Kerala School of Mathematics Kozhikode.
Workshop on Set Theory and Logic Feb, 18 – Mar, 01; 2013.
B. V. Rao Chennai Mathematical Institute.

Posets and Axiom of Choice

In these lectures we shall discuss certain aspects of partially ordered sets, called posets; linearly ordered sets, called losets and the Axiom of Choice.

1. POSETS:

Let X be a non-empty set and R be a binary relation on X. This means that R is a subset of $X \times X$. If $(x, y) \in R$, we write xRy. Otherwise we say $\neg(xRy)$. The relation R is said to be a *partial order* if
(i) $\forall\, x, y, z(xRy, yRz \Rightarrow xRz)$; $(ii) \forall\, x \neg(xRx)$.

A set with a partial order is called a *partially ordered set* or simply a *poset*. In place of R, one uses a suggestive notation $<$. Thus in place of xRy, we write $x < y$. With this change of notation the above two conditions can be expressed as follows.
(i) $\forall\, x, y, z(x < y, y < z \Rightarrow x < z)$; $(ii) \forall\, x \neg(x < x)$.

The condition (i) is called transitivity, and rule (ii) is called antireflexivity. If we have a partial order, sometimes one defines the relation $\leq$ as follows. Say $x \leq y$ if $x < y$ or $x = y$. Then this relation obeys the following three rules.
(a) $\forall\, x, y, z(x \leq y, y \leq z \Rightarrow x \leq z)$.
This is directly verified in all cases. If $x < y$ and $y < z$, then by (i) we get $x < z$ and hence $x \leq z$. If $x = y$ and $y = z$, then $x = z$ so $x \leq z$. Other cases are similar.
(b) $\forall\, x(x \leq x)$.
Since $x = x$ we have $x \leq x$.
(c) $\forall\, x, y(x \leq y, y \leq x \Rightarrow x = y)$.
If you say $x = y$ is not true then the two hypothesis imply that $x < y$ and $y < x$ hold. But then (i) says $x < x$ holds and (ii) says this can not hold.

Conversely, suppose we have a binary relation $\leq$ satisfying the above conditions (a), (b), and (c). Define $x < y$ if $x \leq y$ and $\neg(x = y)$. Then this relation satisfies conditions (i) and (ii). Indeed if $x < y$ and $y < z$ hold then all the four relations $x \leq y$, $y \leq z$, $x \neq y$ and $y \neq z$ hold. The first two of these imply $x \leq z$. We need to show $x \neq z$ in order to conclude transitivity. Note the tricky point that $x \neq y$ and $y \neq z$ do not imply that $x \neq z$. If posible let $x = z$. Then $x \leq y$ and $y \leq z$ mean that $x \leq y$ and $y \leq x$ which implies that $x = y$ contradicting $x \neq y$. Finally, the second clause in the definition of $<$, shows that $\neg(x < x)$ holds.

Sometimes you see the definition of partial order as a relation satisfying (a), (b) and (c). The above argument tells you that there is a correspondence between such a definition and the definition we have adapted.

A partial order $<$ is said to be a *linear order* if the following holds:
$\forall\, x, y\ (x < y\ \vee\ x = y\ \vee\ y < x)$.
That is, any two elements are comparable: either they are equal or one is smaller than the other. A set with a linear order is called a *linearly ordered set* or *totally ordered set* or simply a *loset*. Here are some examples of posets and losets.

1. $X = R$, the set of real numbers with usual order. Of course a subset of R with the order restricted to the subset is also an example. For instance one can take $X = \{0, 1, 2 \cdots\}$, the set of natural numbers or $X = Z$ the set of all integers or $X = Q$, the set of rational numbers or $X = [0, 1] \cup [2, 3) \cup (3, 4]$.

2. X is the collection of all subsets of R. Define $A < B$ if $A \subset B$ and $A \neq B$. Again a subcollection with order restricted to that collection is also an example. For instance X could be the collection of all open subsets of R, or the collection of closed subsets of R or X could be the collection of all countable subsets.

3. $X = R^2$, the set of pairs (x, y) of real numbers. Define $(x_1, y_1) < (x_2, y_2)$ if $x_1 < x_2$ and $y_1 < y_2$. This is called coordinate-wise order. You should be careful in understanding such definitions. It appears as if we are defining $<$ in terms of again $<$. The order between pairs is being defined using the known order on real numbers.

Here is a different order on the same set. Define $(x_1, y_1) < (x_2, y_2)$ if either (i) $x_1 < x_2$ or (ii) $x_1 = x_2$ and $y_1 < y_2$. This order is called dictionary order. This is a linear order.

4. X is the set of all bounded real valued functions on $[0, 1]$. Say $f < g$ if $f(x) < g(x)$ for all x. Again subset of X with the order restricted is also an example. For instance, we can take the subset of all continuous functions.

Let $(X, <)$ be a loset. If there is an element $a \in X$ such that for all $x \in X$ either $a < x$ or $a = x$, we say that a is *first element* of X. The condition amounts to saying $\forall x \neg(x < a)$. Similarly if there is an element $b \in X$ such that for all $x \in X$ either $x < b$ or $x = b$, we say that b is *last element* of X. This amounts to saying that $\forall x \neg(b < x)$. Such elements, if exist, are necessarily unique. First and last elements are also called *end points.*

A loset X is said to be dense-in-itself if between any two distinct elements there is another one. More precisely, $\forall a\, \forall b(a < b \Rightarrow \exists x\; a < x < b)$.

