w\rangle = \sum_k z_k \overline{w_k}; \quad \langle z, z\rangle = ||z||^2 = \sum |z_k|^2.$$

$$d(z, w) = \sqrt{\langle z - w, z - w\rangle} = \sqrt{\sum |z_k - w_k|^2}.$$

Here $\overline{w}$ is the complex conjugate of w. Recall if $w = a + ib$ where a, b are real then its conjugate is $a - ib$.

This space has an orthonormal basis consisting of n vectors. Thus there are n vectors $e_1, e_2, \cdots, e_n$ such that

$$\langle e_k, e_l\rangle = 1 \text{ if } k = l; \quad \text{and} \quad = 0 \text{ if } k \neq l.$$

Any vector v can be uniquely written as

$$v = \sum_1^n \langle v, e_k\rangle e_k = \sum \hat{v}_k e_k.$$

It follows from properties of inner product that

$$||v||^2 = \sum_1^n |\hat{v}_k|^2.$$

interesting point, trivial but should be noted, is that this is true whatever orthonormal basis you take Let us take an integer $1 \leq m < n$ and consider only the partial sum

$$v^* = \sum_1^m \hat{v}_k e_k.$$

Then clearly v^* is in the subspace spanned by the first m-basis vectors, that is,

$$v^* \in \text{ span } \{e_1, e_2, \cdots, e_m\} = S; \quad \text{say.}$$

The vector $v - v^*$ is orthogonal to every vector in S. This is because

$$\langle v - v^*; \sum_1^m a_k e_k\rangle = \langle v; \sum_1^m a_k e_k\rangle - \langle v^*; \sum_1^m a_k e_k\rangle$$

$$= \sum_1^m \overline{a_k}\langle v, e_k\rangle - \sum_1^m \overline{a_k}\langle v^*, e_k\rangle = 0;$$

because, $\langle \sum_{l=1}^{m} \hat{v}_l e_l, e_k \rangle = \hat{v}_k = \langle v, e_k \rangle$.

Since $v = (v - v^*) + v^*$ and $v - v^* \perp v^*$ we see

$$||v||^2 = ||v - v^*||^2 + ||v^*||^2.$$

Thus $||v^*||^2 \leq ||v||^2$. We could have got it by direct computation because

$$||v^*||^2 = \sum_1^m |\hat{v}_k|^2 \leq \sum_1^n |\hat{v}_k|^2 = ||v||^2.$$

There is a subtle point in this so called 'direct computation'. It is that, we assume that v itself is a linear combination of the basis vectors. You might think, what else can it be; it has to be a linear combination; every vector is and so on. However you will realize the subtlety later.

In a sense v^* captures the entire part of v in S (whatever it may mean) so that what remains is orthogonal to S. Not only this, v^* is the vector closest to v in the subspace S. That is

$$d(v, v^*) \leq d(v, w) \quad \forall \;\, w \in S.$$

This is seen as follows.

$$v - w = (v - v^*) + (v^* - w)$$

and the two terms on the right are orthogonal we see

$$||v - w||^2 = ||v - v^*||^2 + ||v^* - w||^2 \geq ||v - v^*|^2$$

Fourier series:

The plan is to understand the statement of Fourier:

Every wave is a superposition of sine and cosine waves.

Let us consider the interval $[0, 1]$. Let f be a function on this space with $f(0) = f(1)$. We can extend it as a periodic function of period one on the real line, in a unique way, namely, define $f(x+1) = f(x)$ for all $x \in R$. More precisely it is defined by $f^*(x)$ is the given $f(x)$ on $[0, 1)$; $f^*(x) = f(x-1)$ on $[1, 2]$ and $f^*(x) = f(x-2)$ on $[2, 3]$ etc. Similarly on the negative side. It is much easier for you to think of its graph. First imagine on $[0, 1]$ and then extend the curve to all of R.

Conversely given any function f^* on R which is periodic of period one its restriction to $[0, 1]$ gives us a function f on $[0, 1]$ with the property $f(0) = f(1)$. Moreover in this process, f is continuous iff f^* is continuous.

Such functions on R have a wavy graph they are called waves. Actually you can define what are waves using certain differential equations but let us not go too far now. The simplest waves we know from high school are $s(x) = \sin(2\pi x)$ and $c(x) = \cos(2\pi x)$. Of course any multiple of these functions is also a wave. The functions $\sin(4\pi x)$, $\cos(4\pi x)$ are also such functions. More generally $\sin(2k\pi x)$ and $(\cos(2k\pi x)$ are such functions.

By superposition we mean sum. Thus the function $f + g$ is superposition of f and g. Generally, you do not use this term. This originates in music, where a sound is superposition of some basic sounds; you may call 'notes' and so on. This phrase originates again in differential equations where if you have two solutions f and g of an equation like $f'' + f = 0$, then their sum is also a solution — just as if you have two vectors which are solutions of a homogeneous matrix equation $Av = 0$, then so is their sum. For the differential equation written above you see 'basic' solutions are $\sin x$ and $\cos x$ and other solutions are just linear combinations of these two. This is irrelevant for us now. I am only trying to explain the word 'superposition'.

Thus the statement of Fourier amounts to saying that any periodic function of period one is a sum of linear combinations of the functions: $\sin 2\pi kx$ and $\cos 2\pi kx$. Thus mathematically, the statement amounts to saying that given periodic function of period one, there are numbers (a_n) and (b_n) such that

$$f(x) = a_0 + (a_1 \sin 2\pi x + b_1 \cos 2\pi x) + (a_2 \sin 4\pi x + b_2 \cos 4\pi x) + \cdots$$
$$+(a_k \sin 2k\pi x + b_k \cos 2k\pi x) + \cdots.$$

Such a series is called Fourier series.

I am carrying the baggage $2\pi k$ instead of k because I am wanting period one. If you want period 2π then you just look at $\sin kx$ and $\cos kx$. It is a matter of standardization. If you want to take the interval $[0, 37]$ then we consider the functions $\sin(2\pi kx/37)$ and $\cos(2\pi kx/37)$. Of course $k = 0, 1, 2, 3, \cdots$..

There are several problems. What is the meaning of the series? How should we understand it? In what sense is the equality to be understood? Is such a thing true? If not every function; which functions are like this? So on.

Prior to Fourier problems concerning series of sines and cosines did arise in understanding waves and sounds. Fourier was studying, in a systematic manner, conduction of heat and arrived at this problem. You should recall that last semester we did discuss heat equation, found out that the 'normal density'

$$p(x,t) = \frac{1}{\sqrt{2\pi t}} e^{-x^2/2t}$$

is fundamental solution, in a precise sense. If you remember, the problem was to describe at every time $t > 0$, the distribution of heat in an (two sided) infinite rod when you know the initial distribution of heat on the rod. The problems for a finite rod are more difficult.

Before making systematic study of the above type of series, let me add most of mathematics originated here. In fact understanding certain problems in Fourier series was the genesis of Set Theory by Cantor. Understanding these issues are behind even in matters like convergence of series and so on. Of course, series were used earlier too.

Exponentials:

The whole theory (at least for us now) can be regarded as an extension of what we learned about C^n earlier. Let us start considering $C[0,1]$. For the time being forget about periodic etc. So it is not necessary to have $f(0) = f(1)$ at this moment —later when theorems appear, we need. But the change of attitude is that *we consider complex valued functions.*

Let us quickly recall. $f : [0,1] \to C$ then for each x, $f(x)$ is a complex number and hence is $f_1(x) + if_2(x)$ where $f_1(x)$ and $f_2(x)$ are real numbers. Thus given a complex valued function, there are two real valued functions f_1 and f_2 so that $f = f_1 + if_2$. The function f_1 is called the real part of f and f_2 is called the imaginary part of f. You must note that these functions are real valued.

Conversely, given two real valued functions, the above equality provides you a complex function. Unless stated to the contrary, when we write $f = f_1 + if_2$ we mean that these are the real and imaginary parts. Also, bringing in the notion of convergence in C, you see that f is continuous iff its real part and imaginary part are continuous.

For later use, let us also make an observation. Suppose you have a sequence ($f^n = f_1^n + if_2^n$) of complex valued functions and a function $f = f_1 + if_2$. Then f^n converges to f uniformly iff the real parts and imaginary parts converge uniformly. This follows from a simple observation. For a complex number $z = a + ib$

$$\max\{|a|, |b|\} \quad \leq \quad |z| \quad \leq \quad |a| + |b|.$$

Thus if $f^n \to f$ uniformly, then $||f_1^n - f_1|| \leq ||f^n - f||$ and hence $f_1^n \to f^1$ uniformly. Similarly $f_2^n \to f_2$ uniformly.

conversely, if the real and imaginary parts converge uniformly, then

$$||f^n - f|| \leq ||f_1^n - f_1|| + ||f_2^n - f_2||$$

showing that $f^n \to f$ uniformly.

We define for any real number θ,

$$e^{i\theta} = \cos\theta + i\sin\theta.$$

Then by induction and using the sine and cosine formulae, we can show

$$e^{in\theta} = \cos n\theta + i\sin n\theta.$$

This is called De Moivre's formula and you must have seen it in high school.

Actually, one defines for every complex number z, the series

$$\exp\{z\} = 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \cdots + \frac{z^n}{n!} + \cdots$$

This series converges. Indeed take the disc $\{|z| \leq M\}$. Let $S_n(z)$ be the n-th partial sum of the above series. Then the sequence $\{S_n(z)\}$ is uniformly Cauchy in the disc because for $n \geq m$

$$|S_n(z) - S_m(z)| \leq \sum_{m+1}^{n} \frac{M^k}{k!} \to 0 \text{ as } m, n \to \infty.$$

If you take θ to be real number, then the above definition along with the fact that $i^2 = -1$ tells us

$$\exp\{i\theta\} = 1 + i\theta + \frac{(i\theta)^2}{2!} + \frac{(i\theta)^3}{3!} + \cdots + \frac{(i\theta)^n}{n!} + \cdots$$

$$= 1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} + \cdots + (-1)^n \frac{\theta^{2n}}{(2n)!} + \cdots$$

$$+i \left\{ \theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} + \cdots + (-1)^n \frac{\theta^{2n+1}}{(2n+1)!} + \cdots \right\}$$

$$= \cos\theta + i \sin\theta$$

where the last identification is from the expressions we obtained in our first course for the sine and cosine functions. Thus the definition we used above is same as the natural one defined just now. Further, if z has zero imaginary part, that is, z is real, then this definition coincides with what we learnt earlier.

One can prove Cauchy theorem on products of series and show, exactly as in the real case, that

$$\exp\{z + w\} = \exp\{z\} \exp\{w\}$$

justifying the use of the notation e^z for $\exp\{z\}$. This gives another proof of the De Moivre formula. But we need not depend on the unproved Cauchy rule.

Recall that integration is done by using the real and imaginary parts. Thus for $k \neq 0$,

$$\int_0^1 e^{2\pi ikt} dt = \int_0^1 \cos(2\pi kt) dt + i \int_0^1 \sin(2\pi kt) dt = 0.$$

$C[0,1]$ **as** C^n **for Huge** n**:**

Let us from now consider the space of complex valued continuous functions on $[0,1]$. This will be $C[0,1]$ during the remaining part of our discussion. This will replace C^n, as you will see.

We define some special functions on $[0,1]$ as follows.

$$e_n(t) \quad = \quad e^{2\pi int}, \quad n = 0, \pm 1, \pm 2 \cdots.$$

The calculation of integral above can be recast in terms of these functions.

$$\int_0^1 e_n(t)\overline{e_m(t)} dt = 0; \quad if \ \ n \neq m; \quad = 1 \ \ if \ \ n = m.$$

We defined on C^n inner product

$$\langle z, w \rangle = \sum z_k \overline{w_k}; \quad z = (z_1, \cdots, z_n); \ \ w = (w_1, \cdots, w_n).$$

Analogously let us define an inner product on the vector spar $C[0,1]$ (remember complex valued continuous functions)

$$\langle f, g\rangle = \int_0^1 f(t)\overline{g(t)}dt$$

This is indeed an inner product:

linear in the first argument f,
conjugate linear in the second argument g;
$\langle f, f\rangle = \int |f|^2 \geq 0$
$\langle f, f\rangle = 0$ iff $f = 0$ (remember continuity of f)
$\langle f, g\rangle = \overline{\langle g, f\rangle}$ because $\int \overline{\varphi} = \overline{\int \varphi}$.

Moreover the integrals of exponentials we evaluated above, tells us that the family $\{e_n : -\infty < n < \infty\}$ is an orthonormal system.

The only thing is that in C^n we have a finite orthonormal basis. If you found n orthonormal vectors, then there are no more to extend this system. If you found only m orthonormal vectors and $m < n$ you can find more. You should actually find more because you can not express every vector as linear combination of just these m vectors.

In the present vector space we have found an infinite system of orthonormal vectors. It is not clear if there are any more that can be added to this list. Just as we expanded an vector in C^n in terms of an orthonormal system, we can try to expand a function in terms of this orthonormal system. Hopefully

$$f = \sum_{-\infty}^{\infty} c_k e_k.$$

Back to Fourier Series:

The question then is whether such an expansion is possible and what has it got to do with Fourier series. Well, actually, this is nothing but Fourier series. We started out with series of sines and cosines. We now have exponentials. But obviously $\{\sin nt; \cos nt\}$ can be expressed as linear combinations of $\{\exp(int); \exp(-int)\}$ and vice versa. Thus linear combination of exponentials and linear combination of sines and cosines are exactly the same.

So let us now turn our attention to the set up we have settled upon, namely the vector space $C[0,1]$ with inner product $\int f\overline{g}$ and the orthonormal basis $\{e_n : -\infty < n < \infty\}$.

Suppose we can express $f = \sum c_n e_n$; without even knowing the sense in which this infinite sum is to be interpreted. Then the first question is what should be the coefficients? Going by our understanding of the finite dimensional case we would guess

$$c_n = \langle f, e_n \rangle = \int_0^1 f(t) e^{-2\pi i n t} dt.$$

Notice that conjugate of e_n is e_{-n}.

Of course even if we did not know the finite dimensional case, as far as guess is concerned we would feel

$$\int f \overline{e_n} = \int (\sum c_k e_k) \overline{e_n} = \sum c_k \int e_k \overline{e_n} = c_n.$$

where we used the fact that the system (e_n) is orthonormal. Of course whether we can interchange the infinite sum with integral is unclear; any way the meaning of the infinite sum itself is unclear. This is only a thought process to guess what the coefficients should be.

Now we define Fourier series of a function $f \in C[0,1]$ to be the series

$$\sum \hat{f}_k e_k; \qquad \hat{f}_k = \int_0^1 f(x) e^{-2\pi i k x} dx.$$

The numbers $\hat{f}_k$ are called the Fourier coefficients of f; and more precisely this is the k-th Fourier coefficient.

Of course, yet there is no meaning for the infinite series. There are several ways of giving meaning. The right stage for this drama is the Lebesgue integral, but we shall not enter that stage. We shall continue to work with familiar continuous functions and familiar Riemann integral.

Towards giving a meaning to the series, let us define partial sums

$$S_N(x) = \sum_{-N}^{N} \hat{f}_k e^{2\pi i k x}; \quad N \geq 1.$$

Following usual procedure of giving a meaning to an infinite series, we ask if the above sequence of partial sums converges.

The fact that we have taken 'symmetric' partial sums will hurt you. But do not worry. We can deal other sums as well after the smoke clears. But in any case that should not be main issue.

Let us first think of uniform convergence. Whether point wise or uniform, one thing is clear. All e_n and all S_N are period one functions. So limit, when exists, is also of period one. Thus we are advised to restrict to period one functions — if we are aiming at uniform convergence.

Here is the first main theorem of the theory.

Theorem: Suppose f is a period one function with (period one) continuous derivative. Then S_N converges to f uniformly.

Let me add period one simply means $f(0) = f(1)$. Its derivative is period one means $f'(0) = f'(1)$. This is same as saying that when you extend f to all of R as period one functions then it is continuously differentiable on R. If you are restricting your attention to $[0, 1]$ then differentiability does not mean derivative is period one, even though original function is so. For example $f(x) = (x - 1/2)^2$ is a period one function, takes same values at zero and one. It is C^1 function on the interval $[0, 1]$. However $f'(0) = -1 \neq +1 = f'(1)$.

We start with some general observations, trying to imitate the finite dimensional case. Let $f \in C[0, 1]$ *not necessarily period one function.* not necessarily differentiable. Recall that now the space $C[0, 1]$ has metric d;

$$d^2(f, g) = \langle f - g, f - g\rangle = \int_0^1 |f - g|^2.$$

1^o. Does S_N capture 'all of f' in the span of $\{e_k : -N \leq k \leq N\}$? Yes.

Denote this subspace by L. Then we claim that $f - S_N \perp L$. to see this take any $h = \sum c_k e_k$ in L.

$$\langle f - S_N, \sum c_k e_k\rangle = \langle f, \sum c_k e_k\rangle - \langle \sum \hat{f}_k e_k, \sum c_k e_k\rangle$$

$$= \sum \overline{c_k}\langle f, e_k\rangle - \sum \hat{f}_k \overline{c_k} = 0$$

Remember we had exactly the same result with same proof in C^n.

2^o. S_N is the closest to f in the span of the vectors $\{e_k : -N \leq k \leq N\}$, this sub space is denoted by L above.

To see this take any $h = \sum c_k e_k$ in this subspace. Need to show

$$||f - S_N|| \leq ||f - h||.$$

To show this observe $f-h=(f-S_N)+(S_N-h)$ and $S_N-h \in L$ so that this is orthogonal to $f-S_N$ from 1^o. Thus pythogoras tells

$$||f-h||^2 = ||f-S_N||^2 + ||S_N-h||^2 \geq ||f-S_N||^2$$

with equality iff $S_N = h$.

In the above calculation, by pythogoras we mean, if $f \perp g$ then

$$\begin{aligned} ||f+g||^2 &= \langle f+g, f+g\rangle = \langle f,f\rangle + \langle f,g\rangle + \langle g,f\rangle + \langle g,g\rangle \\ &= ||f||^2 + ||g||^2 \end{aligned}$$

3^o.

$$\sum_{-N}^{N} |\hat{f}_k|^2 \leq ||f||^2 = \int_0^1 |f(x)|^2 dx.$$

This is known as *Bessel's inequality.*

This is immediate from the fact

$$f = (f-S_N) + S_N; \quad f - S_N \perp S_N$$

and Pythogoras to get

$$||f||^2 = ||f-S_N||^2 + ||S_N||^2; \quad ||S_N||^2 \leq ||f||^2$$

But

$$||S_N||^2 = \langle \sum \hat{f}_k e_k, \sum \hat{f}_k e_k\rangle = \sum \hat{f}_k \overline{\hat{f}_k} = \sum |\hat{f}_k|^2.$$

Same result was in C^n. Now you go back and see 'direct computation' i mentioned there and how it does not make sense here.