The set Z of all integers is not dense-in-itself (with usual order). For example, the set Q is dense-in-itself; and it has no end points. Of course, R also has these two properties. But Q is countable and is the only such set. We shall prove this in the next theorem.

Theorem 1 (Characterization of Q): Let $(X, <)$ be a dense-in-itself countable loset without end points. Then X is isomorphic to Q, that is, there is a bijection $f : X \to Q$ such that $x < y$ iff $f(x) < f(y)$.

You should note that we are denoting partial order by $<$ both on X and Q. Before we prove this theorem, let us see some interesting consequences. Recall that a subset $A \subset R$ is *dense* if every non-empty open interval contains a point of A. You should not confuse this with the concept of 'dense-in-itself'. To avoid confusion, 'dense-in-itself' sets are also called η_0-sets.

Theorem 2 (countable dense subsets of R): Let A and B be two countable dense subsets of R. Then there is a homeomorphism of R which sends A to B. More precisely, there is a homeomorphism $\varphi : R \to R$ such that $\varphi(A) = B$.

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 $a \in A_x, a' \in A, a' < a$ implies 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 is $\sup A_x$. Finally, $x \notin A_x$. Similar statements holds for B_y.

Define $\varphi(x) = \sup f(A_x) = \sup\{f(a) : a \in A_x\}$. This supremum is defined 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 $a, b \in A$ such that $x < a < b < x'$. Thus every point of $f(A_x)$ is smaller than $f(a)$ so that $\varphi(x) \leq f(a) < f(b) \leq \varphi(x')$.

(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 homeomorphism. Just note that any increasing bijection φ of R is a homeomorphism. Being bijection it is first of all strictly increasing. For reals $a < b$, we have

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

To proceed to the next application, recall that a Cantor set in R is a non-empty subset $P \subset R$ such that P is compact, every point of P is a limit point of P and P does not include any non-trivial interval. Thus if $p \in P$ then there is a sequence of points $p_n \in P$ such that $p_n \to p$ and $p_n \neq p$ for all n. Also given any numbers $a < b$, there exists a number c such that $a < c < b$ and $c \notin P$. The standard Cantor set C is the following. Put $C_0 = [0,1]$; $C_1 = [0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$. In general, suppose that C_n is defined as disjoint union of 2^n intervals each of length $1/3^n$. Remove the open middle $1/3$ part of each of these intervals and define C_{n+1} to be the union of the 2^{n+1} closed intervals so obtained. Set $C = \cap C_n$.

Of course you can think of many Cantor sets. For example, you can start with $[0,1]$, divide into 5 subintervals of equal length and take the first, third and fifth closed parts at each stage. You can in fact change the number of intervals into which you want to divide at each stage.

Theorem 3 (Cantor sets are topologically same): Let P and Q be two Cantor sets. Then there is a homeomorphism of R which takes P onto Q.

Proof: We start with two observations.

Every non-empty open set can be expressed in a unique way as a countable (may be finite) disjoint union of (non-empty) open intervals. Indeed, let U be a non-empty open set. Take a point $x \in U$ and consider the union of all open intervals which include the point x and which are contained in U. It is easy to see that union of intervals which have a point in common is again an interval and if they are all open then so is their union. Let this be denoted by I_x. Since $x \in I_x$ we see that the union of all these open intervals is U.

For two distinct points $x, y \in U$, the intervals I_x and I_y are either same or disjoint. Indeed if they have a point in common then their union is again an open interval which contains both x and y. The construction of the intervals now shows that $I_x = I_y$. Consider the distinct intervals I_x. Since each of these intervals must contain a rational number and the set of rational numbers is countable, we have only countably many distinct intervals I_x. Thus U is the disjoint union of these countably many distinct intervals I_x.

If U is expressed as a disjoint union of (non-empty) open intervals, then first of all they must be countable in number; by argument used above. Thus $U = \cup J_n$. Fix any n and consider J_n. Let $x \in J_n$. We show $J_n = I_x$. By construction of I_x, it is obvious that $J_n \subset I_x$. If it is a proper subset, then an end point of J_n, say a, must be in I_x. Since $a \in I_x$ we conclude that $a \in U$. But then it must be in some other J_m. But $a \in J_m$ and J_m is open says that an interval around a must be contained in J_m. But a being end point of J_n we get contradiction to the fact that J_n and J_m are disjoint. This completes the proof the observation.

Here is the second observation. Let P be a Cantor set. Let P^c be the disjoint union of countably many open intervals $\{I_n : n \geq 1\}$. From the above discussion, this family of intervals is uniquely determined. Let $\mathcal{I}$ be the collection of these intervals. Define an order on $\mathcal{I}$ as follows; say $I < J$ if $x < y$ for all $x \in I$, and all $y \in J$. Then $\mathcal{I}$ is a dense-in-itself loset with first and last elements. We shall now prove this assertion.

Since two distinct I and J from $\mathcal{I}$ are disjoint open intervals, if we have $x < y$ for one $x \in I$ and one $y \in J$ then $x < y$ holds for all $x \in I$ and all $y \in J$. It is obvious that we have a partial order. For any two numbers x and y exactly one of the following holds: $x < y$ or $x = y$ or $y < x$. Consequently, for two disjoint open intervals I and J, exactly one of the following holds: $I < J$, $I = J$, $J < I$. Thus the order so defined is a linear order.

Since P is compact, P has a minimum, say, a and a maximum, say, b. Easy to see that the interval $(-\infty, a)$ is in $\mathcal{I}$ and is the first element. Similarly, the interval (b, ∞) is in $\mathcal{I}$ and is the last element.

Suppose we have two intervals $I = (a, b)$ and $J = (c, d)$ in $\mathcal{I}$ and $I < J$. In particular, $x \in I$ and $y \in J$ implies $x < y$. As a result $b \leq c$. Note that $b \in P$, in fact end points of these intervals are in P. If $b = c$, then clearly b is not a limit point of P; this is the only point of (a, d) in P! So we must have $b < c$. Since we can not have the entire interval $[b, c]$ to be subset of P, there must be a point $x \notin P$, and $b < x < c$. But then there must be an interval $I' \in \mathcal{I}$ with $x \in I'$. Easy to see that $I < I' < J$.

Thus $\mathcal{I}$ with the order prescribed above is a countable loset which is dense-in-itself and has first and last elements.

We shall now prove the theorem. Let $\mathcal{I}_P$ and $\mathcal{I}_Q$ be the collections of disjoint open intervals whose union is P^c and Q^c respectively. These are order isomorphic, because you can map the first element of one to the first element of the other, last element of one to the last element of the other; realize that the remaining parts obey hypothesis of Theorem 1 and if you fix an order preserving isomorphism between them, the combined map (remember we already mapped first and last elements) is an order isomorphism.

Fix an order preserving isomorphism $f : \mathcal{I}_P \to \mathcal{I}_Q$. Note that this only means that we have associated with each interval I of the first collection, an interval $f(I)$ of the second collection. Thus when we write $f(I)$ below, do not interpret it as $\{f(x); x \in I\}$, there is no such thing as $f(x)$. We have not defined any map on the real line yet. Here is how we define the required homeomorphism on R. We define h piece-wise.

Let $(-\infty, a)$ be the first element of $\mathcal{I}_P$ and $(-\infty, a')$ be the first element of $\mathcal{I}_Q$. Fix h an increasing homeomorphism of the first onto the other —- by nature it must be strictly increasing and so by defining $h(a) = a'$ it will be a homeo of the corresponding closed intervals.

Similarly, if (b, ∞) and (b', ∞) are the last elements, define an increasing homeo h on the first onto the other — by nature it has to be strictly increasing and so defining $h(b) = b'$ makes it a homeo of the corresponding closed intervals. Let $I = (c, d)$ be any other interval of $\mathcal{I}_P$ and $f(I) = (c', d') \in \mathcal{I}_Q$.

Fix any increasing homeo $h : I \to f(I)$. If we define $h(c) = c'$ and $h(d) = d'$, then it will be a homeo of the closed intervals as well.

Thus we have defined h on the union of all open intervals in $\mathcal{I}_P$ (actually on closed intervals, but do not bother, we will recover the same soon). In otherwords h is a homeo of P^c onto Q^c. We need to define h for points of P. So take $p \in P$. Put $h(p) = \sup\{h(x) : x \in I \in \mathcal{I}_P, I < p\}$. Here we used the suggestive notation $I < p$ to mean that $x < p$ for some (and hence, for all) $x \in I$. Remember since $p \in P$, it is not in any interval of the family $\mathcal{I}_P$.

We claim that h is a strictly increasing function on R. To see this, let $x < y$.

Case 1: none of them is in P. Either they are in the same interval $I \in \mathcal{I}_P$ or in different intervals $x \in I$ and $y \in J$. In the first case, by definition of h on I, it is strictly increasing and so $h(x) < h(y)$. In the second case the hypothesis $x < y$ forces $I < J$. Since $h(x) \in I$ and $h(y) \in J$ we conclude $h(x) < h(y)$.

Case 2: both are in P. Since P does not contain any interval, there must be $I \in \mathcal{I}_P$ with $x < I < y$. Take any point $z \in I$ and remember that h maps I to $f(I)$ to conclude $h(x) < h(z) < h(y)$. The only thing you need to note is that no point of I participates in defining $h(x)$ where as every point of I participates in the definition of $h(y)$.

Case 3: $x \notin P$ and $y \in P$. So there is an $I \in \mathcal{I}$ such that $x \in I$ and $I < y$. Since I is an open interval, fix $x < z \in I$. Clearly $h(x) < h(z) < h(y)$. The remaining case is dealt with in a similar fashion.

We claim that h is onto R. To see this, let $y \in R$.

If $y \notin Q$, then there is $J \in \mathcal{I}_Q$ containing y. If $I \in \mathcal{I}_P$ is such that $f(I) = J$, then h being a homeo of I onto J, we see that there is an x with $h(x) = y$.

Let $y \in Q$. A look at the definition of h for points of P, tells us that the first and last points of Q are images of the corresponding points of P. So let $y \in Q$ be not an end point.

Define $p = \sup\{a : a \in I \in \mathcal{I}_P; f(I) < y\}$.

Since $y \in Q$, this set under consideration is non-empty (contains points of first interval of $\mathcal{I}_P$) and bounded above (does not contain any point of the last interval). Thus the sup is well defined. For $I \in \mathcal{I}_P$ we have $f(I) \in \mathcal{I}_Q$. Since $y \in Q$, either $f(I) < y$ or $y < f(I)$. Accordingly, either every point of I belongs to the set defining p or no point belongs. Hence the sup can not be in any interval of $\mathcal{I}_P$. Thus $p \in P$. We claim that $h(p) = y$.

It is a good idea to recall that $h(p) = \sup\{h(x) : x \in I \in \mathcal{I}_P, I < p\}$.