4^o. For any $f \in C[0,1]$;

$$\sum_{-\infty}^{\infty} |\hat{f}_k|^2 \leq ||f||^2 = \int_0^1 |f|^2.$$

This immediately follows from the inequality for all finite sums obtained above.

Observe that convergence of the series tells us that for any $f \in C[0,1]$ $\hat{f}_k \to 0$ as $k \to \pm\infty$.

5^o. Let now f and f' be period one functions. Then

$$\widehat{f'}_k = -2\pi i k \hat{f}_k.$$

This is a simple consequence of integration by parts.

$$\widehat{f'}_k = \int_0^1 f'(x)e^{-2k\pi ix}dx = f(x)e^{-2k\pi ix}\Big|_0^1 - \int_0^1 f(x)(-2k\pi i)e^{-2k\pi ix}dx.$$

The first term on right side is zero because the functions take same values at both end points.

Thus differentiating f is transformed to multiplication in the Fourier domain.

This is an extremely powerful result as the next observation shows.

6^0. Let f be as in the theorem. Then the sequence (S_N) is uniformly Cauchy.

Take $n < m$ positive integers, then

$$|S_n(x) - S_m(x)| = |\sum_{n<|k|\leq m} \hat{f}_k e_k| \leq |\sum_{n<|k|\leq m} |\hat{f}_k|$$

$$\leq \left| \sum_{n<|k|\leq m} \widehat{f'}_k \frac{1}{2\pi ik} \right|$$

$$\leq \frac{1}{2\pi} \sqrt{\sum_{n<|k|\leq m} |\widehat{f'}_k|^2} \ \sqrt{\sum_{n<|k|\leq m} 1/k^2}$$

$$\leq \frac{1}{2\pi} \sqrt{\int |f'|^2} \ \sqrt{\sum_{|k|>n} 1/k^2}.$$

Denote $M = \sqrt{\int |f'|^2}/(\pi)$ we have

$$\sup_x |S_n(x) - S_m(x)| \leq M \sqrt{\sum_{k>n} 1/k^2}$$

Since the series $\sum 1/k^2$ is convergent the tail sums can be made small and thus the right side above can be made small for all large values of n.

This shows that the sequence of functions (S_N) is Cauchy uniformly.

7^o. If we now show that the sequence S_N converges point wise to f we would have proved our theorem. To this end, let us understand the partial sums.

$$S_N(x) = \sum_{-N}^{N} \hat{f}_k e_k = \sum \int f(y)e^{-2\pi kiy}dy \ \ e^{2\pi ikx}$$

$$= \int f(y) \sum_{-N}^{N} e^{2\pi i k(x-y)} dy.$$

Let us denote

$$D_N(\theta) = \sum_{-N}^{N} e^{2\pi i k\theta}.$$

Then we have

$$S_N(x) = \int_0^1 f(y) D_N(x-y) dy.$$

Using formula for sum of finite geometric series,

$$D_N(\theta) = e^{-2\pi i N\theta} \frac{1 - e^{2\pi i(2N+1)\theta}}{1 - e^{2\pi i\theta}}.$$

When $\theta = 0$, this is to be interpreted as $(2N+1)$.

$$D_N(\theta) = \frac{e^{-2\pi i N\theta} - e^{2\pi i(N+1)\theta}}{1 - e^{2\pi i\theta}}$$

$$= \frac{e^{-2\pi i(N+1/2)\theta} - e^{2\pi i(N+1/2)\theta}}{e^{-\pi i\theta} - e^{\pi i\theta}}$$

$$= \frac{\sin \pi(2N+1)\theta}{\sin \pi\theta}.$$

As a consequence of the orthogonality of the exponential functions, we also get from the above summation,

$$\int_0^1 \frac{\sin \pi(2N+1)\theta}{\sin \pi\theta} d\theta = 1.$$

Thus

$$S_N(x) = \int_0^1 f(y) \frac{\sin \pi(2N+1)(x-y)}{\sin \pi(x-y)} dy$$

Change the variable, notice integrand is a period one function so that you need not worry about range of integration, any interval of length one would give the same answer.

$$S_N(x) = \int_{-1/2}^{1/2} f(x-y) \frac{\sin \pi(2N+1)y}{\sin \pi y} dy.$$

Let us from now on fix a point x. We shall show $S_N(x)$ converges to $f(x)$. Remembering that D_N integrates to one we see

$$S_N(x) - f(x) = \int_{-1/2}^{1/2} [f(x-y) - f(x)] \frac{\sin \pi(2N+1)y}{\sin \pi y} dy.$$

Observe that the function

$$\varphi(y) = \frac{f(x-y) - f(x)}{\sin \pi y}$$

is a nice continuous function on $[-1/2, 1/2]$. Its value at zero is

$$\varphi(0) = f'(x) \cdot \frac{1}{\pi} \cdot (-1).$$

Thus we can write

$$S_N(x) - f(x) = \int \varphi(y) \left[e^{\pi i(2N+1)y} - e^{-\pi i(2N+1)y} \right] \frac{1}{2i} dy$$

$$= \int \varphi(y) e^{\pi i y} \frac{1}{2i} \ e^{2N\pi i y} dy - \int \varphi(y) e^{-\pi i y} \frac{1}{2i} \ e^{-2N\pi i y} dy$$

The first term on right side is $(-N)$-th Fourier coefficient of some continuous function and hence converges to zero as $N \to \infty$. Similarly, second term is N-th Fourier coefficient of a continuous function which again converges to zero as $N \to \infty$.

This shows that for each x, $S_N(x) - f(x) \to 0$.

Since (S_N) is already shown to be uniformly Cauchy, we conclude that S_N converges to f uniformly.

This completes proof of the theorem.

8^o. Let f be any function as in the proof of the theorem, that is, both f and f' are continuous of period one.

$$S_N \to f \qquad \text{uniformly}$$

Thus

$$\overline{S_N} \to \overline{f} \qquad \text{uniformly}$$

Hence

$$|S_N|^2 = S_N \overline{S_N} \to |f|^2 \qquad \text{uniformly}$$

And thus finally

$$\int_0^1 |S_N|^2 \to \int_0^1 |f|^2.$$

But

$$||S_N||^2 = \sum_{-N}^{N} |\hat{f}_k|^2 \to \sum_{-\infty}^{\infty} |\hat{f}_k|^2.$$

so we conclude

$$\sum |\hat{f}_k|^2 = \int_0^1 |f|^2.$$

This is called *Plancherel equality.* remember we have proved an inequality earlier, for all functions f, not necessarily f of the theorem.

9^o Since polynomials are dense, it is reasonable to believe that the above equality remains to hold for all f, not necessarily for f satisfying the theorem. This is indeed true and we shall prove shortly. But why are we interested in this? Let me convince you of its importance.

10^o.
Consider the function

$$f(x) = x, \qquad 0 \le x \le 1.$$

This is a continuous function and of course not periodic, takes different values at zero and one. However from what we said above,

$$\sum |\hat{f}_k|^2 = \int_0^1 |f|^2.$$

Let us see what it means.

$$\int_0^1 x^2 dx = 1/3.$$

$$\hat{f}_0 = \int_0^1 x dx = 1/2.$$

For $k \neq 0$, integration by parts gives

$$\int_0^1 x e^{-2\pi i k x} dx = -\frac{1}{2\pi i k}$$

Thus we have

$$\frac{1}{4} + 2\sum_1^\infty \frac{1}{4\pi^2 k^2} = \frac{1}{3}$$

In other words

$$\sum_1^\infty \frac{1}{k^2} = \frac{\pi^2}{6}.$$

It is very interesting, in the very first course we showed that the series $\sum(1/k^2)$ converges, but we had to wait for an year to find out what the sum is.

One can use this method to evaluate sum of even powers of $1/k$. for example

$$\sum_1^\infty \frac{1}{k^4} = \frac{\pi^4}{90}.$$

Try your hands on it.

Unfortunately, we still do not know sum of odd powers.

It remains to show that Plancherel equality hods for all $f \in C[0,1]$. This is not too difficult but needs some careful analysis.

The first step consists of understanding the map that sends f to its Fourier coefficients. Let $X = C[0,1]$, space of complex valued continuous functions on the interval $[0,1]$. We equip it with the metric

$$d(f,g) = \int_0^1 |f-g|^2.$$

Note we are not taking sup metric.

Let $Y = l_2$ be the space of all complex sequences $z = (z_n; -\infty < n < \infty)$ such that $\sum |z_n|^2 < \infty$. We considered this space earlier but there are two main differences. First, earlier we considered only real sequences. Now we are forced to consider complex sequences. So let us do, we have no choice! Second difference is that earlier we considered one sided sequences, but now we are forced to consider two sided sequences. that is, functions defined on Z rather than on N. We have no choice!

You pause for a moment and make sure that you understand. Luckily, a series of positive numbers converges iff the (finite) partial sums are bounded. Then the sup of the partial sums is the sum of the infinite series. This is so even if you have two sided sequences. We said $\sum |z_n|^2 < \infty$. You can interpret the infinite sum as limit of the sequence

$$s_N = \sum_{-N}^{N} |z_k|^2.$$

Even if you take, for example,

$$t_N = \sum_{-2N}^{3N} |z_|^2$$

then this sequence t_N converges and converges to the same limit as above. But, if you are getting confused, you can just keep in mind the symmetric sums s_N. It is OK at first attempt.

The Cauchy Schwarz inequality still holds. After all, the C-S inequality for infinite sums was obtained as limit from the finite case. The only problem is that we proved for real numbers. Suppose you have complex numbers $(a_k : 1 \leq k \leq n)$ and $(b_k : 1 \leq k \leq n)$,

$$|\sum a_k b_k| \leq \sum |a_k||b_k| \leq \sqrt{\sum |a_k|^2}\sqrt{\sum |b_k|^2}.$$

Here the first inequality is simply that mod of sum is smaller than sum of mod. The second inequality is usual C-S for real numbers, no more complex numbers!.

It is the C-S that shows

$$d(z, w) = \sqrt{\sum_{-\infty}^{\infty} |z_k - w_k|^2}$$

is a metric on the space l_2.

Exactly as in the case of the l_2 that we considered, this is also complete. Take a Cauchy sequence; show each coordinate converges; using that it is Cauchy in your metric show that this coordinatewise limit is actually in your space (show finite partial sums are bounded); show that the convergence takes place in the l_2-metric.

Define the map $T : X \to Y$ by

$$Tf = \{\hat{f}_k : -\infty < k < \infty\}$$

We already showed, right after Bessel inequality, that for $f \in C[0, 1]$, the sum $\sum |\hat{f}_k|^2 < \infty$. In other words, $Tf \in l_2$. Let D be the set of all $f \in C[0, 1]$ that satisfy the theorem. We shall show that D is dense in the metric d on the space X. Clearly, D is a linear subspace of X. Also

$$T(f - g) = Tf - Tg$$

What we proved amounts to saying that the map T is an isometry of D into l_2.

To proceed further let us recapitulate something that we discussed already once. Let X and Y be two metric spaces and Y be complete. Suppose D is a dense subset of X and T is an isometry of D to Y. Then we can extend T as an isometry of X to Y. Proof is trivial. Take any $x \in X$. Since D is dense, take $p_n \in D$ such that $p_n \to x$. Hence (p_n) is Cauchy; use T is isometry to conclude (Tp_n) is Cauchy in Y; use Y is complete to get its limit and declare it as Tx. This is good definition because if you take $q_n \to x$ in D; then $d(p_n, q_n) \to 0$; T being isometry $d(Tp_n, Tq_n) \to 0$; thus (Tp_n) and (Tq_n) have same limit. Finally, this extension is isometry because if you take x and y and $p_n \to x$ and $q_n \to y$ then using the fact that metric is 'continuous' we have

$$d(Tx, Ty) = \lim d(Tp_n, Tq_n) = \lim d(p_n, q_n) = d(x, y).$$

(Did you realize we are using the same symbol d for metric everywhere!)

Returning to our problem at hand, the Fourier coefficient map T from $D \subset C[0, 1]$ to l_2 is an isometry and hence can be uniquely extended as an isometry on all of $C[0, 1]$ to Y. Interestingly, we already have the existing map T defined on all of $C[0, 1]$. Is this extension same as the existing map? If so it is an isometry already and this is exactly what we need!

We shall sort out this next time and complete this discussion.

Plancherel:

We have $X = C[0,1]$, complex valued continuous functions on $[0,1]$ with metric
$$d(f,g) = \sqrt{\int_0^1 |f-g|^2}$$
We have l^2, space of (two sided) infinite sequences which are square summable with metric
$$d(a,b) = \sqrt{\sum |a_k - b_k|^2}.$$
We have map from X to Y
$$Tf = \{\hat{f}_k : -\infty < n < \infty\}.$$
We have the set D of all functions that satisfy the convergence theorem. That is, all C^1 functions on $[0,1]$ with $f(0) = f(1)$ and $f'(0) = f'(1)$.

We claim that D is dense in $C[0,1]$. Remember the metric is d above, not the sup metric. Of course, D can not be dense in sup metric. This is so because convergence in sup metric, in particular, implies point-wise convergence and so period one stays in the limit (though not differentiability).

Let f be given. Assume that f is real valued. Let $\epsilon > 0$ be given. By Weierstrass we can get a polynomial P with sup distance from f smaller than ϵ. Since length of interval of integration is one, we see $d(f,P) < \epsilon$.

To continue, here is an observation. Given $-\infty < a < b < c < d < \infty$ there is a C^1 function φ on R such that $\varphi(x)$ is

zero for $x \leq a$;

then increases to one as x increases to b;

then remains one up to c;

then decreases to zero as x further increases to d;

then remains zero for $x \geq d$.

Such a function, actually C^∞ function was constructed last year. But since we do not need C^∞ now, we can do it in a simpler way as follows.

Define ψ on R:

zero below a;

between a and b its graph is an isosceles triangle with base $[a, b]$ and height $+2/(b-a)$ thus graph is above x-axis;

again zero between b and c;

between c and d its graph is an isosceles triangle with base $[c, d]$ and height $-2/(d-c)$ thus graph is below x-axis;

from d onwards zero again.

Now take indefinite integral

$$\varphi(x) = \int_{-\infty}^{x} \psi(t)dt.$$

By fundamental theorem of calculus, $\varphi' = \psi$ and is hence continuous and φ has all the required properties.

Now to continue with our earlier argument, let P be a real polynomial on $[0, 1]$ Suppose $|P| \leq M$. Take φ of the above paragraph with

$$a = 0; \quad b = 1/8M; \quad c = 1 - (1/8M); \quad d = 1$$

and take the product $g = P\varphi$. Observe that g and g' are continuous and periodic; g and P agree on a large part of the interval (namely, on $[b, c]$) and direct calculation shows you $d(P, g) < \epsilon$

By triangle inequality, we have proved the following. Given any real continuous f on $[0, 1]$, there is $g \in D$ with $d(f, g) < \epsilon$. Complex case follows. If $f = f_1 + if_2$ complex valued, get g_1 and g_2 for f_1 and f_2 with $\epsilon/2$ and argue that by taking $g = g_1 + ig_2$ we have $d(f, g) < \epsilon$ and $g \in D$.

All this goes to show that D is a dense subset of $C[0, 1]$.

Observe that we have already shown, after Bessel inequality,

$$\sum |\hat{f}_k|^2 \leq \int |f|^2$$

Since $T(f - g) = Tf - Tg$, conclude

$$d(Tf, Tg) \leq d(f, g)$$

In other words, T is a continuous map. But remember that T is an isometry on D. If you now take $f, g \in C[0, 1]$; you can get $f_n \in D$; $f_n \to f$ and $g_n \in D$; $g_n \to g$. Since both T and the distance function are continuous, we get

$$d(Tf, Tg) = \lim d(Tf_n, Tg_n) = \lim d(f_n, g_n) = d(f, g)$$

This shows that T is an isometry on all of X into Y.

since $T0 = 0$, taking $g = 0$ we get Plancherel for all $f \in C[0,1]$.

Thus we have completed proof of Plancherel identity and thereby completed proof of the fact $\sum(1/k^2) = \pi^2/6$.

Jacobi identity:

For real number $t > 0$ define

$$\vartheta(t) = \sum_{-\infty}^{\infty} e^{-n^2\pi t.}$$

This is called theta function. This series is convergent because $\exp\{-\pi t\} < 1$ and geometric series $\sum a^n$ converges and the above series is dominated by this geometric series.

This function appears in many contexts: Riemann zeta function, number theory, statistical physics. Here is an useful identity due to Jacobi.

$$\vartheta(t) = \frac{1}{\sqrt{t}}\,\vartheta(1/t).$$

We shall prove this identity now. Consider the function

$$f(x) = \sum_{-\infty}^{\infty} e^{-(x-n)^2/2t}; \qquad 0 \le x \le 1.$$

This series is uniformly convergent. In fact, fix any $x \in [0,1]$.
If $|n| = 2$, then $(x-n)^2 \ge 1^2$; if $|n| = 3$, then $(x-n)^2 \ge 2^2$; and so on; thus this series is dominated (beyond the terms $n = 0, \pm 1$) again by a convergent series of numbers and is hence uniformly convergent.