First note that if $x \in I \in \mathcal{I}_P$ and $I < p$ then by definition of p we have $f(I) < y$. Since $h(x) \in f(I)$, we conclude that $h(x) < y$. So $h(p) \leq y$. Let now $z < y$. Since $y \in Q$, there is an interval $J \in \mathcal{I}_Q$ such that $J < y$ and either $z \in J$ or $z < J$. For example, when y is right end point of an interval in $\mathcal{I}_Q$ and z is in that interval the first of the alternaqtives mentioned occurs. Pick $I \in \mathcal{I}_P$ with $f(I) = J$. Since $f(I) = J < y$, conclude that I appears in the definition of p, so all the numbers $h(x)$ for $x \in J$ appear in the definition of $h(p)$. In other words all numbers in J appear in the definition of $h(p)$. Hence $z < h(p)$. This is true for any $z < y$. So $y \leq h(p)$.

Thus $h(p) = y$ completing the proof that h is onto R. Since h is strictly increasing, it is a homeomorphism. Clearly $h(P) = Q$. ▮

Proof of Theorem 1:

Let us start with an observation. Let S and T be two finite subsets of A with $S < T$. This means $s < t$ for every $s \in S$ and every $t \in T$. Then there is $a \in A$ with $S < a < T$. Since S and T are finite sets take maximum of S and minimum of T and use the fact that A is dense-in-itself. In case one of the sets is empty, the hypothesis that there are no end points makes it possible.

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

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

so that the map $f(a_i) = b_i$ is the required isomorphism. This will be so if we make sure that for each k
$f(a_i) = b_i$ is order preserving on $\{a_1, \cdots, a_k\}$ onto $\{b_1, \cdots, b_k\}$. (*)

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

$$S = \{b_i : i \leq 2, a_i < a_2\}; \quad T = \{b_i : i \leq 2, a_2 < a_i\}.$$

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* the x are listed as a's and all the q are 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. Also at each stage (*) holds. This completes the proof. ▮.

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 $x < a$ or $x = a$. 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 $a < b$ or $a = b$. 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 $b < x$ or $b = x$. 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 $c < b$ or $c = b$. 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 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 $S = [0, 1)$ You should not conclude that this set has no supremum, in fact 2 is its supremum as far as the loset X is concerned. As a loset X is isomorphic to $[0, 2]$. Here is a simple observation regarding existence of infimums and supremums.

Theorem 4 (lub and glb): 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). ∎.

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*. A loset is *separable* if there is a countable set D such that every non-empty open interval contains a point of D.

Say that a loset is *boundedly complete* if any of the above conditions hold. Clearly R has no end points and is dense-in-itself and is separable. Of course Q also has these properties. R has an additional property, namely, it is boundedly complete. In fact R is the only such loset.

Theorem 5 (characterization of R): A dense-in-itself, separable, boundedly complete loset without end points is order isomorphic to R.

Proof: This is repetition of the arguments of Theorem 2 — 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 is its first point then the open interval $\{x \in X; x < d\}$ must be empty showing that d is first point of X too. 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 X is dense-in-itself. Thus there is $d \in D$ such that $a < d < b$.

Since D is a countable dense-in-itself set without end points, we can fix an order preserving isomorphism $f : D \to Q$ and repeat earlier arguments.▮.

Example:

$[0, 1]$ is
separable and dense-in-itself; usual set of rational numbers witnesses these assertions,
boundedly complete; for any non-empty subset, the usual sup in reals would work as sup in this loset too,
but violates 'no end points' condition; the elements zero and one are its end points.

Example:

$(0, 1] \cup [2, 3)$ is
separable; set of rational numbers witnesses this assertion,
boundedly complete; for any non-empty set, its usual sup works as sup in this loset too,
has no end points; zero and one are not in this set,
but violates 'dense-in-itself' condition; there is a gap between one and two.

Example:

Q is
separable; the set itself is countable,
dense-in-itself,
has no end points,
but violates 'completeness'; for example, the set
$\{1, 1+1, 1+1+\frac{1}{2!}, 1+1+\frac{1}{2!}+\frac{1}{3!}, \cdots\}$
has no sup though bounded above.

Example:

$[0,1] \times [0,1] - \{(0,0),(1,1)\}$ with dictionary order is

dense-in-itself; given two points $(x,y) < (x',y')$, the point $(\frac{x+x'}{2}, 0)$ is in between when $x < x'$; if $x = x'$ then take average of other coordinates,

has no end points; given a point, to get something smaller divide first coordinate by two if it is non-zero, otherwise divide second coordinate by two,

boundedly complete; given a non-empty set S bounded above here is its sup: take sup of first coordinates of points of S, let it be x^*, if there is no point of S with this first cordinate then $(x^*, 0)$ is $\sup S$; if there are points of S with first coordinate x^*, then take those and take $y^* =$ sup of the second coordinates of all points whose first coordinate is x^*, then (x^*, y^*) is $\sup S$; just note that the set is bounded above tells you $x^* = 1$ and $y^* = 1$ is not possible ,

but violates 'separability'; for each $x, 0 < x < 1$, define the interval I_x consisting of all points strictly between $(x,0)$ and $(x,1)$, these are uncountably many disjoint non-empty open intervals.

We conclude with one more characterization of the reals using its order properties. We need some definitions. Let X be a loset. A *cut* means a partition $X = L \cup U$ into non-empty sets such that $a \in L, b \in U \rightarrow a < b$.

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

A cut is said to be a *gap* if L has no upper bound in L and U has no lower bound in U.