In particular this series defines a continuous function on $[0,1]$. This is also periodic, $f(0) = f(1)$; it is the same sum, just the terms get shifted. Let us compute its Fourier coefficients.

$$\hat{f}_k = \int_0^1 \sum_{-\infty}^{\infty} e^{-(x-n)^2/2t} e^{-2\pi i k x} dx;$$

Because of uniform convergence you can interchange the order of integration and summation;

$$= \sum_{-\infty}^{\infty} \int_0^1 e^{-(x-n)^2/2t} e^{-2\pi i k x} dx;$$

Now a change of variable and the realization that $\exp\{2\pi ikn\} = 1$ for any integer n will tell us the following. For example take the term $n = -1$

$$\int_0^1 e^{-(x+1)^2/2t} e^{-2\pi ikx} dx = \int_1^2 e^{-y^2/2t} e^{-2\pi iky} dy$$

(too lazy to write for general $\pm n$) we see

$$\hat{f}_k = \int_{-\infty}^{\infty} e^{-y^2/2t} e^{-2\pi iky} dy.$$

Fortunately, this is something we had already known, done under the name characteristic function of the normal distribution. Recall

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-x^2/2} e^{iux} dx = e^{-u^2/2}.$$

Changing the variable x to $y/\sqrt{t}$ we get

$$\frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{\infty} e^{-y^2/2t} e^{iuy/\sqrt{t}} dy = e^{-u^2/2}.$$

Remember $t > 0$ is fixed. Use this formula with $u = -2\pi k\sqrt{t}$ to see

$$\frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{\infty} e^{-y^2/2t} e^{-2\pi ikty} dy = e^{-4\pi^2 k^2 t/2} = e^{-2\pi^2 k^2 t}.$$

or

$$\int_{-\infty}^{\infty} e^{-y^2/2t} e^{-2\pi ikty} dy = \sqrt{2\pi t} \;\; e^{-2\pi^2 k^2 t}.$$

Thus returning to our calculation

$$\hat{f}_k = \sqrt{2\pi t} \;\; e^{-2\pi^2 k^2 t}.$$

What an achievement! we have been able to evaluate a rather complicated looking integral, the integrand is itself an infinite series.

Suppose that someone tells us that the Fourier expansion is valid for our function, that is

$$f(x) = \sum_{-\infty}^{\infty} \hat{f}_n \;\; e^{2\pi inx}.$$

Then we see that for every $x \in [0, 1]$

$$\sum_{-\infty}^{\infty} e^{-(x-n)^2/2t} = \sum_{-\infty}^{\infty} \sqrt{2\pi t} \;\; e^{-2\pi^2 n^2 t} e^{2\pi inx}.$$

Read this equation for $x = 0$ to get

$$\sum_{-\infty}^{\infty} e^{-n^2/2t} = \sum_{-\infty}^{\infty} \sqrt{2\pi t} e^{-2\pi^2 n^2 t}.$$

or

$$\frac{1}{\sqrt{2\pi t}} \sum_{-\infty}^{\infty} e^{-n^2/2t} = \sum_{-\infty}^{\infty} e^{-2\pi^2 n^2 t}.$$

read this equation for $t/2\pi$ instead of t.

$$\frac{1}{\sqrt{t}} \sum_{-\infty}^{\infty} e^{-n^2\pi/t} = \sum_{-\infty}^{\infty} e^{-n^2\pi t}.$$

This is precisely what we are looking for.

The validity of Fourier expansion can be justified by differentiating the series for $f(x)$ term by term and showing that the series so obtained is uniformly convergent. Recall that if the term-by-term derived series is uniformly convergent then the original series can be differentiated term-by-term. We also can see that the derived series is continuous and period one function. All this just depends on the fact that when $0 < a < 1$, the series $\sum_{n \geq 1} na^n$ converges. However we shall not pause to verify the details.

Returning to the first convergence theorem, it is also possible to prove convergence theorem for not necessarily smooth functions. Let f be a continuous periodic function. Define the partial sums S_N as earlier for $N = 0, 1, 2, \cdots$. Now take their averages

$$\sigma_n = \frac{1}{n}\{S_0 + S_1 + \cdots + s_{n-1}\}$$

Then σ_n converges uniformly to f on $[0, 1]$. This is known as Fejer's theorem and proved by realizing that

$$\sigma_n(x) = \int f(y) F_n(x - y) dy$$

where F_n is Fejer kernel

$$F_n(\theta) = \frac{1}{n} \left[\frac{\sin n\pi\theta}{\sin \pi\theta}\right]^2$$

As you may have noticed we are dealing with expressions like

$$\int f(y) K(x - y) dy.$$

This leads to a general theory called convolutions. Thus the above function of x is called convolution of the two functions f and K.

Thus S_N is convolution of f and D_N; whereas σ_n is convolution of f and F_n. The point is that these functions D_N or F_n put more and more 'mass' at the point zero as N or n becomes large; thus when you translate by x they put more and more mass at x and thus capture the value of f at x. This concept has a beautiful theory behind it that goes by the name of approximate identities. But we shall discuss no more.

This completes our discussion of Fourier series.

HA:

Three problems from home assignments need to be sorted out.

1. (♠) :

There is again an error, now in the last home work problem.

As pointed out by Uma, it is not a metric. Here is the idea, if you have two points on the circle, their distance is the length of the 'smaller part' of the arc between them. My error lies in saying that for points in $(0,1)$ the distance is the good old one. No.

Here is the correction and very brief solution.

Let $X = [0,1]$ usual metric. Let $Y = (0,1) \cup \{\spadesuit\}$ Thus Y has all points of X except zero and one, instead it has one extra point. This is the bag containing zero and one. Here is the metric

$$d^*(s,t) = \min\{|t-s|; \;\; s \wedge t + 1 - (s \vee t)\}; \quad 0 < s,t < 1$$

$$d^*(s,\spadesuit) = d^*(\spadesuit, s) = \min\{|s-0|, |s-1|\} = s \wedge (1-s); \quad 0 < s < 1.$$

$$d^*(\spadesuit,\spadesuit) = 0.$$

If $0 < s < t < 1$ then you can go from s to t directly travelling distance $t-s$ or you can go from s to 0 which is same as 1 and then to t, you get $s + (1-t)$. This explains the first formula. The second formula is similarly explained: to go from s to 0 which is same as 1, take the shortest route.

Note that in the definition of d^* the two quantities whose minimum is being taken add to one so that the distance is always at most $1/2$.

The questions are: Does this satisfy rules for being distance function? Is the space homeomorphic to the circle? There appear to be several smart ways.

Probably, the best way is to map the space Y onto the circle by $f(s) = (\cos 2\pi s, \sin 2\pi s)$ and $f(\spadesuit) = (1, 0)$. This is clearly one-one and onto. Arc length is a distance on the circle. Arc length means distance between two points is defined as the angle they make with the origin; you take the one that is at most π. This distance gives you the usual notion of convergence on the circle.

Now argue that the map f is nearly an isometry; well it is not, it multiplies distance exactly by 2π.

The essential point of this exercise is the following. In the space Y, if you take a point x different from $\spadesuit$, then a sequence converges to x iff it converges in the space X already. Further, in the space Y, a sequence of points, different from $\spadesuit$ converges to $\spadesuit$ iff they converges to either zero or one in the space X.

2. det one matrices:

The problem is to show that the space of matrices of determinant one is connected. We fix a k and consider only matrices of order $k \times k$. Recall the concept of convergence is entry wise; that is, a sequence of matrices M_n converges to a matrix iff for each i, j their (i, j)-th entries converge. This is same as identifying the space with an appropriate subset of Euclidean space of dimension $k(k-1)/2$.

The plan is to show that it is path connected. Probably you can do easily using the Lie group formalism, or simple group theory and connected components, I have not thought about it. Probably, you can also do it by any one of several canonical reductions, like echelon form etc. But there are several hands on calculations, via: orthogonalize, normalize, (Gram-Schmidt) and rotate — go to identity matrix from any where. This is lengthy but you can easily walk through. Another one is to go to signed permutation matrix and then go to identity. I shall outline the later.

In what follows, path always means continuous path. Continuity is easy for you to verify and so we do not mention. Remember Image of a continuous function defined on an interval is a connected set. Thus if you exhibit a path from a given matrix (of the space) to identity matrix, then the space is connected. Because then, there is a path between any two matrices. In case there is a disconnection of the space as $A \cup B$, union of two disjoint non-empty closed sets; then a path from any point of A to any point of B gets disconnected.

So let M be a given $k \times k$ matrix of determinant one. We shall show a path from it to the identity matrix. this is done in two steps. Let $e_1, e_2, \ldots, e_k$ be the standard orthonormal basis. We show a path from M to the matrix where each row is $\pm e_i$ and all i appear. Second step is to show a path from there to I.

Just keep in mind that non-singularity means just that the rows are independent. Determinant is a continuous function of the matrix. Thus rows of M are independent and span all of k-dimensional space R^k.

Let r_1 be the first row of M. There is one e_i which is not in the span of the other rows. Fix one such. Let $r_i = ae_i + v$ where v is in the span of the other vectors and more importantly, $a \neq 0$. Let $f(\lambda)$ be the matrix with first row $ae_i + \lambda v$ and other rows as they are. No matter what λ is, this is not in the span of the vectors $\{r_2, r_3, \cdots, r_n\}$ Because if it were then ae_i would be and $a \neq 0$ tells e_i would be. Thus independence of rows of M and the fact that this new first row is not in the span of the rest tells you that rows of $f(\lambda)$ are also linearly independent and hence this is non-singular. If g is f divided by its determinant (divide one row); so that det remains one; you have a path from the given matrix to a matrix whose first row is ae_i with $a \neq 0$. All this is to make sure that your path lies in your space.

Now consider the matrix with first row λe_i where λ ranges between a and 1 in case $a > 0$; and ranges between a and -1 in case $a < 0$(and divide by det). You get a path to the matrix whose first row is $\pm e_i$. Remember i and the sign $\pm$ is not at our disposal, depends on the given matrix.

Thus given matrix has a path to a matrix with first row $\pm e_i$ for an i and the other rows remain as they are. Most important, the path is in our space. Now look at the second row and span of the other $(k-1)$ rows and argue exactly as above. Remember $\pm e_i$ is already in the first row now, the new vector you capture is some $\pm e_j$ with $j \neq i$. Get a path to a matrix whose first two rows are $\pm e_i, \pm e_j$.

Continue by induction to get a path from the given matrix to a matrix whose rows are $\pm e_1, \pm e_2, \cdots, \pm e_k$ in some order. Of course if these rows are $e_1, \cdots, e_k$ then we have I. But we do not know. Instead of going from here, we shall come from I to this matrix with a path. We start with two simple observations.

Consider only 2×2 matrices. There is a path from the matrix

$$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

to I. In fact the path is

$$\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

from $\theta = 0$ to $\theta = \pi$ (may be in the reverse direction). Note that all the matrices above are det one matrices (rotation).

Second observation is the following; again only 2×2 matrices. There is a path from

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

to I. In fact the the same path as above, but now from $\theta = 0$ to $\theta = \pi/2$. similarly, there is a path from

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

to I.

Now let us return to our original problem. We start with I and make a path to a matrix A of det one whose rows are $\pm e_i$. Just to make writing simpler, let us denote f_i is $+e_i$ if it appears in the matrix A; f_i is $-e_i$ if this appears in A. Thus f_i appears as arrow in A and this is $\pm e_i$.

Step 1: A consists of rows $f_1, f_2, \cdots, f_k$ in that order.

Thus A is diagonal and -1 appears at an even number of places, because det is one. Thus take two places and in that coordinate axes apply the rotation mentioned above to I to make those negative. Repeat until you achieve the matrix A. Since -1 is at even number of places, changing two at a time yields final result.

Step 2: Suppose we used e_i in the j-th row and $-e_j$ at i-th row and others are e_p at p-th row $(p \neq i, p \neq j)$. Or we have used $-e_i$ in the j-th row and e_j at i-th row and others are e_p at p-th row $(p \neq i, p \neq j)$. Thus one negative sign and one row permutation applied to I, so that the determinant remains one. In both cases the rotations in the corresponding coordinate axes will give the desired result.

Step 3: general case. Since A consists of rows f_i in some order, there is a permutation π of $\{1, 2, \cdots, k\}$ such that $f_{\pi(i)}$ appears in the i-th row. Any permutation is a composition of transpositions and each transposition changes the sign of the determinant. Using these two facts and the rotations in two coordinate axes at a time we can complete the proof.

The main point is either you have even number of minus signs and even number of transpositions OR odd number of minus and odd number of transpositions.

This can be made precise and something needs to be written, but I am not writing because I believe there must be a better way. Probably you show that det positive matrices is connected and our space is an obvious continuous image of it.

3. Baire category and polynomials:

Let f be an infinitely differentiable function on $[0, 1]$.

$(\exists\ k)\ (\forall\ x)\ f^{(k)}(x) = 0$ implies f is a polynomial.

In fact it is a polynomial of degree at most k that appears in the existential quantifier. This you already know.

It is a simple 'backward' argument. Take such k. since derivative of $f^{(k-1)}$ is zero conclude $f^{(k-1)}$ is a constant a_0. conclude $f^{(k-2)}$ is $a_0 + a_1 x$. Then conclude $f^{(k-3)}$ must be $a_0 + a_1 x + a_2 x^2$ etc.

Of course instead of etc you use induction. We have done it.

The problem to be settled is

$(\forall\ x)\ (\exists\ k)\ f^{(k)}(x) = 0$ implies f is a polynomial.

Thus for each x there is some number $n(x)$ such that the $n(x)$-th derivative at x equals zero.

I had already mentioned that this is not an easy problem. Actually the idea is simple. Let us start with some observations. The first two observations spell out the idea and the rest is just an elaboration using bare.

1^o. Suppose we know $f^{(5)} = 0$ on $(1/3, 2/3)$ and $f^{(8)} = 0$ on $(1/4, 1/2)$. Then actually $f^{(5)} = 0$ on the union $(1/4, 2/3)$.

This is because f is a polynomial $P(x)$ in $(1/3, 2/3)$ and a polynomial $Q(x)$ in $(1/4, 1/2)$. Thus in the common part f equals both P and Q. In other words the two polynomials equal on the common part, which is a non-degenerate interval. This implies that the two polynomials must be same.

2^o. Suppose (i) at every point x of C Cantor set $f^{(8)}(x) = 0$ and (ii) f in each open interval complementary to C, we know f is a polynomial (depending on the interval). Thus in one deleted interval it may be poly of degree at most 20 and in other deleted interval it may be poly of degree at most 100. Then we claim that actually f is a polynomial of degree at most 8.

You see the interesting formulation. If $f^{(8)}(x)$ is zero for all points of an open interval you can say that f is a poly of degree at most 8 on that open interval. But if $f^{(8)}(x)$ is zero for all points of Cantor set, it does not make sense to say that f is a poly of degree at most 8 at all points of the Cantor set.

To get clear idea whether such a thing is possible at all, take the function $h(x) = d(x, C)$ which is continuous and is zero exactly at points of C. (We can give many examples of such continuous functions which are zero exactly at points of C). If

$$f(x) = \int_0^x h(y)dy$$

then derivative of f is zero exactly for points of the Cantor set. You can keep taking indefinite integrals. pause and think.

Let us prove the assertion 2^o. The proof needs a couple of steps.

First we claim that not only $f^{(8)}(x)$ but $f^{(k)}(x)$ is zero for all $k \geq 8$ and all $x \in C$. The crucial point is that f is differentiable any number of times and every point of C is a limit point of C. Given $x \in C$, the second property enables you to take $p_n \in C$, $p_n \neq x$ and $p_n \to x$. The first property enables you to calculate

$$f^{(9)}(x) = \lim \frac{f^{(8)}(p_n) - f^{(8)}(x)}{p_n - x} = 0.$$

Thus $f^{(9)}(x) = 0$ for all $x \in C$. Now use induction.

The plan is to show that at every point $f^{(8)}(x) = 0$. Of course this is granted to you for points in C. Need to show for points outside C, that is, in each of the deleted intervals.

So let us fix any interval (a, b) complementary to C. Thus a and b are in C. Let us say f is a poly of degree k on this interval. In case $k \leq 8$, then there is nothing for you to do; by differentiation, you get $f^{(8)}(x) = 0$ on this interval. Suppose $k > 8$. For instance say $k = 9$. Then $f^{(8)}(x)$ is a continuous function on $[a, b]$ and equals zero at end points and its derivative, namely $f^{(9)}$ is zero inside the interval should tell you, by integration, $f^{(8)} \equiv 0$ on all of $[a, b]$.

In the general case you repeat the same argument. Suppose $k > 8$. Remember: we know higher derivatives are all zero at points of C. Thus $f^{(k-1)}$ is zero at the end points a and b of the interval $[a, b]$ and it is continuous and $f^{(k)}$ is zero inside. Thus $f^{(k-1)} \equiv 0$ on $[a, b]$. Now argue $f^{(k-2)} \equiv 0$ on $[a, b]$ and continue till you reach 8. This completes the proof of 2^o.

3^0. Let us return to our problem. Set

$$A_k = \{x : f^{(k)}(x) = 0\} \qquad k =, 1, 2, 3, \cdots.$$

Then each A_k is closed because each derivative is continuous and $\cup A_k = [0, 1]$ by hypothesis.

The sets A_k, as of now, may neither increase nor decrease.

For example, $f(x) = x^3 + x$ has $f'(0) \neq 0$ but $f''(0) = 0$.

For $g(x) = x^2$ we have $g'(0) = 0$ but $g''(0) \neq 0$.

However, their interiors A_k^o are increasing. This is because, if the k-th derivative is zero on an open interval, then the later derivatives are zero on that open interval.

But interestingly, suppose A_k^0 is disjoint union of non-empty open intervals $I_p = (a_p, b_p) : p = 1, 2, \cdots$. Then none of these end points can be interior points at any time later. This is precisely the content of observation 1^o. Thus any future interiors that get added are disjoint to what you already have.

This observation has the following implication. Let

$$V = \cup A_k^o$$

say V is the disjoint union of open intervals $I_n : n \geq 1$. Remember any (non-empty) open set is union of disjoint non-empty open intervals in only one way. Then each of these I_n is already contained in some A_k^o.

This in turn has the following implication. In each I_n our f is a polynomial.

Let us denote $P = [0,1] - V$. There are two possibilities: either P is empty or not. Suppose it is empty, then V is all of $[0,1]$. In other words the increasing union A_k^o covers all of the compact $[0,1]$ and so must already equal this for some k and this then completes what we wanted to set out to prove.

Let us now assume that $P \neq \emptyset$. We claim that it must be a perfect set. Indeed it is closed being complement of open set V. If it has a point which is not a limit point, then, we have two intervals (a,b) and (b,c) in V but b is not. But then as mentioned above $f^{(i)} \equiv 0$ in (a,b) and $f^{(j)} \equiv 0$ in (b,c). If $m = i \vee j$ then $f^{(m)} = 0$ on both the intervals and hence, by continuity, on all of (a,c). Contradiction because then $(a,c) \subset A_m^o$.