A cut is said to be *Dedikind cut* if exactly one of the following two happens: L has upper bound in L OR U has a lower bound in U.

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.

Theorem 6: A separable loset without end points in which every cut is a Dedikind cut is order isomorphic to R.

Proof: In view of the previous theorem, we only need to show that the loset X is dense-in-itself and boundedly complete. Take any $a < b$. If there are no points in between, then $L = \{x : x \leq a\}$ and $U = \{x : b \leq x\}$ is gap. Thus X is dense-in-itself.

Let A be any non-empty bounded set. Put

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

If we take any element of X, then either $(\forall a \in A)(x < a)$ in which case it is in L or its negation $\neg(\forall a \in A)(x < a) = (\exists a \in A)(a \leq x)$ holds in which case it is in U. Also $x \in L$ and $x' < x$ implies $x' \in A$. Thus it is a cut.

Suppose that L has an upper bound in L, say, c. Then $c < a$ for all $a \in A$. Hence c is a lower bound for A. Suppose that b is any lower bound for A and if possible let $c < b$. Then by the dense-in-itself property proved above, fix a point $c < x < b$. Since b is a lower bound for A, we see that $x < a$ for all $a \in A$. Thus $x \in L$. But $c < x$ contradicts that c is an upper bound for L. Thus nothing larger than c is a lower bound of A. In other words c is inf A.

Similarly, in case U has a lower bound in U, then that can be shown to be infimum of A. ▮.

Suppose we consider the set N with usual order. Then it has an interesting property: every non-empty subset has a minimum, that is, an infimum which belongs to the given subset. On the other hand R does not have this property. For instance the interval $(0, 1)$ has infimum but no minimum.

Recall that this property of having minimum for non-empty sets is used in proofs by mathematical induction or in defining functions by induction. For example, consider proving the fact, $1 + 2 + \cdots + n = \frac{n(n+1)}{2}$ for every integer $n \geq 1$. One way is to say: prove it for $n = 1, 2, 3$ and say 'so on'. The only trouble is that there is no one to assure us that 'so on' can be believed and that the proof works for every n. This approach has another subtle problem, namely, you will be able to prove for each given n and not for all n. The subtlety is that if you want to execute the proof as suggested, you need at least one line for every n and hence infinitely many sentences are needed to prove it for all n.

Another way is: prove for $n = 1$; assume that it is proved for $n = 1, 2, \cdots, k$ and using this prove for $n = k + 1$, say the proof is complete. Why is this method acceptable? Consider the set of n for which the stated formula is not correct. If this set is empty, we are done. If it is non-empty, then it has a first element, say, p. Can $p = 1$?? No, because we verified for $n = 1$. But then it is true for $n = 1, 2, \cdots, p - 1$, so must be true for $n = p$ as well. We have taken a simple illustration, in this case there is a one line proof, namely add these numbers in reverse order to get that double the sum is $n(n + 1)$.

Losets which posess this property of having a minimum for every non-empty subset are important and have a special name. A loset $(W, <)$ is said to be a *well ordered* set or simply *woset* if given a non-empty subset $S \subset W$ there is an element $a \in S$ such that for each $x \in S$ and $x \neq a$ we have $a < x$. Such an element is also called *first element* of S.

It is hard to see uncountable wosets, though there are plenty and unending. It is easy to see countable wosets. Of course N and any non-empty subset are wosets. In particular initial segments of N are wosets. Here are some more countable wosets with self explanatory order.
$\{0, 1, 2, \cdots, \infty\}$; $\{0, 1, \cdots, \infty, \infty + 1.\}$;
$\{0, 1, \cdots, \infty, \infty+1, \infty+2, \cdots, \infty+\infty, \infty+\infty+1, \cdots, \infty+\infty+\infty.\}$. Here the notation $+$ is a suggestive notation, but at this moment you see no addition

floating around. In common parlance $\infty + \infty = \infty$ and we definitely do not intend this here. To avoid this confusion, we should have used ω instead ∞.

A tricky point we have avoided so far, necessary for technical reasons (and you need not pay any attention unless you reach the stage of technical demand) is the following. We always started with non-empty set X in defining posets. We should allow the emptyset $\emptyset$ also as a possible candidate for X. But of course, $X \times X = \emptyset$ and there is only one binary relation on this X, namely, $R = \emptyset$. We regard this also as poset, loset, and woset. Of course, you would wonder why we are making fuss about nothing. Beacuse, this 'nothing' is precisely 'zero'. But you need not bother about it now. All posets below are non-empty, in the sense, if it is given to us we assume that it is non-empty; if we construct we need to show that it is non-empty.

Axiom of Choice:

There are several equivalent formulations of this axiom. Here are some.

(1) Given any 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 $A \in \mathcal{A}$.

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. 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 el-

ements, 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.

(4) Every non-empty set can be well-ordered, that is, given a non-empty set X, there is a binary relation $<$ on X so that $(X, <)$ 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 the 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'.

It is natural to ask if we need any axiom at all. Is it not obvious that we can select one element from a non-empty set. Yes, from one given non-empty set we can select a point, simply because it is non-empty. However, if you have a family of sets how can you make a new set as stated in (1). You might say, why not select one element from each of the sets and call this set of points so selected as S. The trouble is, you did not give me any rule as to which point is to be selected. If I ask you ‘is a in your set” you would not know the answer. If you do not know which points are there and which are not, how can you say you have a set?

This is beautifully illustrated by Bertrand Russell. If you have a large collection of pairs of shoes, you can select one from each pair. But if you have collection of pairs of new socks, there is no way to select one from each pair.