Thus P is a non-empty perfect set. Of course

$$P = (A_1 \cap P) \ \cup \ (A_2 \cap P) \ \cap \ (A_3 \cap P) \ \cup \cdots.$$

But P is a polish space in its own right and these are closed subsets of P and thus one of them must contain an open subset of P. In other words there is an k and open interval (a,b) such that

$$(a,b) \cap P \subset A_k \cap P$$

. Now argument 2^o leads to a contradiction. Every point of P in (a,b) is a limit point of P and so not only $f^{(k)}(x) = 0$ for all $x \in (a,b) \cap P$ we actually have

$$f^{(m)}(x) = 0; \quad \forall m \geq k \quad \forall x \in (a,b) \cap P.$$

This will help you to show (either by differentiation or by integration, as the case may be) to show $f^{(k)}(x) = 0$ for all x in any of the deleted intervals. This ultimately shows that $f^{(k)}(x) = 0$ for all $x \in (a,b)$. In other words $(a,b) \subset A_k^o$ and would have no point in P. This is a contradiction.

This semester we shall not continue with calculus. So this exercise set does not represent our topics. It is intended to bring you back to the mood and also remind you that there are many interesting matters we did not discuss.

1. Show that there are no real numbers a and b such that $\frac{1}{a}+\frac{1}{b}=\frac{1}{a+b}$. But there are complex numbers satisfyng it.

2. Let a and b be real numbers. Show

$$\int_0^\infty \frac{1}{1+x^2}\,\frac{x^b-x^a}{(1+x^a)(1+x^b)}\,dx \;=\; 0.$$

Deduce

$$\int_0^\infty \frac{1}{1+x^2}\,\frac{1}{1+x^a}\,dx \;=\frac{1}{2}\int_0^\infty \frac{1}{1+x^2}\,dx \;=\; \frac{\pi}{4}.$$

3. Fibonacci numbers are defined by $F_0=F_1=1; F_{n+2}=F_{n+1}+F_n$ for $n\geq 0$. Show $\sum\limits_0^\infty F_n t^n \;=\; \frac{1}{1-t-t^2}$.

4. Let f and g be functions on R to R which are differentiable as many times as needed below. Show Leibniz's rule:

$$D^p(fg) \;=\; \sum_{k=0}^{p}\binom{p}{k} D^k f\; D^{p-k} g.$$

Calculate the first ten Bernoulli numbers defined by

$$\frac{t}{e^t-1} \;=\; \sum_0^\infty \frac{B_n}{n!} t^n.$$

Show

$$\sum_{j=0}^{n-1} e^{jt} \;=\; \frac{e^{nt}-1}{t}\,\frac{t}{e^t-1} \;=\; f(t)g(t) \quad \text{say.}$$

Show

$$\sum_{j=0}^{n-1} j^p \;=\; \sum_{k=0}^{p}\binom{p}{k}(D^k f)(0)\;(D^{p-k}g)(0).$$

$$=\;\sum_{k=0}^{p}\binom{p}{k}\frac{n^{k+1}}{k+1}\;B_{p-k} \;=\; \sum_{k=0}^{p}\binom{p+1}{k+1}\frac{n^{k+1}}{p+1}\;B_{p-k}.$$

Deduce

$$1+2+3+\cdots+n \;=\; \frac{1}{2}n^2+\frac{1}{2}n.$$

$$1^2+2^2+3^2+\cdots+n^2 \;=\; \frac{1}{3}n^3+\frac{1}{2}n^2+\frac{1}{6}n.$$

$$1^3+2^3+3^3+\cdots+n^3 \;=\; \frac{1}{4}n^4+\frac{1}{2}n^3+\frac{1}{4}n^2.$$

$$1^{10}+2^{10}+3^{10}+\cdots+n^{10} \;=$$

$$\frac{1}{11}n^{11}+\frac{1}{2}n^{10}+\frac{5}{6}n^9-n^7+n^5-\frac{1}{2}n^3+\frac{5}{66}n.$$

5. Let f be C^1 on $[1,\infty)$. Show

$$\sum_1^n f(k) \;=\; \int_1^n f(x)\,dx \;+\; \frac{f(1)+f(n)}{2} \;+\; \int_1^n \left(x-[x]-\frac{1}{2}\right) f'(x)\,dx$$

$P_1(x)=x-1/2$ on $[0,1)$, extended periodically to all of R. Deduce

$$\sum_1^n f(k) \;=\; \int_1^n f(x)\,dx \;+\; \frac{f(1)+f(n)}{2} \;+\; \int_1^n P_1(x)f'(x)\,dx$$

This is known as the first derivative formula (of Euler).

Let us now assume f is C^2. Set $P_2(x) = x(x-1)+1/6$ on $[0,1)$, extended periodically to all of R. Show

$$\sum_1^n f(k) = \int_1^n f(x)\,dx+\frac{f(1)+f(n)}{2}+\frac{f'(n)-f'(1)}{12}-\frac{1}{2}\int_1^n P_2(x)f''(x)\,dx$$

This is the second derivative formula.

These are Euler Summation formulae.

P_2 is continuous and determined by $P_2'=2P_1$; $\quad \int_0^1 P_2(x)dx=0$.

Here are special cases. Read Err as Error.

(a) $f(x)=1/x$ gives $\sum\limits_1^n 1/k=\log n + C + Err(n)$.

$$C = 1-\int_1^\infty \frac{x-[x]}{x^2}\,dx; \quad Err(n)=\frac{1}{2n}+\int_n^\infty \frac{P_1(x)}{x^2}\,dx.$$

(b) $f(x) = \log x$ gives

$$\log n! = (n + 1/2) \log n - n + C - Err(n).$$

giving Stirling formula $n! \sim e^C e^{-n} n^{n+1/2}$. I hope you remember that $C = \log \sqrt{2\pi}$.

(c) Fix a number $s > 0$ and $s \neq 1$. Then $f(x) = 1/x^s$ gives

$$\sum_{k=1}^{n} \frac{1}{k^s} = \frac{n^{1-s}}{1-s} + C_s + s \int_n^\infty \frac{t - [t]}{t^{s+1}}\, dt$$

where

$$C_s = 1 + \frac{1}{s-1} - s \int_1^\infty \frac{t - [t]}{t^{s+1}}\, dt.$$

Letting $n \to \infty$, note that C_s for $s > 1$, is nothing but the Riemann zeta function $\sum 1/k^s$. This last summation does not make sense for $0 < s < 1$ where as the above expression for C_s makes sense.

(d) $f(x) = \log x/x$ gives

$$\sum \frac{\log k}{k} = \frac{1}{2}(\log n)^2 + \frac{1}{2}\frac{\log n}{n} + C - Err(n)$$

6. Let C be the set of all pairs (a, b) of real numbers with usual addition $(+)$ and multiplication given by $(a, b) \times (c, d) = (ac - bd, ad + bc)$. Show that $(C, +, \times)$ is a field.

   Let M be the set of all 2×2 matrices of the form $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$ where a, b are real numbers. With usual addition and multiplication of matrices, show that it forms a field.

   Let Q be the set of all polynomials $p(t)$ in one variable t with real coefficients. Consider usual addition and multiplication of polynomials. Say $p_1 \sim p_2$ if there is a polynomial p such that $p_1 - p_2 = (t^2 + 1)p$. Let P be the set of equivalence classes. Show that the equivalence relation respects (?) the operations of addition and multiplication. Show that P is a field.

   Show that the three fields above are isomorphic in a 'canonical way'.

   For even integer $n \geq 1$, show that there are $n \times n$ invertible real matrices A and B such that

$$A^{-1} + B^{-1} = (A + B)^{-1}.$$

7. Let Q be R^4 with points written as $(a, b, c, d) = a + bi + cj + dk$. Use coordinate-wise addition. Multiplication is 'prescribed' by
$i^2 = j^2 = k^2 = ijk = -1$.
Show $ij = k, jk = i, ki = j$ and $ji = -k, kj = -i, ik = -j$.
Show that Q is nearly a field, it misses only commutativity of multiplication. It is called skew-field.

   Show M, the set of 4×4 real matrices of the form $\begin{pmatrix} a & -b & -c & -d \\ b & a & d & -c \\ c & -d & a & b \\ d & c & -b & a \end{pmatrix}$
is 'canonically' isomorphic to the above skew field. This skew field is called 'Quaternions.'

8. Let X be any set and $P(X)$ be its power set, that is, collection of all subsets of X. Define for $A, B \subset X$, their symmetric difference, by. $A \Delta B = (A - B) \cup (B - A) = \{x : x \quad \text{is in exactly one of the sets.}\}$.

   Show that $A_1 \Delta A_2 \Delta A_3 \Delta \cdots \Delta A_n$ consists of points that belong to an odd number of these sets; in whichever order you operate.

   Show that $P(X)$ is a group under the operation Δ.

   Let $H \subset P(X)$ be the collection of finite sets. Show that H is a subgroup.

   Let $H \subset P(X)$ be the collection of countable sets. Show that H is a subgroup. If H is the collection of countably infinite sets then is it a subgroup.

9. Let G be a group and X a non-empty set. The collection of functions on X to G, denoted by G^X, is a group with pointwise (?) operations.

   The set $2 = \{0, 1\}$ is a group under addition modulo 2. What is the relation between the group 2^X and the earlier $P(X)$ group.

10. Instead of finite union, you can define union of *any* family of sets. Suppose $\mathcal{C}$ is a collection of sets. Then their union, $\cup\mathcal{C}$, is the set of all objects x such that $x \in C$ for some sets$C \in \mathcal{C}$. Similarly, $\cap\mathcal{C}$ is the collection of all objects x such that $x \in C$ for every $C \in \mathcal{C}$.
    Prove DeMorgan's laws

$$(\cup\mathcal{C})^c = \cap\mathcal{C}^c; \quad (\cap\mathcal{C})^c = \cup\mathcal{C}^c; \quad \text{where} \quad \mathcal{C}^c = \{C^c : C \in \mathcal{C}\}.$$

11. For a sequence of sets (A_n), we define $\limsup A_n$ to be the set of all objects x such that x belongs to A_n for infinitely many values of n. When this happens for an object x then we also say that $x \in A_n$ *frequently*. Similarly, $\liminf A_n$ is the set of all objects x such that there is an n_0 and $x \in A_n$ for every $n \geq n_0$. When this happens for an object x we also say that $x \in A_n$ *eventually*.

    Prove De Morgan's laws:

$$(\limsup A_n)^c = \liminf A_n^c; \quad (\liminf A_n)^c = \limsup A_n^c.$$

If X is the universe of discourse (this means all sets we now consider in this exercise are subsets of this X), we define for a set $A \subset X$ its *indicator function* or *characterstic function* to be the following function defined on X: $I_A(x) = 1$ if $x \in A$ and $I_A(x) = 0$ if $x \notin A$.
This function I_A is also denoted by χ_A.

Show that $\limsup I_{A_n} = I_{\limsup A_n}$; $\quad \liminf I_{A_n} = I_{\liminf A_n}$

Let $A_n = [0, 1]$ if n is even and $A_n = [2, 3]$ if n is odd. Find limsup and liminf.

Let $\{a_n\}$ be a sequence of real numbers and $A_n = (-\infty, a_n)$ and $B_n = (-\infty, a_n]$. Calculate $\liminf A_n$ and $\limsup A_n$ in terms of liminf and limsup of a_n. Do the same for (B_n).

In what follows $N = \{1, 2, \cdots\}$. (Do you understand what exactly is hidden in the dots?).

12. Write a complete proof of the fact: If $m < n$ then there is no bijection between $\{1, 2, \cdots, m\}$ and $\{1, 2, , \cdots, n\}$.

13. If X is a countable set, then show that there is a $1-1$ function on X to (into or onto) N. Conversely, if there is a $1-1$ function on X to N then show that X is countable.

    If X is countable set, show that there is a function on N onto X. Conversely, if there is a function on N onto X, show that X is countable.

    Let X be a countable set and $Y \subset X$. Show that Y is countable.

    Let X_i for $i = 1, 2, \cdots$ be countable sets. Show that $\cup_i X_i$ is a countable set. This is stated as 'countable union of countable sets is countable'.

    (Hint: For each i fix a $1-1$ function $f_i : X_i \to \{1, 2, \cdots\}$. Here is f. Take x in the union. Take the first i such that $x \in X_i$ put

$$f(x) = 2^i 3^{f_i(x)}.$$

    If X and Y are countable, show that $X \times Y$ is also countable.

    Let seq(N) be the set of all finite sequences of integers. Show that it is countable.

    Let Y be the set of all infinite sequences of integers. Show Y is not countable.

14. If X is the set of all 5×5 real matrices and $Y = R^{25}$, exhibit a canonical bijection and conclude $|X| = |Y|$.

If X is the set of all real symmetric 5×5 matrices and $Y = R^{15}$ show that $|X| = |Y|$. If Z is the set of all upper triangular 5×5 matrices then $|Z| = |X|$.

15. If X is an infinite set, then there is a subset $Y \subset X$ which is countably infinite. Prove this.

    If X is an infinite set then show that there is a proper subset $Y \subset X$ and $Y \neq X$ such that $|X| = |Y|$. Conversely, if X is a set and if there is a proper subset $Y \subset X$ and $Y \neq X$ such that $|X| = |Y|$ then show that X is an infinite set.

16. If $|X| = |Y|$ and $|A| = |B|$ then show that $|X \times A| = |Y \times B|$.

    With the same assumption show that $|X^A| = |Y^B|$.

    Here X^A is the set of all functions from A to the set X.

17. Let Y be the set of all ordered pairs of real numbers, in other words, $Y = R^2$. Show $|Y| = |R|$.

    More generally show that $|R^n| = |R|$ for any integer $n \geq 1$.

    Let seq(R) be set of all finite sequences of real numbers and $X =$ the set of all infinite sequences of real numbers. Show both have same cardinality as R.

18. Verify if the following are equivalence relations on the sets prescribed.

    Let $X = R$. Say $a \sim b$ if $|a - b| \leq 25$.

    Let $X = R$. Say $a \sim b$ if $a - b$ is an integer. Say $a \sim b$ if $a - b$ is rational. Say $a \sim b$ if $a - b$ is irrational.

    Let $X = P(R)$. Say $A \sim B$ if $A \Delta B$ is finite. What if finite is replaced by countable? What if finite is replaced by countably infinite?

    Let X be the set of all functions from R to R. Say $f \sim g$ if $f - g$ is a continuous function. What if 'continuosu function' is replaced by 'polynomial'.

    Let $X = R^{29}$. Say $a = (a_1, \cdots, a_{29}) \sim b = (b_1, \cdots, b_{29})$ if there is a permutation π of $\{1, 2, \cdots, 29\}$ such that $a_i = b_{\pi(i)}$ for all i.

    Same X as above. Say $a \sim b$ if there is a permutation π of $\{1, 2, \cdots, 25\}$ such that $a_i = b_{\pi(i)}$ for all i with $1 \leq i \leq 25$.

    Same X as above, say $a \equiv b$ if $\sum a_i = \sum b_i$.

19. First recall The Cantor set $C \subset [0,1]$. Let $\mathcal{I}$ be the collection of all deleted open intervals of $(0,1)$. Or equivalently $[0,1] - C = (0,1) - C$ is an open set and hence can be written, in a unique way, as a disjoint union of open intervals and $\mathcal{I}$ is the collection of exactly these open intervals.

    If $I = (a,b)$ and $J = (c,d)$ are in $\mathcal{I}$ then say that $I \leq J$ if either they are the same interval or $b < c$ (that is, the interval I sits to the left of J). Show that the collection $\mathcal{I}$ with this order is another manifestation (?) of Q.

    Here is another Cantor set. Instead of dividing $[0,1]$ into three parts and removing the middle part, divide into five intervals of equal length and remove second and fourth, keep first, third, fifth. Doing this at each stage get $D \subset [0,1]$. Let $\mathcal{I}$ be the collection of deleted open intervals. Show again that $\mathcal{I}$ looks like Q (as far as order is concerned).

I hope you appreciate what AC is doing for us. Of course, set theory has many 'axioms' which are self evident (no dispute) and hence we do not talk about them. However about AC doubts exist; for very very good reasons. Here I tell you a story where AC is the main character.

20. Let $N = \{1, 2, \cdots\}$ as usual. Let $\mathcal{U} \subset P(N)$. That is, $\mathcal{U}$ is a collection of subsets of N. It is called an ultrafilter if it satisfies the following conditions:
(i) $N \in \mathcal{U}$; $\quad \emptyset \notin \mathcal{U}$;
(ii) $A, B \in \mathcal{U} \Rightarrow A \cap B \in \mathcal{U}$
(iii) $A \in \mathcal{U}, A \subset B \Rightarrow B \in \mathcal{U}$
(iv) $A \subset N \Rightarrow A \in \mathcal{U}$ or $A^c \in \mathcal{U}$.

Thus it is closed under finite intersections, supersets, not all of $P(N)$, and for every set either that or its complement is in it.

A collection $\mathcal{F}$ satisfying only the first three conditions is called a filter. Here are examples of filters. Verify.

All sets A such that A^c is finite. This is called co-finite filter. The word co-finite stands for 'complement of finite'.

All sets A such that A^c has only finitely many even numbers. For example the set of even integers is one such set.

All supersets of $\{23, 48, 56\}$

None of these is ultrafilter. Verify.

Here is an ultrafilter: All sets that contain the number 103. Verify.

You (and me too) can not think of other ultrafilters! — except replacing 103 by another number. These ultrafilters as above are called fixed ultrafilters.

Say that a filter $\mathcal{F}$ is maximal if $\mathcal{F} \subset \mathcal{G}$ and $\mathcal{G}$ is a filter imply that $\mathcal{F} = \mathcal{G}$. Thus a maximal filter is a filter which is not proper subset of another filter. Show that a filter is maximal iff it is an ultrafilter.

Use Zorn and verify there are ultrafilters which are not fixed. Such ultrafilters are called free ultrafilters.

Show that an ultrafilter is free iff it misses all singleton sets. Show that an ultrafilter is free iff it includes co-finite filter.

21. Last semester we discussed closure, boundary of sets. In that story these concepts were only supporting actors, main actors were integration and partitions. In the present scene 'closure' is the main actor. It is a good idea to recall. We do so only for R. If you understand you can do for R^n but no need to do now.

    Let $S \subset R$. A point $a \in R$ is a limit point of S if every open interval around a has infinitely many points of S. That is, for every $\epsilon > 0$ the set $A \cap (a - \epsilon, a + \epsilon)$ is infinite. We denote

$$\overline{S} \;=\; S \cup \text{ limit points of } S.$$

    This is called closure of S.

    Show that $\overline{S}$ is a closed set. Recall closed set means a set that includes all its limit points. Thus you are supposed to show that a limit point of $\overline{S}$ is already in $\overline{S}$ (in other words, it is already a limit point of S).

    Show that if C is a closed set and $S \subset C$ then $\overline{S} \subset C$. Thus $\overline{S}$ is the smallest closed set that includes S. That is, closure of a set is the smallest closed set that includes the set.

    Just to get practice, find closures of the following sets.

    $S = Q$, the set of all rational numbers.

    $S = Z$, the set of all integers.

    $S = \{1,\ 1/2,\ 1/3,\ 1/4,\ 1/5, \cdots\}$.

    $S = \{n + (1/n) : n = 1, 2, \cdots\}$.

    Remember The Cantor set? Take one point from each of the deleted intervals. Let S be the set so obtained. Do not be obsessed with AC now. Alright, take the mid point of each deleted interval. What is closure of S.

    Take the 1/10-th point of each deleted interval, what is its closure? 1/10-th point of (a, b) is the point $(9a + b)/10$. Why is it called so?

    What if you took one point from each, except some 55 of the deleted intervals?

    What if you took 5 points from each deleted interval?

    You understand a friend only if you keep interacting with the friend, not by a cursory hello. Same holds with math concepts too.

22. Now we combine ideas of the two previous exercises. Let $\mathcal{U}$ be a free ultrafilter on N. Let $\vec{x} = (x_i : i \in N)$ be a *bounded* sequence of real

numbers. let
$$\lim \vec{x} = \bigcap_{A \in \mathcal{U}} \overline{\{x_i : i \in A\}}$$

Show that the right side is non-empty and in fact is a singleton. This number is defined as limit of the sequence $\vec{x}$ *along the ultrafilter* $\mathcal{U}$ and is denoted $\lim\limits_{\mathcal{U}} x_n$.

Show that $\liminf x_n \leq \lim\limits_{\mathcal{U}} x_n \leq \limsup x_n$.

Show $\lim\limits_{\mathcal{U}} x_n = a$ iff given $\epsilon > 0$, there is a set $A \in \mathcal{U}$ such that $\{x_i : i \in A\} \subset (a - \epsilon, a + \epsilon)$.

Show that, $\lim x_n = a$ in the sense of last year Calculus iff given $\epsilon > 0$ there is a set A in the co-finite filter such that
$\{x_i : i \in A\} \subset (a - \epsilon, a + \epsilon)$.

Do you see similarity between calculus definition and the present definition?

Limit along the ultrafilter remains same if you change finitely many terms of the sequence. Show this.

Show that $\lim\limits_{\mathcal{U}} x_n$. is a limit point of the sequence (x_n). Remember limit point of a sequence is any number a such that whatever $\epsilon > 0$ you take $x_n \in (a - \epsilon, a + \epsilon)$ for infinitely many n.

Show that if the sequence (x_n) actually converges (in the sense you have learnt last year) then the limit is same as limit along the ultrafilter.

Show that for any bounded sequences,

$$\lim_{\mathcal{U}} (x_n + y_n) = \lim_{\mathcal{U}} x_n + \lim_{\mathcal{U}} y_n; \qquad \lim_{\mathcal{U}} (57x_n) = 57 \lim_{\mathcal{U}} x_n.$$

$$\lim_{\mathcal{U}} (x_n y_n) = \lim_{\mathcal{U}} x_n \cdot \lim_{\mathcal{U}} y_n. \qquad \forall n \;\; x_n \leq y_n \Rightarrow \lim_{\mathcal{U}} x_n \leq \lim_{\mathcal{U}} y_n.$$

23. Given one specific sequence, and one specific limit point of this sequence, it is possible to choose one ultrafilter so that limit along this ultrafilter for this sequence equals this given number. Have I confused you?

24. Even if your ultrafilter is a fixed ultrafilter, the above prescription of limit works; its conclusion can be explained in simpler terms and the limit could change if one term of the sequence is changed and the concept is useless. I do not want to spoil the suspense, Think.

Limit of a sequence, as learnt in Calculus last year, associates a number — namely its limit — with ‘convergent’ sequence. But the above procedure, with free ultrafilter, associates a number with *every* bounded sequence. In fact it associates a limit point of the sequence and so does not destroy what you learnt in Calculus. Of course, you might say use AC to pick one of its limit points. But the profound fact is this: the above selection respects addition, multiplication, monotonicity etc.

There is absolutely no easy way to achieve this and this is what you should appreciate.

25. We showed that there are discontinuous functions $f : R \to R$ satisfying $f(x+y) = f(x) + f(y)$ for all x, y. Suppose I want the function to satisfy both the earlier equation and also $f(xy) = f(x)f(y)$ for all x, y then what are the solutions. Answer: $f(x) \equiv 0$ and $f(x) \equiv x$. No more!

    Here is a way to prove. If $f(1) = 0$ show $f \equiv 0$. Let now $f(1) \neq 0$.

    Show (i) $f(x) \geq 0$ if $x \geq 0$; (ii) $f(x) \geq f(y)$ if $x \geq y$; (iii) $f(1) = 1$; (iv) $f(q) = q$ for rational q.

    If for some x we have $f(x) > x$ take q strictly in between and argue $f(q) \geq f(x)$, recognize this contradiction.

    Similarly there is no x with $f(x) < x$.

26. Let $S \subset R$ be open. Show that $|S|$ is either 0 or c.

    Let $S \subset R$ be closed. Its cardinality can be any of the following: $0,\ 1,\ 2,\ 3,\ 4, \cdots,\ \aleph_0,\ c$. Give examples of each.

    Can it be anything else? No. To see this, assume S is not countable.

    First suppose that every point of S is a limit point of S. Show that there are two disjoint closed intervals I_0 and I_1, each having points of S. Show that there are two disjoint closed intervals within each I_k having points of S. Continue and conclude $|S| = c$.

    To understand the general case, say that an open interval J is small if it has rational end points and $S \cap J$ is countable. Let V be the union of small intervals. Show $S_1 = S \cap V^c$ is closed, uncountable, every point of S_1 is a limit point of S_1. Conclude the proof.

27. We consider $V = R^{37}$ as a vector space, as usual, over the field R.

    An ordered basis is a sequence of vectors $\langle v^1, v^2, \cdots, v^{37} \rangle$ where $v^i \in R^{37}$ for all i; which forms a basis for V. Let X be the set of all ordered bases of V. Show $|X| = c$.

    Let G be the set of all 23-dimensional subspaces of V. Show $|G| = c$.

    Show that each of the following collections in R^2 has power c: the set of equilateral triangles, set of rectangles with sides parallel to the axes, the set of all rectangles, set of circles, set of regular hexagons.

28. Show that the set of *all* functions from R to R has power strictly larger than c. This means there is a injection from R to this set but no bijection.

29. Let X be the space of infinite sequences of real numbers. That is

$$X = \{(x_1, x_2, \cdots) : x_i \in R \quad \text{for all} \quad i\}.$$

Show that $|X| = c$.

Let $C(R)$ be the collection of all real valued continuous functions on R. We define a map from $C(R)$ to X as follows. Fix once and for all an enumeration of Q, rationals: $q_1, q_2, q_3, \cdots$.

Here is the map. For $f \in C(R)$ associate the sequence

$$s(f) \quad = \quad \{f(q_1), f(q_2), f(q_3), \cdots\}.$$

Show this is an injection and conclude that $|C(R)| = c$.

What is the power of the set of real valued continuous functions on R^5.

What is the power of the set of continuous functions on R^{17} to R^{71}.

The set of real valued continuous functions on R has the same power as R.

Let P_1 be the set of polynomials in one variable x with real coefficients. $|P_1| = c$. Show.

Let P_k be the set of polynomials in k variables $x_1, x_2, \cdots x_k$ with real coefficients. $|P_k| = c$. Show.

What if P_k is the set of polynomials in k variables as above but with rational coefficients.

Let Homeo(R) be the set of homeomorphisms of R onto itself. Show that its power is c.

30. Let M be the set of monotone increasing functions on R to R. That is $f \in M$ if $f : R \to R$ and $x < y$ implies $f(x) \leq f(y)$. The function need not be strictly increasing. Show $|M| = c$.

Just raise your heels and stretch your hands to reach this result.

First: recall from last semester (prove again) that $f \in M$ implies that f has only countably many discontinuities.

Second: Fix a countable set $D \quad = \quad \{d_1, d_2, \cdots\} \quad \subset \quad R$.

For $f \in M$, associate the sequence of numbers

$$s(f) = \{f(q_1), f(d_1), f(q_2), f(d_2), \cdots\}.$$

Denote by M_D those functions in M whose set of discontinuities are contained in D. That is, those functions which are continuous at every $x \in R - D$. Show that the above map is injective on this set and conclude $|M_D| = c$.

How many countable subsets of R are there? What is $|R \times R|$?

Use answers to the above two queries to conclude $|M| = c$.

Show that the set of all monotone functions (increasing or decreasing) from R to R has power c.

31. How large a collection of non-empty subsets of N can you get so that any two are disjoint? That is, I want a family of sets $(A_\alpha : \alpha \in \Delta)$, each $\emptyset \neq A_\alpha \subset N$; such that for distinct indices $A_\alpha \cap A_\beta = \emptyset$. How large can the set Δ be? Show that it must be countable.

    What if we need the non-empty sets to be only almost disjoint? That is, I want a family $(A_\alpha; \alpha \in \Delta)$ of non-empty subsets of N so that for distinct indices $A_\alpha \cap A_\beta$ is a *finite set*, could be but need not be empty. How large can Δ be? Answer: $|\Delta| = c$. To see this, First argue $|\Delta| \leq c$.

    Fix an enumeration $\{r_1, r_2, \cdots\}$ of rational numbers. You can do this explicitly without using AC (but this is beside the point). For $x \in R$, let us define a set as follows. x_1 is the first rational in $(x-1, x+1)$. x_2 is the first rational that occurs after x_1 which is in the interval $(x - 1/2, x + 1/2)$. In general x_k is the first rational that occurs after x_{k-1} which is in the interval $(x - 1/k, x + 1/k)$. Let

    $$A_x = \{x_1, x_2, \cdots\}$$

    $A_x \subset Q$; if $x \neq y$ then $A_x \cap A_y$ is finite.

    Returning to the problem we started with, show $|\Delta| = c$.

32. Let P be a partially ordered set and $C \subset P$ be a chain.

    Understand the difference between the following statements:

    (1) $\neg\ \exists(p \in P)\ \forall(x \in C)\ (x < p)$.

    (2) $\exists(p \in P)\ \neg\ \forall(x \in C)\ (x < p)$.

    (3) $\exists(p \in P)\ \forall\ (x \in C)\ \neg\ (x < p)$.

Want to say: there is no point of P which is larger than every element of C. Which of the above says this? why?

I was too lazy in the class!

Justify why the other two statements do not say that?

33. Consider the set $X = [0,1] \times [0,1] - \{(0,0);(1,1)\}$ with dictionary order.

    Show that it has no first point; no last point; between two different points there is some thing strictly in between; every non-empty subset which is bounded above has a supremum; has no countable dense set.

    Is this order isomorphic to R?

    Suppose you considered all of $[0,1] \times [0,1]$ with dictionary order. Which of the above properties hold?

    Suppose you considered $R \times R$ with dictionary order. Which of the above properties hold?

    Suppose you consider $Z \times [0,1)$ with dictionary order. Which of the above properties hold?

34. Test which of the following properties hold for the losets given beow:
(a) has first element;
(b) has last element;
(c) has another point in between two different points;
(d) has countable dense set;
(e) every non-empty bounded subset has sup.

$$(i)\ \ S = [0,1] \qquad (ii)\ \ S = [0,1) \qquad (iii)\ \ S = (0,1]$$

$$(iv)\ \ S = (0,1] \cup [2,3); \qquad (v)\ \ S = Q$$

$$(vi)\ \ S = [0,1] \times [0,1] - \{(0,0),(1,1)\} \quad \text{with dictionary order.}$$

35. Verify the details left out in class: We defined multiplication xy for $x > 0$ and $y > 0$ and observed its properties. Extend the definition as follows:
$x < 0$ and $y > 0$: xy is defined as: $-[(-x)y]$.
$x > 0$ and $y < 0$: xy is defined as $-[x(-y)]$.
$x < 0$ and $y < 0$: xy is defined as $(-x)(-y)$.
$x = 0$ or $y = 0$: xy is defined as 0.
This definition satisfies all the properties required of multiplication, Show. Remember, you need not go to cuts. Use known things.

As you noticed probably, the connection of multiplication with order is only that product of two positive numbers is positive; and the rest (whatever) is a consequence of just this.

36. In our definition of R we have defined linear order which is friendly with addition and multiplication.

Instead, sometimes following is taken: There is a subset $P \subset R$ with the following properties.

(i) for all x; exactly one holds: $x = 0$, $x \in P$, $-x \in P$.

(ii) $x \in P, y \in P$ implies $x + y \in P$ and $xy \in P$.

Show that they and us are doing the same thing, that is, (i) starting with P as above define $\leq$ satisfying our conditions; conversely, (ii) starting from our $\leq$ exhibit P. show also that if you start with P, use (i) and then (ii) for that $\leq$; you get back your starting P.

Subsets of R are called 'unary' relations in R. Subsets of $R \times R$ are called binary relations in R — like $\{(x, y) : x \leq y\}$. Subsets of $R \times R \times R$ are called ternary relations in R— like $\{(x, y, z) : x + y = z\}$. In general you use the word, n-ary relation.

Did you realize how a binary operation like addition is also a relation, it is ternary relation.

37. Recall the definition of Cauchy sequence and definition of convergence of a sequence in R.

    Show that (x_n) is Cauchy iff for any given rational $r > 0$, there is an n_0 such that $|x_m - x_n| < r$ for all $m, n > n_0$. Show that (x_n) converges to x iff given rational $r > 0$ there is an n_0 such that $|x_n - x| < r$ for all $n > n_0$.

    Show that given any real number x, there is a Cauchy sequence (q_n) of rational numbers which converges to x.

    These two statements appear to be below our standard, prove them anyway. They are profound for the following reasons.

    (i) we need only rationals to define Cauchyness and convergence for sequences. $(*)$

    (ii) we need only Cauchy sequences of rational numbers to define real numbers. $(**)$

38. Consider the set of real numbers R. Recall the 'least upper bound axiom':

    (VI) Every bounded non-empty subset has supremum.

    Show that this can be replaced by the 'completeness axiom': $(* * *)$

    (VIa) Every Cauchy sequence converges.

    These three simple observations $(*), (**), (* * *)$ form basis for another construction of real number system due to Cantor. This method rescues us many times in Analysis. Shall do soon.

39. Here is an interesting field that builds on ultrafilter. usually algebraists are uninterested in this, they have field of rational functions which are not ordered fields. Analysts do not bother either, because many do not realize importance of this seed. Logicians have refined this idea as well as this example thoroughly (whatever this may mean).

    Let $\mathcal{U}$ be a free ultrafilter on $N = \{1, 2, \cdots\}$. Let R_0 be the set of all infinite sequences $s = (s_n)$ of real numbers. Here are some examples of sequences for you to see.

$$s_n \equiv \sqrt{555}; \quad t_n = 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \cdots + \frac{1}{n!}..$$

$$u_n = n; \qquad u_n^* = n^2; \qquad w_n = \frac{1}{n}; \qquad w_n^* = \frac{1}{n^2}.$$

Here $n = 1, 2, 3, \cdots$..

Let us define $s \sim t$ if $\{n : s_n = t_n\} \in \mathcal{U}$.

Show this is an equivalence relation.

Let the space of equivalence classes be denoted by R^*.

Remember we have addition and multiplication of sequences term by term. We can define these operations on R^*. Take x and y. These are equivalence classes. Take one sequence $s \in x$ and one $t \in y$ and define $x + y$ to be the equivalence class containing the sequence $s + t$. Show this is a good definition, that is, it does not depend on s and t as long as you choose from x and y.

Define multiplication in a similar way.

Show that we have a field.

Define order on R^* as follows. Say $x \leq y$ iff $\{n : s_n \leq t_n\} \in \mathcal{U}$ where $s \in x$ and $t \in y$. Show that this is a good(?) definition.

Show that R^* is a loset with this order.

Show we have an ordered field.

Identify usual set of real numbers as constant sequences. In other words, if $a \in R$ define $\varphi(a) \in R^*$ to be the equivalence class containing the constant sequence with each term equal to a.

Show that this is an embedding of $(R, +, \cdot, \leq)$.

Thus this ordered field R^* contains usual R.

This is the identification we use. Thus when we say the number $\sqrt[7]{33} \in R^*$ we mean the equivalence class containing the corresponding constant sequence.

Using the notation of the examples of sequences given at the beginning, let $[s]$ be the equivalence containing the sequence s.

Show $[u] > k$ for each $k \in R$. In some sense $[u]$ is an infinite integer.

For each $\epsilon > 0; \epsilon \in R$, show that $0 < [w] < \epsilon$. In some sense $[w]$ is an 'infinitesimal', strictly positive but smaller than all ϵ's *we know*.

So also is $[w^*]$.

Show the infinitesimal $[w^*]$ is smaller than the infinitesimal $[w]$.

We shall not discuss more about this field.

40. Suppose (X, d) is a metric space. Let us define

$$d_1(x, y) = \min\{d(x, y), 1\}.$$

Show that d_1 is also a metric.

Show that $x_n \to x$ in d iff $x_n \to x$ in d_1.

Show that a sequence is Cauchy in d iff it is Cauchy in d_1.

Show that a set is open in the metric d iff it is open in the metric d_1.

Show that the metric d_1 is bounded, whether d is bounded or not.

Suppose that in taking the minimum above, if I took minimum with 0.0001 how do the answers to above questions change?

Start with a metric space (X, d). Instead of the above, define

$$d_2(x, y) = \frac{d(x, y)}{1 + d(x, y)}.$$

Test whether all the above statements are still valid for d_2.

41. Consider the function $\tan x : (-\pi/2, +\pi/2) \to R$. It is strictly increasing, onto and one-one. Let $\tan^{-1}$ be the inverse map. Let us define

$$d_1(x, y) = |\tan^{-1} x - \tan^{-1} y|; \qquad x, y \in R.$$

Show that d_1 is a metric. Show that it is equivalent to d, that is, $x_n \to x$ in d iff $x_n \to x$ in d_1.