Suppose, $\mathcal{K}$ is the collection of non-empty compact subsets of R, set of real numbers. Then we have a choice function, you can say select the maximum from each set or you can say select the minimum from each set. But for arbitrary collection of sets, we do not have such a rule.

You might still wonder why we should accept this for one set. After all even if we have one set, we did not prescribe any rule how to select. This is simply because I can not prescribe rule without having the set before me. Once you give me a set I can give the rule. Would not this argument hold for a family of sets too. No. There are ways of clearly prescribing a family of sets for which there is no clear rule of picking points. The real reason lies on axioms of set theory, more precisely, on how we agreed to make sets. We shall not enter into the details.

You will see later that even though this axiom is intuitively ‘acceptable’, there are some consequences which are ‘doubtful’.

Theorem 7: Statements (1), (2), (3) and (4) are equivalent.

Proof of Theorem 7: Assume (1). We prove (2). Let $\{A_\alpha : \alpha \in \Delta\}$ be any family of non-empty sets. Put $B_\alpha = \{\alpha\} \times A = \{(\alpha, x) : x \in A_\alpha\}$. These are disjoint non-empty sets. Get S such that $S \cap B_\alpha$ is singleton for each α. Define $g(\alpha)$ by $\{g(\alpha)\} = S \cap B_\alpha$. Take $f(\alpha)$ to be the second coordinate of $g(\alpha)$. In this proof, we have brought the additional set Δ. Actually it is not necessary. We could take Δ to be $\mathcal{A}$ itself. Instead of B_α simply take $\{A\} \times A = \{(A, x) : x \in A\}$; for each $A \in \mathcal{A}$. This may be confusing at first sight.

Assume (2). Given a family $\mathcal{A}$ of disjoint sets, take f and put S to be range of f. This proves (1).

Assume (3). To prove (1), let $\mathcal{S}$ be the collection of sets S such that $|S \cap A| \leq 1$ for each $A \in \mathcal{A}$. This is non-empty because, take one $A \in \mathcal{A}$ and using the fact it is non-empty take one $a \in A$ and put $S = \{a\}$. Order this family by inclusion. Given a chain, their union is its upper bound. So get maximal element, say, S. Of course, for each $A \in \mathcal{A}$ we have $S \cap A$ is at most one point. Suppose for one $A \in \mathcal{A}$ we have $A \cap S = \emptyset$. Then pick one $a \in A$ and observe that $S \cup \{a\}$ contradicts maximality of S.

Assume (2). We shall prove (3). Let us define a chain C to be maximal if there is no element larger than everything in the chain C. More precisely, $\neg(\exists p)(\forall x \in C)(x < p)$. Carefully note that the chain may not be 'physically' the largest. For example, in the poset P consisting of subsets of R with at most four elements, the following are maximal chains.
$\emptyset, \{3\}, \{3, 49\}, \{3, 49, 99\}, \{3, 49, 99, -31\}$.
OR the chain consisting of two elements $\{1\}; \{1, 2, 3, 4\}$
OR the chain consisting of just one element $\{1, 2, 3, 4, \}$

It is enough to prove that there is a maximal chain. Suppose we have done this. Then take a maximal chain C and its upper bound a. If there is b with $a < b$, then maximality of the chain is contradicted.

To show that there is a maximal chain the intuitive idea is to start with an element, take one larger than it, take one larger than this and so on. If you fail at some stage you got your maximal chain because you could not get anything larger than what you already got. You should stop at some stage because you are taking points from your poset P — you definitely stop when all points of P are exhausted. This argument can be made precise using well-ordered sets (which exist irrespective of axiom of choice). However we follow the route of Hausdorff.

Here is the idea. Consider R, set of real numbers. Suppose you want to make a set wich includes integers $1, 2, \cdots$ etc and no others. You can say, take 1, then take 2, and so on. Here is another way. Consider all sets C having the property: $1 \in C$; $a \in C \Rightarrow a + 1 \in C$. For example R itself is one such OR the set Q of rational numbers OR the set of all half integers, $\{n/2 : n \in Z\}$. Take the intersection of all these sets.If you think about it, this intersection is none other than N. The second method appears silly, are we complicating life? You see that the word 'so on' is not used in this method and a way to recognize N is given, namely it is the smallest set satisfying the conditions stated above.

Here is how we construct maximal chain. Let $\mathcal{C}$ be the collection of all chains. Just for convenience let us put the empty chain also in this collection. [If you feel uncomfortable with this arrangement, you need not do this, When I use empty set later on, you start with a fixed singleton]. For each $C \in \mathcal{C}$ let us fix an element $c \in C^c$ such that $x < c$ for all $x \in C$; provided there is one such. Here is how we do it.

For each chain C, let $C' = \{p \in C^c : x < p \text{ for all } x \in C\}$. Of course, this set may be empty. Let $\mathcal{C}_0$ be the collection of all chains C for which $C' \neq \emptyset$ and $\mathcal{C}'_0$ is the collection of sets $\{C' : C \in \mathcal{C}_0\}$. Then $\mathcal{C}'_0$ is a family of non-empty sets indexed by $C \in \mathcal{C}_0$. Fix a function f as stated in (2). Let us define $g(C) = f(C')$ for $C \in \mathcal{C}_0$.