Do you think the following is correct: A sequence is cauchy in d iff it is Cauchy in d_1.

42. Given any open set U in $C[0, 1]$ with the sup metric, show that there is a polynomial with rational coefficients which belongs to U.

43. In a metric space show $|d(x, z) - d(y, z)| \leq d(x, y)$. More generally $|d(x, z) - d(y, u)| \leq d(x, y) + d(z, u)$.

44. Let $C[0,1]$ be the collection of real continuous functions on $[0,1]$. define

$$d_1(f,g) = \int_0^1 |f(x) - g(x)|dx.$$

$$d_2(f,g) = \left[\int_0^1 |f(x) - g(x)|^2 dx\right]^{1/2}.$$

Show that these are metrics. More generally, fix $p > 1$ and show that the following is a metric

$$d_p(f,g) = \left[\int_0^1 |f(x) - g(x)|^p dx\right]^{1/p}.$$

45. Let l_2 be the space of all (infinite) sequences $x = (x_n : n \geq 1)$ of real numbers such that $\sum |x_n|^2 < \infty$. Show that this is a linear space. Define

$$d_2(x,y) = \left[\sum |x_n - y_n|^2\right]^{1/2}.$$

Show that this is a metric.

Let l_1 be the space of all sequences $x = (x_n : n \geq 1)$ of real numbers such that $\sum |x_n| < \infty$. Show that this is a linear space. Define

$$d_1(x,y) = \sum |x_n - y_n|.$$

Show that this is a metric.

Give examples of sequences which are in l_2 but not in l_1. Do you think there are sequences which are in l_1 but not in l_2.

What happens if we consider sequences of complex numbers in both the above l_2 and l_1. Show that they are linear spaces and metric spaces.

More generally, consider for a fixed $p > 1$, the space l_p of all complex sequences $z = (z_n : n \geq 1)$ such that $\sum |z_n|^p < \infty$.

Show that this is a linear space and the following is a metric on the space.

$$d_p(z,w) = \left[\sum |z_n - w_n|^p\right]^{1/p}.$$

You can consider the space l_∞ also. It is the space of all bounded complex sequences. That is, the space of all sequences $z = (z_n : n \geq 1)$ with $\sup |z_n| < \infty$. Show that the following is a metric on this space.

$$d_\infty(z,w) = \sup_n |z_n - w_n|.$$

46. Consider

$$X = [0,1]^\infty = \{x = (x_1, x_2, \cdots) : x_i \in [0,1]; \;\; i = 1,2,3,\cdots.\}$$

$$d(x,y) = \sum \frac{|x_i - y_i|}{2^i}.$$

Show that d is a metric.

Show that $x^n \to x$ in the metric d iff $x_i^n \to x_i$ for each $i \geq 1$. That is, iff coordinate-wise convergence holds.

You can also take

$$X = R^\infty = \{x = (x_1, x_2, \cdots) : x_i \in R; \;\; i = 1,2,3,\cdots.\}$$

Show that if you define d exactly as above then the series may not converge. Define

$$d_1(x,y) = \sum \frac{1}{2^i} \frac{|x_i - y_i|}{1 + |x_i - y_i|}.$$

If you have done last part of exercise 43, you will not be frightened with this expression.

Show that this is a metric. Show that $x^n \to x$ iff $x_i^n \to x_i$ for each $i \geq 1$. That is, iff coordinate-wise convergence holds.

47. Let $X = Q$, the set of rational numbers.

We define, for $x \in Q$, order of x as follows: If x is an integer, then $O(x)$ is the largest power of 7 that divides x. If $x = a/b$ rational (with a, b integers), then $O(x) = O(a) - O(b)$. Show that this definition does not depend on how you represent the rational as a fraction.

For $x \in Q$, we define $||x||$ as follows:

$$||0|| = 0; \qquad ||x|| = 7^{-O(x)} \quad x \neq 0$$

note the negative sign for exponent.

Show $\qquad (i) ||x|| = 0$ iff $x = 0$.

$(ii) ||x + y|| \leq ||x|| + ||y||. \qquad (iii) ||xy|| = ||x|| ||y||.$

Define $d(x,y) = ||x - y||$. Show that this is a distance on Q. Actually this satisfies a better condition than triangle inequality.

Let $x \geq 1, y \geq 1$ be integers. Show $d(x,y) \leq 1/7^n$ iff $x = y \bmod(7^n)$.

Calculate for $n, m \in Z$,

(i) $d(7^n, 7^m)$ (ii) $d(7^{-n}, 7^{-m})$

Wonder: $7, 7^2, 7^3, 7^4, \cdots \to 0$.

Convince yourself that this metric takes (apart from zero) only the values

$$\cdots, \cdots \frac{1}{7^3}; \ \frac{1}{7^2}; \ \frac{1}{7}; \ 1; \ 7; \ 7^2, \ 7^3, \cdots\cdots.$$

Try to think which fellows are sitting exactly at a distance 1/7 from zero. Who are sitting at a distance 7 from zero.

48. Calculate interior, closure, closure of interior, interior of closure and so on of the following sets:

    (a) $X = R$.
    $A = C^c$ where C is cantor set in $[0, 1]$.
    $A = (0, 1) \cup (1, 2)$.
    $A = (0, 1) \cup (1, 2) \cup \{6, 7, 8, 9, \cdots.\}$
    $A = (0, 1) \cup (1, 2) \cup \{\text{rationals in } (3, 4)\} \cup \{6, 7, 8, 9, \cdots.\}$

    (b) $X = R^2$ and the sets are $A \times A$; take A to be each of the above sets.

    (c) $X = C[0, 1]$ and A is the set of all polynomials; or A is the set of all continuous functions taking values in $(0, 1)$; or A is the set of all continuous functions taking values in $[0, 1]$.

In what follows (X, d) is a metric space. A metric gives, distance between points, provides one measurement. You can use this to define several other measurements. Naturally, they depend on this metric.

49. For non-empty set $A \subset X$ and for $x \in X$ define

$$d(x, A) = \inf\{d(x, a) : a \in A\}.$$

In a sense, measures the least distance you need to travel from your location x to reach the town A.

    (a) $X = R$ usual metric $|x - y|$. Calculate $d(x, A)$ when
    $A = [0, 1]$. $x = 25$ or $x = -25$ or $x = 1/2$
    $A = (0, 1)$. x same as above.

    (b) $X = R^2$ Euclidean metric.
    $A = x_1$-axis. $x = (4, 5)$; or $x = (1, 0)$; or $x = (0, 1)$
    $A =$ the line $x_1 + x_2 = 4$. Same points as above.
    $A = B(0, 1/2)$. Same points as above.

    (c) $X = C[0, 1]$ with sup metric.
    A is the set of all functions $z \in X$ with $z(0) = 0$. x is the function $x(t) = t^2$; or $x(t) = \cos(2\pi t)$; or $x(t)$ is a polynomial in t.
    A is the set of all functions $z \in X$ with $z(0) = 0 = z(1)$. And x as above.

(d) $X = R$ with $d(x,y) = 0$ or 1 according as $x = y$ or not. This is called discrete metric. Calculate $d(x, A)$ for each $x \in R$ and each non-empty $A \subset R$.

(e) Show that $|d(x, A) - d(y, A)| \leq d(x, y)$.

(f) Show that x is in the closure of A iff $d(x, A) = 0$.

(g) How does $d(x, A)$ change when A is made larger?

(h) If A is compact, then show that the inf is actually minimum. Do you think that if the infimum is attained then A should be compact?

50. Suppose A and B are two non-empty sets. The distance between the sets is defined as

$$d(A, B) = \inf\{d(a, b) : a \in A, b \in B\}.$$

In a sense the smallest length of bridge needed to connect the two towns A and B. calculate $d(A, B)$ for the sets given below.

(a) $X = R$.
A is set of rationals in $[0, 1]$ and B is set of irrationals in $(2, 3)$.
$A =$ Cantor set in $[0, 1]$ and
$B = [4/27,\ 5/27] \cup (11/27,\ 16/27) \cup [64/81,\ 71/81]$.
$A = \{1, 2, 3, \cdots\}$ and
$B = \{1 + 1, 2 + (1/2), 3 + (1/3), 4 + (1/4), \cdots\}$

(b) $X = R^2$.
A is unit disc and $B = [2, 3] \times [2, 3]$
A is the x_1-axis and
$B = \{(x_1, x_2) : x_1 > 0, x_2 > 0, x_1 x_2 = 1\}$

(c) $X = R$ with discrete metric. Describe $d(A, B)$ for all pairs of nonempty subsets.

(d) Show that $d(A, B) \neq 0$ implies that $A \cap B = \emptyset$. Do you think the converse is true?

(e) Show $d(A, B) = \inf\{d(a, B) : a \in A\} = \inf\{d(b, A) : b \in B\}$.

(f) Show that distance between two sets is same as the distance between their closures.

(g) Do you think triangle inequality holds:
$d(A, C) \leq d(A, B) + d(B, C)$?

(h) If $A_1 \subset A; B_1 \subset B$, how do you compare $d(A, B)$ and $d(A_1, B_1)$.

(i) Show that if A and B are compact, then the inf is actually minimum. Do you think that if the inf is attained then at least one of the sets should be compact?

51. For a non-empty set A, we define the diameter of A by

$$\delta(A) = \sup\{d(x, y) : x, y \in A\}.$$

In a sense, measures how far are the farthest corners of the town A from each other. Calculate diameters of the sets below.

(a) $X = R$. A is the interval $[0, 1]$; or the Cantor set; or set of rational numbers in $(0, 1)$; or set of irrational numbers in $(0, 1)$; or the interval $(0, \infty)$.

(b) $X = R^2$. A is a line segment; or a triangle; or the open unit disc; or closed unit disc; or ellipse (with semi-major axis a and semi-minor axis b); or the unit square $[0, 1] \times [0, 1]$.

(c) $X = C[0, 1]$. A is the set of all x whose graph lies in the unit square; or set of all x whose graph lies in the unit disc.

(d) $x = R$ with discrete metric. A is the ball of radius $1/2$ around 5; or ball of radius $3/4$ around 5; or ball of radius 2 around 5; or ball of radius 5 around 5.

(e) Show that $\delta(B(a, r)) \leq 2r$. Do you think equality always holds? Show diameter of a set A is same as the diameter of its closure.

(f) Show that if A is compact then sup is actually maximum. Do you think that if the sup is attained then A should be compact?

52. A set $A \subset X$ is said to be bounded if its diameter is finite. Emptyset by convention, is bounded. But anyway, let us consider non-empty sets only.

Sometimes the following definition is used. A set A is bounded relative to a point $x \in X$ if $\sup\{d(x, a) : a \in A\} < \infty$, that is, there is a number M such that $d(x, a) \leq M$ for all $a \in A$.

If A is bounded relative to one point, show it is bounded relative to any other point. Show that this definition is same as saying that the diameter is finite.

Show that any Cauchy sequence is bounded.

Show that union of finitely many bounded sets is bounded.

Do you think every closed bounded set is compact?

Consider R^n with any of the l_p distances $1 \leq p \leq \infty$. Show that this definition of boundedness coincides with the notion of boundedness we adapted last year in R^n. recall $A \subset R^n$ is bounded if $||x|| \leq M$ for all $x \in A$.

53. Here is another metric spacce. Let $X = l_\infty$ be the space of all bounded sequences of real numbers. Define for $x, y \in X$.

$$d(x,y) = \sum_{n \geq 0} |x_n - y_n| e^{-5} 5^n \frac{1}{n!}.$$

This appears complicated, you would not understand whether the number $\exp\{-5\}$ is necessary at all. The essential point is: this is nothing but expectation of $|x-y|$ w.r.t. the Poisson probabilities you have learnt.

Show that this is a metric. do you think the space is complete?

54. We showed last year that $\sqrt{p}$, for prime $p > 1$, is irrational number. Thus $\sqrt{2}, \sqrt{3}, \sqrt{17}, \cdots$ are all irrational numbers. Of course there are many others. We should attend to some fillable gaps in our understanding.

(a) You can use the same reasoning to show that for any integer $k > 1$, either $\sqrt{k}$ is an integer or is irrational. Thus if $\sqrt{k}$ is rational then it must already be integer. Show this.

(b) Here is how you prove e is irrational.
Let if possible $e = a/b$ ratio of two positive integers.
Show then that

$$\frac{1}{b+1} + \frac{1}{(b+1)(b+2)} + \frac{1}{(b+1)(b+2)(b+3)} + \cdots$$

is an integer. Show $b > 1$ and the above can not be an integer.

(c) Here is how you prove π is irrational. (Proof by Ivan Niven)
Let if possible $\pi = a/b$ ratio of positive integers. Define, for a positive integer n, two functions

$$f(x) = \frac{x^n(a-bx)^n}{n!}$$

$$F(x) = f(x) - f^{(2)}(x) + f^{(4)}(x) - \cdots + (-1)^n f^{(2n)}(x).$$

Show

$$f(x) = f\left(\frac{a}{b} - x\right).$$

Show that the function $n!f$ has integer coefficients.
Show that f and its derivatives have integer values for $x = 0$ and $x = \pi = a/b$. Show

$$\frac{d}{dx}\{F'(x)\sin x - F(x)\cos x\} = F''(x)\sin x + F(x)\sin x = f(x)\sin x.$$

$$\int_0^\pi f(x)\sin x dx = F(\pi) + F(0) \quad \text{is an integer.}$$

$$0 < f(x)\sin x < \frac{\pi^n a^n}{n!}; \quad 0 < x < \pi.$$

Note that f depends on n. Argue that if you choose n large enough the above inequation leads to a contradiction.

(d) Actually the above numbers are transcendental, proof is not easy.

(e) We do not know if γ, Euler constant, is irrational.
When you read such a sentence, you should pause, recapitulate definition of γ and understand what is said; though there is nothing to prove in here.

55. Giving explicit examples of nowhere differentiable continuous functions is not easy. First such example was by Karl Weierstrass. Here is one from Rudin.

Define the function φ on R by $\varphi(x) = |x|$ for $-1 \leq x \leq 1$ and $\varphi(x+2) = \varphi(x)$ for all $x \in R$. Show $|\varphi(t) - \varphi(s)| \leq |t - s|$.

Define
$$f(x) = \sum_{n=0}^{\infty} \left(\frac{3}{4}\right)^n \varphi(4^n x).$$

Show f is continuous.

Fix any $x \in R$. Let
$$\delta_m = \pm\frac{1}{2}\,\frac{1}{4^m}.$$

Show you can choose the sign for δ_m so that there is no integer between $4^m x$ and $4^m(x + \delta_m)$. We fix this sign now. Having fixed m, set

$$\gamma_n = \frac{\varphi(4^n(x + \delta_m)) - \varphi(4^n x)}{\delta_m}.$$

Show that $\gamma_n = 0$ for $n > m$ while $|\gamma_n| \leq 4^n$ for $0 \leq n \leq m$. Show

$$\left|\frac{f(x + \delta_m) - f(x)}{\delta_m}\right| \geq 3^m - \sum_{0}^{m-1} 3^n \geq \frac{1}{2}(3^m + 1).$$

Show f is not differentiable at x.

56. Let (X, d) be a metric space.

(a) Consider a subset $C \subset X$ (always non-empty, in this context). Restrict d to pairs of points in C; still use d for this. then (C, d) is a metric space. Show.

(b) Let (X, d) be complete. If $C \subset X$ is closed then show that (C, d) is a complete metric space.
Conversely, let $C \subset X$ and (C, d) be complete. Then show that C is a closed subset of X.

(c) Let (X, d) be a complete metric space. Let $U \subset X$ be open set. ¿From the above you know that (U, d) is not complete unless U is closed.

However, without changing the notion of convergence, we can change the metric for U as follows. Keep in mind our examples of metrics on $(0, 1]$ and $(0, 1)$ discussed earlier.

First show that $f(x) = d(x, C)$ where $C = U^c$ is a continuous function and is never zero on U. Secondly, show that

$$\rho(x, y) = d(x, y) + \left| \frac{1}{d(x, U^c)} - \frac{1}{d(y, U^c)} \right|; \quad x, y \in U$$

or (not same, though same notation is used),

$$\rho(x, y) = d(x, y) + \min \left\{ \frac{1}{2^9}; \left| \frac{1}{d(x, U^c)} - \frac{1}{d(y, U^c)} \right| \right\}; \quad x, y \in U$$

is a metric on U. Thirdly, Show that $d(x_n, x) \to 0$ iff $\rho(x_n, x) \to 0$, as far as U is concerned. finally show that (U, ρ) is complete.

(d) Thus for example complement of finite sets in the real line can be given a 'complete metric' without changing notion of convergence.

(e) Let (X, d) be a complete metric space. Let $A = \cap U_n \subset X$ where each U_n is open. Thus A is intersection of countably many open sets. Put $d^*(x, y) = \min\{1, d(x, y)\}$. Put

$$\rho(x, y) = d(x, y) + \sum_n \min \left\{ \frac{1}{2^n}, \left| \frac{1}{d^*(x, U_n^c)} - \frac{1}{d^*(y, U_n^c)} \right| \right\}; \quad x, y \in A.$$

Show (A, ρ) is a complete metric space; d-convergence is same as ρ-convergence.

(f) The set of irrational numbers or the set of transcendental numbers in R is a possible candidate for above A.

(g) Read part (e) again. Following says we can not do any better.

Suppose $A \subset X$ and you can give a metric ρ for A such that the following two hold: (i) For points of A, ρ-convergence is same as d-convergence. (ii) (A, ρ) is complete metric space. Then A must indeed be countable intersection of open sets in X.

We shall not prove, involves some work (beautiful though).

57. Let f be any continuous function on $[0, 1]$. Let f_1 be any function whose derivative is f. Let, in general, f_n be any function whose derivative is

f_{n-1}. (Of course, given f_{n-1} you have freedom to choose your favourite constant in getting f_n; no more!)

Show the following: if $(\exists n)\ (\forall x) f_n(x) = 0$ then $f \equiv 0$.

Show the following: if $(\forall x)(\exists n) f_n(x) = 0$ then $f \equiv 0$.

58. Consider $C[0,1]$. A function is said to be everywhere oscillating if it is not monotone on any (non-degenerate) interval. Pause. Can you picture such a function? Very very difficult. If I did not put continuity, it is trivial to imagine!

    There are many many everywhere oscillating functions. More precisely, Complement of the set of everywhere oscillating functions, in $C[0,1]$, is countable union of small sets.

59. (Principle of uniform boundedness) Suppose that (X, d) is a complete metric space. Let F be a collection of real valued continuous functions on X.

    Assume that the collection is point-wise bounded. that is, given $x \in X$, the set $\{f(x) : f \in F\}$ is a bounded subset of R.

    Show that there is an (non-empty) open set U on which these functions are uniformly bounded. That is, $\{f(x) : x \in U; f \in F\}$ is a bounded subset of R.

60. Let (X, d) be a metric space.

    Recall $\overline{A}$ is closure of A and A^o is interior of A.

    Check: $\overline{A \cup B} = \overline{A} \cup \overline{B}$. Is $(A \cup B)^o = A^o \cup B^o$?

    Check $(A \cap B)^o = A^o \cap B^0$. Is $\overline{A \cap B} = \overline{A} \cap \overline{B}$?

    Show $(\overline{A})^c = (A^c)^0$. and $(A^o)^c = \overline{A^c}$

    Recall that closure of a set A consists of points which are in A or which are close to A — recall precise definition. Boundary of a set A is defined to be the set of all points which are close to A as well as A^c. Precise definition of boundary is $\partial A = \overline{A} \cap \overline{A^c}$.

    We needed and used this concept in discussing integration in several variables. But we considered only rectangles and regions within a nice closed curve.

    Show $\partial A = \overline{A} - A^0$.

    Show $\partial(A \cup B) \subset \partial A \cup \partial B$ and $\partial(A \cap B) \subset \partial A \cup \partial B$.

    Equality does not hold in general. Give examples.

    Do you think $\partial(A \cap B) \subset \partial A \cap \partial B$.

    Do you think $A \subset B$ implies $\partial A \subset \partial B$?

    What is ∂Q where Q is the set of rationals in R.

    What is boundary of Cantor set in R? What is boundary of complement of Cantor set in R? What is boundary of a disc in R^2? boundary of a circle in R^2 boundary of a rectangle in R^2?

    Do you think A and $\overline{A}$ and A^0 all have the same boundaries?

61. Consider $X = C[0, 1]$ with the metric $d(x, y) = \int_0^1 |x(t) - y(t)|dt$. Show the space is not complete? Do you think that convergence in this metric is same as convergence in the sup metric? If A is a non-empty subset which is open and closed in this space then it must be either $\emptyset$ or all of X. Show.

62. Many a times we start with a metric space but soon restrict attention to some subset of it. This happened, for example, in the proof that R is not union of infinitely many disjoint non-empty closed sets.

Let (X,d) be a metric space. Let $Y \subset X$. Keep the same metric on Y. Then (Y,d) is called a subspace of X.

For $a \in Y$, the ball $B^Y(a,r)$ is just $B^X(a,r) \cap Y$. Show.

A set $U \subset Y$ is open in Y iff there is a set $V \subset X$ which is open in X and $U = V \cap Y$. Similarly, a set $C \subset Y$ is closed in Y iff there is a set $F \subset X$ closed in X such that $C = F \cap Y$. Show.

Give an example where a proper subset of Y is open in Y but not open in X. Can you give such an example with Y open in X?

Give an example of a proper subset of Y which is closed in Y but not closed in X. Can you give such an example with Y closed in X?

Let $A \subset Y$. Show that the closure of A in Y is nothing but closure of A in X intersected with Y. That is,

$$\overline{A}^Y = \overline{A}^X \cap Y$$

$$\text{Is } (A^o)^Y = (A^o)^X \cap Y?$$

Give an example where a subset of Y is small in X but not small in Y. Can you give an example of a set $A \subset Y$ which is small in Y but not small in X?

In the following find $\overline{A}^Y$.
$X = R; \quad Y = Q; \quad A = \left\{ \sum_{k=0}^{n} \frac{1}{k!}; \ \ n \geq 1 \right\}$
$X = R^2; \quad Y = [0,1) \times (0,1]; \quad A$ is the set of all points (x,y) with each of x, y being either $1/n$ or $1-(1/n)$.

63. Recall that a metric space is connected if empty set and whole space are the only sets which are both closed and open. We say that $Y \subset X$ is connected if (Y,d) is connected.

Show that if $Y \subset X$ is connected then $\overline{Y}^X$ is also connected. Do you think the converse is true? That is $Y \subset X$ and $\overline{Y}^X$ is connected then Y is connected.

Consider U an open subset of R^2. Show that it is path connected, that is, given two points a and b in U, there is a continuous function γ on $[0,1]$ taking values in U such that $\gamma(0) = a$ and $\gamma(1) = b$.

Such a space is called path connected. More precisely, (X,d) is path connected if given any two points a and b in X there is a continuous function γ on $[0,1]$ such that $\gamma(0) = a$ and $\gamma(1) = b$.

Show that a path connected space is connected. Now do you see why R^{35} and $C[0,1]$ with sup metric or d_1 or d_2 or d_p metrics are connected? Show that l_2 (infinite square summable sequences) is connected.

Consider $X = R^2$. Consider the following set.

$$Y = \{(x, \sin(1/x)) : 0 < x \leq 1\} \cup \{(0, y) : 0 \leq y \leq 1\}.$$

Show Y is connected.

Show this Y is not path connected. Thus connected open sets are path connected but connected closed sets need not be path connected. This is also an example of connected space which is not path connected.

64. In R every non-empty open set is union, in a unique way, of countably many non-empty open intervals. Of course, in R^2 you can not talk of intervals.

Show the following: Every non-empty open set in R^2 is union, in a unique way, of countably many non-empty open connected sets.

65. Here is an application of Baire (a Diophantine approximation).

    Suppose that $\{t_n\}$ is a strictly increasing sequence of positive numbers increasing to ∞. This is given to us. Then there is a large set S of real numbers such that if you take $x \in S$, then the set

    $$\{t_n x + m : n \geq 1; m \in Z\}$$

    is dense in R. of course, we can not exactly specify for which numbers x the above set is dense. However, Kronecker tells that when $t_n = n$ then the above set is dense for every irrational number x.

66. Here is another application of Baire, similar to what we did for integrals. Suppose that f is a C^∞ function on $[0, 1]$.

    If $(\exists k)$ $(\forall x)$ $f^{(k)}(x) = 0$ then f is a polynomial. Show.

    If $(\forall x)$ $(\exists k)$ $f^{(k)}(x) = 0$ then f is a polynomial. Show (not easy).

    In particular, there is *one* k such that the k-th derivative vanishes identically.

67. We learnt Cantor intersection theorem; Baire's theorem; Banach fixed point theorem in complete metric spaces. but we have not shown that some of the standard spaces we saw are actually complete. Yes, we know R is complete.

    (a) Show that R^n is complete.

    (b) Recall that R^∞ is the space of all infinite sequences $x = (x_n)$ of real numbers. The metric is (exercise 46)

    $$d_1(x, y) = \sum \frac{1}{2^i} \frac{|x_i - y_i|}{1 + |x_i - y_i|}$$

    according to which the convergence is again coordinate-wise convergence.

    Show that this space is complete. Take a Cauchy sequence $\{x^n\}$ where

    $$x^n = (x_1^n, x_2^n, x_3^n, \cdots, x_j^n, \cdots\}$$

    Show that for every j the sequence of numbers $\{x_j^n : n \geq 1\}$ is a Cauchy sequence. Let its limit be x_j.

    Let $x = (x_j)$. Show that $x^n \to x$ in the space R^∞.

(c) Show $C[0,1]$ with sup metric is complete. Take a Cauchy sequence $\{x_n\}$ — remember this is Cauchy in sup metric.

show that for each t the sequence of numbers $\{x_n(t)\}$ is cauchy.

Show that there is a function x such that $x_n(t) \to x(t)$ for each t.

Fix $\epsilon > 0$. Fix N (using hypothesis) so that

$$\sup\{|x_n(t) - x_m(t)|; 0 \le t \le 1\} < \epsilon/2; \quad n, m \ge N$$

Show that $|x_n(t) - x(t)| \le \epsilon/2$ for each $n \ge N$ and each t. Conclude that

$$\sup\{|x_n(t) - x(t)|; 0 \le t \le 1\} \le \epsilon/2; \quad n \ge N$$

Conclude that (x_n) converges to x uniformly. Conclude that x is continuous and hence $x \in C[0,1]$.

Show that $d(x_n, x) \to 0$.

(d) Recall that l_2 is the space of all sequences $x = (x_j)$ of real numbers such that $\sum x_j^2 < \infty$. Recall

$$d(x,y) = \sqrt{\sum (x_j - y_j)^2}.$$

Show that this space is complete. Take Cauchy sequence (x^n) where x^n is the sequence

$$x^n = (x_1^n, x_2^n, x_3^n, \cdots, x_j^n, \cdots).$$

Show that for very j, the sequence of numbers $\{x_j^n : n \ge 1\}$ is a Cauchy sequence. Let x_j be its limit.

Using the fact that Cauchy sequences are bounded (and Minkowski) conclude that there is a number K such that

$$\sum_j (x_j^n)^2 \le K; \quad n \ge 1$$

Show that $\sum_{j=1}^{J} x_j^2 \le K$ for every $J \ge 1$. Conclude that $x \in l_2$.

Fix $\epsilon > 0$. Fix N so that $d(x^n, x^m) < \epsilon/2$ for $n, m \ge N$. Show that for each $J \ge 1$

$$\sqrt{\sum_{j=1}^{J} (x_j^n - x_j)^2} \le \epsilon/2; \quad n \ge N.$$

Conclude that $d(x^n, x) < \epsilon$ for $n \geq N$; or $x^n \to x$ in l_2.

You should appreciate the ease with which we try an inequality for finite sums (limits can be taken for finite sums) and then conclude the same for infinite sums. pause and understand this sentence.

(e) Test whether you understood the above two calculations by showing that the space l_1 is complete. Recall this is the space of all sequences $x = (x_j)$ with $\sum |x_j| < \infty$ and

$$d(x, y) = \sum_j |x_j - y_j|.$$

It is also true that all l_p spaces $(p > 1)$ are complete but you need not bother about it now (not for exam either). That will be for a later functional analysis course, not now.

68. Let (X, d) be a metric space. Recall that a subset $D \subset X$ is dense if every non-empty open set contains a point of D, or equivalently, every ball contains a point of D. A metric space is separable if there is a countable dense set.

You know that Q is dense in R and that Q is countable.

(a) Show that the set Q^n; the set of n-tuples of ration numbers is dense in R^n.

Show that this set Q^n is countable.

(b) Recall that R^∞ is the space of all infinite sequences $x = (x_n)$ of real numbers. The metric is (exercise 46)

$$d_1(x, y) = \sum \frac{1}{2^i} \frac{|x_i - y_i|}{1 + |x_i - y_i|}$$

according to which the convergence is again coordinate-wise convergence.

Let D be the set of all sequences with finitely many non-zero terms and those terms are rationals. Show that D is dense in R^∞.

Show that D is countable.

(c) Show that the set P of polynomials with rational coefficients is dense in $C[0, 1]$.

Show that this set P is countable.

If you are fed up of polynomials, here is another set. For rational numbers $0 = t_0 < t_1 < t_2 < \cdots < t_k = 1$ and rational numbers

$r_0, r_1, \cdots, r_k$ let x be the function whose graph consists of straight lines joining (t_i, r_i) and (t_{i+1}, r_{i+1}) for $i = 0, 1, \cdots, k-1$. Let D be the set of all these function as $k \geq 1$ varies over integers and t_i and r_i vary over rationals as stated. Show that D is dense. Show D is a countable set.

(d) Consider l_2. Let D be the set of all sequences $x = (x_n)$ which have only finitely many non-zero terms and those terms are rationals. Show that this set is dense in l_2.

Show that D is countable.

(e) Show that the same set as above is dense in l_1.

69. For any set A, let $l(A)$ denote its set of limit points. $N = \{1, 2, 3, \cdots.\}$

(a) Let $A = \{(i, 0, 0, 0) : i \in N\} \subset R^4$. Then A is closed and $l(A) = \emptyset$.

(b) Let $B = A \cup \{(i, \frac{1}{j}, 0, 0) : i, j \in N\} \subset R^4$ Then B is closed, $l(B) \neq \emptyset$ but $l(l(B)) = \emptyset$.

(c) Let $C = B \cup \{(i, \frac{1}{j}, \frac{1}{k}, 0) : i, j, k \in N\} \subset R^4$. Then C is closed, $l(C), l(l(C))$ are non-empty but $l(l(l(C)))$ is empty.

(d) Let $D = C \cup \{(i, \frac{1}{j}, \frac{1}{k}, \frac{1}{m}) : i, j, k, m \in N\} \subset R^4$. Guess and show.

(e) It is not difficult to get similar sets in R itself. Try, if you like challenges.

****

The most important scientific tool of all is not anything you can buy. It is your own mind. Your thoughts and ideas are the keys that can unlock the mysteries. In the search for understanding, *questions* are perhaps the most powerful *force* of all.

– "The hidden world of forces" by Jack R White

70. I am not sure if you understood the concept of ball in a metric space. This is crucial. I have plotted some of them in class.

    (a) Consider R^2 with the Euclidean metric. Plot the unit ball around the origin. Plot ball of radius $1/2$ around the point $(1, 3)$. Plot in each case closed ball as well as open ball.

    (b) Do the same when the metric is $d(x, y) = |x_1 - y_1| + |x_2 - y_2|$.

    (c) Do the same when the metric is $d(x, y) = \max\{|x_1 - y_1|, |x_2 - y_2|\}$.

    (d) Do the same when the metric is $d(x, y) = \sqrt[3]{|x_1 - y_1|^3 + |x_2 - y_2|^3}$.

    (e) Consider the real line R with the metric $d(x, y) = \min\{|x - y|, 1\}$ Plot ball of radius $r = 1/2$ with centre $x = 13$. Plot ball of radius $r = 1$ same centre. Plot ball of radius $r = 1.001$ with same centre. What if I said plot ball of same radius $r = 1.001$ with centre $x = 31$.

    (f) Consider the open unit interval $X = (0, 1)$ with metric $d(x, y) = |x - y|$. Is the set $B = \{x : 1/3 < x < 1/2\}$, that is, the interval $(1/3, 1/2)$ a ball? If so what is its centre and radius?

        Is the set $(0.99, 1)$ a ball? There are several centres and radii posible. Think and discover.

    (g) Let X be a finite set and d be *any* metric on X. Show that every singleton set is a ball. What is the centre and radius?

    (h) Consider the ellipse $B = \left\{(x_1, x_2) : \frac{1}{4}x_1^2 + \frac{1}{9}x_2^2 < 1\right\}$. Can you give a metric for R^2 without changing convergence so that the above ellipse is ball of radius one with centre $(0, 0)$?

    (i) Take a $k \times k$ positive definite symmetric matrix. Show that
        $d(x, y) = \sqrt{(x - y)^t A(x - y)} = \sqrt{\sum a_{ij}(x_i - y_i)(x_j - y_j)}$.
        is a metric on R^k and convergence in d is usual one.

71. Just as we have $B(x, r) = \{y : d(x, y) < r\}$ ball around a point we also have ball around a set. For a non-empty set $S \subset X$,

$B(S,r) = \{y : (\exists\, x \in S)\ \ d(x,y) < r\} = \bigcup_{x \in S} B(x,r)$.

If K is a compact and U an open set such that $K \subset U$ show that there is an $r > 0$ such that $B(K,r) \subset U$.

72. Let K be a compact and C a closed set in (X,d). If $K \cap C = \emptyset$ then show $d(K,C) > 0$. Recall $d(A,B) = \inf\{d(x,y) : x \in A;\ y \in B\}$.

    Do you think the statement remains true if both sets are closed?

73. S_n is a decreasing sequence of non-empty compact sets in a metric space. Show $\cap S_n$ is non-empty.

    Do you think this remains true if the sets are just closed instead being compact? What if all are closed and one is known to be compact?

74. Let (X,d) be a compact metric space. Show:

    (*) If $\mathcal{C}$ is a family of closed sets such that intersection of any finitely many of them is non-empty then there is a point common to all of them (in other words, intersection of all of them is non-empty) .

    Conversely show: a metric space satisfying (*) is compact.

75. If every real valued continuous function is bounded, show that the space is compact. All this drama is on a metric space, a nice stage.

    If every real valued bounded continuous function attains its supremum then show that such a function attains its infimum too. Show that when this happens the space is compact.

76. X is a compact metric space and f_n is a sequence of real continuous functions decreasing to a continuous function f point wise. Show that the convergence is uniform. This is known as Dini's theorem and we did it last year for $X = [0,1]$

77. Describe *all* closed additive subgroups of R. Describe all open additive subgroups of R.

E $\exists$ A $\forall$

is it odd how asymmetrical is symmetry
symmetry is asymmetrical how odd it is.

78. Here are some examples of useful metrics on some useful spaces. Of course, others we discussed earlier are useful too!

(a) Consider a (finite) set S and an integer $N \geq 1$. Consider $X = S^N$, sequences (called strings) of length N from the set S. denote $x = (x_1, \cdots, x_N)$ for points in X. Put

$$d(x, y) = |\{i : x_i \neq y_i\}|,$$

that is, the number of places where the two strings differ. Verify this is a metric.

This distance tells you how many places need to be changed to transform one sequence to the other. This is called *Hamming distance* and is useful in coding theory and DNA analysis.

Obviously the larger the distance the better chances of decoding them correctly; they can 'withstand' one or two transmission errors. Or, a larger distance between DNA strings of two animals will hint that they are different species; other interpretations are also possible.

(b) Let $X = R^N$ and Σ be a symmetric positive definite matrix of size $N \times N$. Define

$$d(x, y) = \sqrt{(x - y)^t \Sigma^{-1} (x - y)}.$$

show that this is a metric on R^N.

This is sometimes called *Mahalanobis distance* and is useful in Statistics.

The points x and y are observations (may be vector consisting of nose length, cheek bone size, forehead width; scalp measurement etc of some skull). We have two observations on two skulls we found. Want to know if they belong to the same tribe. Here Σ comes from an assumed probabilistic model; appearance of its inverse seems mysterious but shall not get into details.