We now define successor for chains as follows. For $C \in \mathcal{C}$, let us put $s(C)$ to be C if $C \notin \mathcal{C}_0$ and equal to $C \cup \{g(C)\}$ if $C \in \mathcal{C}_0$. Thus $s(C)$ is one choice of chain larger than C, provided there is one such. Of course if there is nothing larger than C then $s(C)$ is C itself. Thus note that $s(C)$ is always defined.

A subset $\mathcal{D} \subset \mathcal{C}$ is called *normal family* if it satisfies the following three conditions.
(i) empty chain is in $\mathcal{D}$.
(ii) If $C \in \mathcal{D}$ then $s(C) \in \mathcal{D}$.
(iii) Any union of a subcollection of $\mathcal{D}$ which is a chain, then it is again in $\mathcal{D}$.

Of course, the family $\mathcal{C}$ is a normal family. So there are normal families.

Intersection of two normal families is again a normal family. Let $\mathcal{D}_1$ and $\mathcal{D}_2$ be normal families. Since $\emptyset$ is in both, it is in the intersection. If C is in the intersection, then it is in both the families and hence so is $s(C)$ and hence $s(C)$ too is in the intersection. If you take union of elements in the intersection, then it is union of elements in each of the families too. so if it is a chain then it is in both the families and hence is in the intersection. Thus $\mathcal{D}_1 \cap \mathcal{D}_2$ satisfies the three conditions.

In fact intersection of all normal families is a normal family. Exactly the same argument as above shows this. Let us denote it by $\mathcal{N}$. **Enough to show that the union of elements of $\mathcal{N}$ is a maximal chain**.

Note that elements of $\mathcal{N}$ are chains. That is, they are certain subsets C of the poset P where any pair of elements of C are comparable in the given partial order $<$. **It is enough to show that for any two chains C_1 and C_2 in $\mathcal{N}$ one is a subset of the other.** If this is done, then union of sets in $\mathcal{N}$, say $\widetilde{C}$ will be a chain. Indeed, suppose $x, y \in \widetilde{C}$ be distinct elements. Say $x \in C_1$ and $y \in C_2$; both C_1 and C_2 are in $\mathcal{N}$. If for example $C_1 \subset C_2$ then both x and y are in C_2. But C_2 is a chain, so either $x < y$ or $y < x$. so $\widetilde{C}$ is a chain. By the third condition of normal family, we see that $\widetilde{C} \in \mathcal{N}$. By the second condition $s(\widetilde{C}) \in \mathcal{N}$. The fact that $\widetilde{C}$ is union of *all* sets in $\mathcal{N}$ now says that $s(\widetilde{C}) = \widetilde{C}$. In other words, $\widetilde{C}$ is a maximal chain.

Let us say that an element C of $\mathcal{N}$ is *good* if for any element N of $\mathcal{N}$. either $C \subset N$ or $N \subset C$. From now on, the set inclusion includes equality also, that is $\subset$ means $\subseteq$. **It is enough to show that all sets in $\mathcal{N}$ are good.**

We reduce our problem further. **It is enough to show that if C is a good set and $N \in \mathcal{N}$ then either $N \subset C$ or $s(C) \subset N$.** Suppose this is done. Then we argue as follows.

Consider the collection of all good sets. Clearly emptyset is good, because it is in $\mathcal{N}$ and clearly comparable with every set. Let C be good and $N \in \mathcal{N}$. Then, by the property of good sets stated above, either $s(C) \subset N$ or $N \subset C$ in which case $N \subset s(C)$. In other words $s(C)$ is also good. Now consider a union of good sets. Firstly this union is a chain (because these are good and hence comparable among themselves too) and hence is in $\mathcal{N}$. Let $N \in \mathcal{N}$. Either all these good sets are contained in N in which case so is their union; or one of them contains N in which case their union contains N. In other words union of good sets is also comparable with every element of $\mathcal{N}$ and is hence good. Thus the collection of good sets satisfies all the three conditions of a normal family. But remember $\mathcal{N}$ is the intersection of all normal families and the collection of good sets is a subcollection of $\mathcal{N}$ which is proved to be a normal family. We conclude that every set in $\mathcal{N}$ is good.

We have so far been reducing our problem. Now we finally conclude proof of the theorem by showing that if C is a good set and $N \in \mathcal{N}$ then either $N \subset C$ or $s(C) \subset N$. We repeat the same type of argument used above. Fix a good set C. Consider
$\mathcal{M} = \{N \in \mathcal{N} : N \subset C \text{ or } s(C) \subset N\}$.
Observe that $\emptyset \in \mathcal{M}$ because $\emptyset \in \mathcal{N}$ and $\emptyset \subset C$. Union of any subcollection of $\mathcal{M}$, if it is a chain, then it is in $\mathcal{M}$. Because that is a subcollection of $\mathcal{N}$ and hence the union, being chain, is in $\mathcal{N}$. If each of them is a subset of C, then so is the union. If one of them contains $s(C)$ then their union also contains $s(C)$. Thus the union is in $\mathcal{M}$.

We now show that if $N \in \mathcal{M}$, then so is $s(N)$. Indeeed $N \in \mathcal{N}$ and hence $s(N) \in \mathcal{N}$. C being good it is comparable with all elements of $\mathcal{N}$.
So either (i): $C \subset s(N)$ or (ii): $s(N) \subset C$ or (iii): $C = s(N)$.
Since $N \in \mathcal{M}$ we have either (a): $N \subset C$ or (b): $s(C) \subset N$.
Need to show either (α): $s(N) \subset C$ or (β): $s(C) \subset s(N)$.
If (b) holds, then whatever be (i)-(iii), we see (β) holds.
If (a) holds, then in cases (ii) and (iii) we see that (α) holds.
The only case to be considered is (a) and (i): $N \subset C \subset s(N)$. Remembering that for any chain C, the set $s(C)$ contains at most one extra point, we conclude that $s(N) - N$ is a singleton. Thus either $C = N$ showing (β) holds or $C = s(N)$ showing (α) holds.
Thus the family $\mathcal{M}$ is a normal family and is a subfamily of $\mathcal{N}$ and hence must equal $\mathcal{N}$. In other words, if C is a good set then either $N \subset C$ or $s(C) \subset N$ for each $N \in \mathcal{N}$.