(c) Let X be the set of permutations of $S = \{1, 2, 3, \cdots, N\}$. For two permutations π and η define

$$d(\pi, \eta) = \left| \left\{ (i,j) : i < j; \begin{array}{lll} \pi(i) < \pi(j) & \& & \eta(i) > \eta(j) \quad \text{OR} \\ \pi(i) > \pi(j) & \& & \eta(i) < \eta(j) \end{array} \right\} \right|$$

that is; the number of pairs which are compared by π and η differently.

Show that this is a distance. This is called *Kendall's tau* and is useful in statistics. Its largest value is $N(N-1)/2$. When is this value achieved?

Suppose you and I rank (no ties) fifty selected hotels in Chennai. How different are our rankings? Above distance is one such measure.

(d) Let X be the set of vectors of length N consisting of non-negative numbers whose sum equals one.

$$X = \{(p_1, \cdots, p_N) : p_i \geq 0 \ \ \forall\, i; \ \sum p_i = 1\}.$$

that is, all probability vectors. equivalently, all possible probability models for an experiment which has N outcomes. Define

$$d^2(p,q) = \frac{1}{2} \sum_1^N (\sqrt{p_i} - \sqrt{q_i})^2 = 1 - \sum \sqrt{p_i q_i}.$$

Then d is a distance, called *Hellinger or Bhattacharya* distance. This is useful in functional analysis (Hellinger) and in Statistics (Bhattacharya).

How different are two models? Above is a measure.

(e) Consider $S = \{0, 1, 2, 3, \cdots\}$ and $X = S^N$. The l_1 distance on X, namely,

$$d(x,y) = \sum |x_i - y_i|$$

is also called *block distance*. This is because of the following reason. Imagine houses are located at points of X. roads are laid only along lines parallel to the axes. Thus when you travel you can not go diagonally and so on. To go from house x to y this is the distance you need to travel. Think about it.

79. Here are some examples of metric spaces which are important when you study groups. Do not get panicky. For us these are just some routine examples of subsets of euclidean space.

(a) Let X be R^{400} and M_{20} be the space of 20×20 matrices. We think of them as same in the following way: Given a vector in R^{400}, we break it into 20 consecutive segments each of length 20 and these form the rows of a matrix in M_{20}.

Verify that this is an identification of the two sets(?). Bring the metric of R^{400} to M_{20}. Show matrix multiplication is continuous. (From where to where?)

(b) Show that the set of invertible matrices GL_{20} is an open subset of M_{20}. Show that matrix inverse is a continuous map of GL_{20} to itself. In other words, the group operations are friendly with convergence.

Since this is an open subset of a complete space you can treat it as 'complete' space. Is this connected?

(c) Let SL_{20} be the subgroup of GL_{20} which consists of all matrices A with $|det(A)| = 1$, that is, $det(A)$ equals ± 1. show that this is a closed subset of SL_{20}. Since this is a closed subset, you can treat it as a complete metric space. Is this connected?

What if I considered SL_{20}^{+}, matrices with determinant one. Is it closed? Is it connected?

(d) Suppose I consider the set Sym of symmetric matrices. Is this closed subset of M_{20}? Is this connected?

(e) If I considered P^{+} the set of symmetric positive definite matrices. is it open or closed? Is it connected?

(f) What if I considered normal matrices, that is, matrices with $AA^t = A^tA$. What kind of subset is it?

(g) If I considered the space O_{20} of orthogonal matrices, that is, matrices with $AA^t = A^tA = I$, identity matrix. What kind of subsets is it?

Remember: you can think and solve these problems. You should.

80. Let Q be any bounded closed rectangle contained in R^2. Show that the set of polynomials in x, y is dense in $C(Q)$.

    Generalize to R^k.

81. Here is a non-trivial application of the Stone-Weierstrass theorem. Let (X, d) be a compact metric space. We already know that $C(X)$ the space of real continuous functions on X with sup metric is a complete metric space. Goal: to show that $C(X)$ is separable.

    Thus $C(X)$ will be a Polish space.

    (a) Suppose we could exhibit a sequence of functions $\{f_1, f_2, \cdots\}$ in $C(X)$ which separate points.

    Let $f_0 \equiv 1$ and $D_0 = \{f_n : n \geq 0\}$. Let D_1 be the collection of finite products of functions in D_0. Let D be finite rational linear combinations of functions from D_1.

    Show D is countable. Show D is closed under multiplication.

    Let $\overline{D}$ be closure of D in our space $C(X)$. show that $\overline{D}$ is a vector space; is an algebra. equals $C(X)$. Deduce $C(X)$ is separable.

    (b) Let $p \in X$ and $\epsilon > 0$. Put $f(x) = d(x, p)$. Put $g(x) = \min\{f(x), \epsilon\}$. put $h(x) = g/\epsilon$. Show these are continuous functions.

    (c) Given any open ball $B(p, \epsilon)$ show that there is a continuous functions f with $f(p) = 1$ and $0 < f \leq 1$ on B and $f(B^c) = 0$.

    (d) (done in class) show that there is a *sequence of open balls* such that every open set is a union of some of these balls.

    (e) Show that there is a sequence of functions as required in the first step.

82. Let X be a metric space which is not compact. Goal: to show that the space $C_b(X)$ of bounded real continuous functions with sup metric is not separable.

    (a) Show that the space $C_b(X)$ is complete; not needed for discussing separability but good to know such basic facts; at no extra cost.

(b) Let us consider R. Let $N = \{1, 2, 3, 4, \cdots\}$. For any subset $A \subset N$, show that there is a bounded continuous function f_A on R such that $f_A(x) = 1$ for each $x \in A$ and $f_A(x) = 0$ for each $x \in N - A$. show that $C_b(R)$ is not separable.

(c) Back to metric space. Let A and B be two disjoint non-empty closed subsets of the metric space X. Show

$$f(x) = \frac{d(x, A)}{d(x, A) + d(x, B)}$$

is a real bounded continuous function on X.

(d) If X is not compact, there is a sequence which has no convergent subsequence. Show we can take the sequence to consist of distinct points. show every subset of the sequence (?) is a closed subset of X.

(e) Show $C(X)$ is not separable.

83. Let $K(x, y)$ be a continuous function on $[0, 1] \times [0, 1]$. Let us define a map on $C[0, 1]$ by

$$Tx(s) = \int_0^1 K(s, t)x(t)dt.$$

Show T takes the space to itself, in other words, if x is continuous then so is Tx (done last semester, recall proof).

Show: if $\{x_n\}$ is a sequence in $C[0, 1]$ which is uniformly bounded, then the sequence $\{Tx_n : n \geq 1\}$ is precompact.

84. Suppose that $A \subset C[0, 1]$ is a collection of differentiable functions. Suppose that there is a number M such that for all $x \in A$ and for all $t \in [0, 1]$

$$|x(t)| + |x'(t)| \leq M$$

Show that A is precompact.

85. Here is an application of Arzela-Ascoli. This is called Peano's theorem.

We are given an open set $\Omega \subset R^2$ and a point $(t_0, x_0) \in \Omega$ and a continuous function $F : \Omega \to R$. Goal is to show that we can exhibit an interval around t_0 and a differentiable function x on that interval such that (i) $x(t_0) = x_0$ and (ii) $x'(t) = F(t, x(t))$. The last condition means that for each t in the interval exhibited, the point $(t, x(t)) \in \Omega$ and the stated equality holds.

get a closed ball B around (t_0, x_0) contained in Ω and let $|F| < M$ for all points in this ball. Take $\delta > 0$ so that

$$[t_0 - \delta, t_0 + \delta] \times [x_0 - M\delta, x_0 + M\delta] \subset B.$$

Argue that there is a sequence of polynomials $P_n(x, y)$ such that

$$d(P_n, f) \to 0; \quad |P_n(x, y)| < M \;\; \forall (x, y) \in B$$

Show that there is a solution $x_n(t)$ for (i), (ii) with data (t_0, x_0, P_n).

Write the integral equation satisfied by x_n.

Show there is a subsequence of $\{x_n\}$ converging to, say, x^*.

show that x^* is a solution of original problem.

86. consider R^2 and $F(t, x) = x^{2/3}$ and $(t_0, x_0) = (0, 0)$. Then $x(t) \equiv 0$ and $x^*(t) = t^3/(27)$ both solve $x' = F(t, x(t))$.

87. Let C be a closed subset of R. Show that we can express

$$C = D \cup P; \quad D \text{ countable}; P \text{ closed}; \quad P = \emptyset \text{ or } l(P) = P$$

Recall that $l(P)$ is the set of limit points of P and so the last phrase means that every point of P is a limit point of P. [take D = union of $(a, b) \cap C$ which are countable].

Let P be a closed set for which every point is a limit point. Then P must have same power as c — if non-empty.

You Prove it just like Cantor intersection theorem. Get two disjoint closed balls B_0 and B_1. within each get two disjoint closed balls B_{00}, B_{01} and B_{10}, B_{11}. Stare at these.

Open your thoughts first, then your pen.

88. The algebra generated by $\{1, x^2\}$ is dense in $C[0,1]$. Is it dense in $C[-1,1]$? Suppose I take an integer $k \geq 1$. For what I is the algebra generated by $\{1, x^k\}$ dense in $C(I)$? Of course, I here is a closed bounded interval.

89. f is real continuous on $[0,1]$ and $\int_0^1 f(x)x^n dx = 0$ for integers $n \geq 0$ show $f(x) \equiv 0$.

90.]it Unit balls

    Remember the unit ball in R^k is the set of all vectors x such that $d(x,0) = ||x|| \leq 1$. We know it is compact.

    Analogously, unit ball in $C[0,1]$ is the set of all functions $x = x(t)$ such that $d(x,0) \leq 1$. This means the set of all x with $\sup |x(t)| \leq 1$. Show that this is not compact.

    Similarly the unit ball in l_2 is the set of all sequences $x = (x_n : n \geq 1)$ such that $d_2(x,0) \leq 1$. This means $\sum x_n^2 \leq 1$ Show this is not compact.

    However show that the following set is compact in l_2. All points $x = (x_n)$ with $|x_n|, \leq 1/n$ for all n.

91. If a metric space is totally bounded then show that every sequence contains a subsequence which is Cauchy.

    Thus, duty of totally boundedness is to provide a Cauchy subsequence; duty of completeness is to make it converge; together give compactness.

92. *compact subsets of* R^∞

    The diagonal argument has great potential.

    Remember we made R^∞ the space of sequences of real numbers into a metric space. We gave a metric to the space. Of course convergence is just coordinatewise convergence. Show that a subset $K \subset R^\infty$ is pre compact iff for each n, there is number M_n such that
    $x = (x_n) \in K \Rightarrow |x_n| \leq M_n \ \forall n$.

93. Here is a further generalization of the space $C(X)$ we considered.

    Let X be a compact metric space and Y be a Polish space. Let $C[X,Y]$ be the space of all continuous functions on X with values in Y.

Let $X = [0,1]$ and Y is Cantor set, describe $C[X,Y]$.

Let $X = [0,1]$ and Y the set consisting of the vertical lines at $x = 2$ and $x = 4$ in the plane. Describe $C[X,Y]$. Are you able to imagine it as two copies of $C[0,1]$ sitting side by side.

Let $X = [0,1]$ and Y be the set of complex numbers. Describe $C[x,Y]$ are you able to imagine it as $C[0,1] \times C[0,1]$?

Let $X = [0,1]$ and $Y = C[0,1]$. Identify $C[X,Y]$ with $C([0,1] \times [0,1])$. What does identify mean here?

back to generalities.

Show $f \in C$ implies range of f is a bounded subset of Y.

Show $f \in C$ implies it is uniformly continuous. This means, given $\epsilon > 0$, there is $\delta > 0$ such that

$$d_X(s,t) < \delta \Rightarrow d_Y(f(s), f(t)) < \epsilon.$$

Define for $f, g \in C[X,Y]$;

$$d(f,g) = \sup\{d_Y(f(s), g(s)) : s \in X\}.$$

Show that this is a metric and $C[X,Y]$ is a complete metric space.

94. *Hausdorff metric*

The setup here appears abstract but not the mathematics. I mention to impress upon you the diversity of metric spaces you can think of.

Consider $I = [0,1]$ and let Γ be the collection of all non-empty compact subsets of I. Given two sets K and L in this space show that there is at least one $\epsilon > 0$ such that $K \subset B(L,\epsilon)$ and $L \subset B(K,\epsilon)$.

Hausdorff defines the distance between two sets as the smallest such ϵ. That is,

$$\rho(K,L) = \inf\{\epsilon > 0 : K \subset B(L,\epsilon);\ L \subset B(K,\epsilon)\}.$$

Suppose K and L are singletons $\{x\}$ and $\{y\}$ guess what should be their distance and verify.

Suppose K is a doubleton and L is a singleton. Then what is their distance? What if both are doubletons?

What is the distance between $[0,1/3]$ and $[2/3,1]$. Guess first and proceed.

Show that ρ is actually distance on Γ.

It is not difficult to show that the space Γ is compact, but let us not spend time on it. Remember, in connection with fixed point theorem we came across this space.

95. *Stone-Weierstrass, Complex version*

Let X be a compact metric space and $C[X, C]$ be the space of complex valued continuous functions on X. Let A be a sub algebra of this which contains constant functions and separates points. Suppose moreover that $f \in A$ implies its conjugate $\overline{f} \in A$. Show that A is dense in the space. (Reduce the problem to real case).

Remember if $f(x) = f_1(x) + if_2(x)$ where f_1 and f_2 are real valued then conjugate of f is defined by $\overline{f}(x) = f_1(x) - if_2(x)$.

96. *identifying points*

I said that if you tie the two points zero and one together, the unit interval becomes unit circle. This is actually a precise statement. Let us see what it means.

Let $X = [0, 1]$ usual metric. Let $Y = (0, 1) \cup \{\spadesuit\}$ Thus Y has all points of X except zero and one, instead it has one extra point. This is the bag containing zero and one. Here is the metric $d^*(s, t)$: it is same as $d(s, t)$ when $s, t \in (0, 1)$; if both are the extra point then distance is zero; if $s \in (0, 1)$ then $d^*(s, \spadesuit) = \min\{|s - 0|, |s - 1|\}$.

Show that this is a metric. Show this space is homeomorphic to the circle via the map $h(t) = (\cos 2\pi t, \sin 2\pi t)$. Need not worry how to interpret the value at $\spadesuit$, take it as zero or one; does not matter.

later you will see several such constructions.

Chennai Mathematical Institute

24-09-2014 Midsemester examination Duration: Two hours

BSc Second year Real Analysis B. V. Rao

Justify you statements. Organize your proofs and then write, stating clearly any results you use.

1. Calculate $\limsup A_n$ and $\liminf A_n$ for the sets defined below.
   $A_n = [0,1] \cup [n, \infty)$ if n is odd; and
   $A_n = [-1,0] \cup (-\infty, -n]$ if n is even. [5]

2. Exhibit uncountable family of non-empty subsets of $\{1, 2, 3, \cdots\}$ such that intersection of any two distinct sets of the family is finite. [5]

3. Consider the following linearly ordered set.
   $X = Z \times R$ with dictionary order where Z, set of integers; and R, the set of real numbers have usual order.
   Show that X is order isomorphic to a subset of R.
   Show that X is not order isomorphic to R. [10]

4. X is a non-empty set and $f_1, f_2, \cdots$ is a sequence of functions on X to $[0,1]$, Define

$$d(x,y) = \sum \frac{1}{2^i} |f_i(x) - f_i(y)|; \qquad x, y \in X$$

   which axioms of metric does this satisfy?
   Let (x_n) be a sequence. Show that $d(x_n, x) \to 0$ iff $f_i(x_n) \to f_i(x)$ for each i. [10]

5. Consider the space $C[0,1]$ with sup metric. If $A \subset C[0,1]$ and A is both closed and open; show that it must be either $\emptyset$ or $C[0,1]$. [10]

GOOD LUCK

Chennai Mathematical Institute

26-11-2014 Semestral examination Duration: $2\frac{1}{2}$ hours

BSc Second year Real Analysis B. V. Rao

Justify you statements. Organize your proofs and then write, stating clearly any results you use. Each question carries 10 marks.

1. Let $N = \{1, 2, 3, 4, \cdots\cdots\}$, the set of all positive integers.

   Explain with reasons whether the following sets are countable or uncountable.

   (a) set of all functions from $\{1, 2, 3, 4\}$ to N.

   (b) set of all functions from N to $\{1, 2, 3, 4\}$.

2. Explain with reasons whether the two sets S and T (with usual order) are order isomorphic.

   (a) $S = R$; $\quad T = (0, 1)$.

   (b) $S =$ set of all rationals in $(0, 1)$;

   $T =$ set of all rationals in: $\quad (0, 1) \cup [e, 2e]$.

3. Explain with reasons whether the following subsets of R are compact.

   (a) $S = \{8x^2 \sin(y + x) : \quad x \in [0, 1], \ \ y \in [0, \infty)\}$.

   (b) $S = \{x^5 y^8 : \quad 0 < x \leq 8; \ \ 0 \leq y \leq 50\}$.

4. (a) Let f be a continuous real valued function on $[0, 1] \times [0, 1]$. For each $x \in [0, 1]$, let f_x be defined on $[0, 1]$ by $f_x(y) = f(x, y)$. Consider the map $Tx = f_x$ from $[0, 1]$ to $C[0, 1]$. Show that this is continuous. Here $C[0, 1]$ has sup metric.

   (b) Explain with reasons whether the following statement is correct:

   $f_n \to f$ uniformly on $[0, 1]$ and f_n is C^1 for each n, implies f is also C^1.

5. Let (X, d) be a metric space.

   (a) Let $A \subset X$ be non-empty subset. Define $d(x, A)$ and show that it is uniformly continuous function.

   (b) Given two disjoint closed sets A and B show that there is a real valued continuous function f on X such that $f(x) = 0$ for $x \in A$ and $f(x) = 1$ for $x \in B$ and $0 \leq f(x) \leq 1$ for all x.

6. (a) Calculate all the Fourier coefficients of the following function.

$$f(x) = \sin^2(2\pi x) + \cos(10\pi x); \qquad 0 \leq x \leq 1.$$

(b) Let (X, d) and (Y, ρ) be metric spaces. Consider the metric space $X \times Y$ with metric:

$$\mu((x_1, y_1), (x_2, y_2)) = d(x_1, x_2) + \rho(y_1, y_2).$$

If X and Y are connected, show that $X \times Y$ is also connected.

GOOD LUCK