This completes proof of (3) assuming (2). ∎

We did not discuss (4). We give some applications of the Axiom of Choice.

Recall definitions of vector space, independent set of vectors , maximal independent set which is called basis. For our purpose it is enough to consider vector spaces over reals R or complex numbers or over field of rationals Q. Recall that if B is a basis for the vector space, then every vector is a finite linear combination of basis vectors; with coefficients from the field. Also recall that if we define a function on a basis B of a vector space V taking values in a vector space W, then it can be extended in a unique way as a linear map between the vector spaces. Of course, here both V and W are vector spaces over the same field.

Theorem 8: Every vector space has a basis.

Proof: Consider the collection $\mathcal{C}$ of independent sets B. This is a poset with set inclusion as the partial order. If you take a chain $\mathcal{C}$, it has an upper bound, namely, $B = \cup\{C; C \in \mathcal{C}\}$. If you take $v_1, \cdots, v_k$ from this union, say, $v_i \in B_i \in \mathcal{C}$, then $\mathcal{C}$ being a chain the sets (B_i) are comparable. Thus one of these sets contains the others, and hence, in particular, one B_j contains all the v_i. Since this set B_j is an independent set, the collection of vectors are independent. Thus by Zorn's lemma, the poset has an upper bound, say L. Clearly, if v is not in the span of L, then the set obtained by adjoining v to L is an independent set contradicting the maximality of L. ∎

Theorem 9: There is a function $f : R \to R$ such that $f(a + b) = f(a) + f(b)$ for all $a, b \in R$ which is not continuous.

Proof: Consider R as a vector space over the field of rationals Q. Let H be a basis. Clearly H is uncountable, otherwise their finite rational linear combinations would be countable too, but this is R. Take one element $h \in H$. Define $f(h) = 1$ and $f(v) = 0$ for $v \in H$, $v \neq h$. Extend this by linearity (over Q) to all of R. This will do. This is not continuous because, such a function which is continuous must be of the form $f(x) = cx$ for some c. If $c = 0$ then $f(h) = 1$ is contradicted and if $c \neq 0$ then $f(v) = 0$ for $v \in H$, $v \neq h$ is contradicted. ∎

Theorem 10: There is a bijection $f : R \to R^2$ such that $f(a + b) = f(a) + f(b)$ for all $a, b \in R$.

Proof: Take any basis H for the vector space R over Q and a basis K for the vector space R^2 over$Q$. Both must have the same cardinality c, of the continuum. If cardinality of H is less than c, then finite rational linear combinations of elements of H also would have cardinality smaller than c, but H spans R. Simialrly K has cardinality c. Thus there is a bijection between the two bases which can be extended as a bijective linear (over the field of rationals) map between the vector spaces. ∎

Theorem 11: Let $(P, <)$ be a poset. We can extend the given order to a linear order. That is, we can define an order $\prec$ on P such that $(P, \prec)$ is a loset and $a \prec b$ if already $a < b$.

Proof: Let $(P, <)$ be a poset. We start with an observation. Suppose there are two distinct elements p and q such that $\neg(p < q)$ and also $\neg(q < p)$. We shall define an order on P as follows

$$a \lhd b \longleftrightarrow (a < b) \vee (a \leq p \& q \leq b).$$

Remember, $a \leq p$ means either $a < p$ or $a = p$. Remember that p and q are fixed and given to us. For any two elements a and b of the poset, we are defining the order $\lhd$ as above.

$\underline{\neg(a \lhd a)}$: We know $\neg(a < a)$ holds. Can $(a \leq p \& q \leq a)$ hold? Since p and q are distinct, equality can not hold at both. Since p and q are not comparable, equality can not hold at exactly one place. In the remaining case we have, $q < a < p$ which again can not hold because p, q are not comparable.

$\underline{a \lhd b, b \lhd c \text{ implies } a \lhd c}$: There are four possibilities in the hypothesis.
$a < b$ and $b < c$ In this case $a < c$ so that $a \lhd c$.
$a < b$ and $(b \leq p \& q \leq c)$. In particular $a < b$ and $b \leq p$, so $a \leq p$. And of course $q \leq c$. Hence $a \lhd c$.
$(a \leq p \& q \leq b)$ and $b < c$. Then $(a \leq p \& q \leq c)$ so that $a \lhd c$.
$(a \leq p \& q \leq b)(b \leq p \& q \leq c)$. Now $q \leq b$ and $b \leq p$ implies $q \leq p$ which is not possible by assumption.
$\underline{p \lhd q}$: Satisfies second clause of the definition.
$\underline{\text{if } a < b \text{ then } a \lhd b}$: First clause of the definition gives this.

Thus given any partial order and two incomparable elements, we can extend the given order to a partial order where these are comparable. Now let us consider the collection $\mathcal{C}$ of all partial orders R on P extending the given order. Say $R_1 \prec R_2$ if $R_1 \subset R_2$ and $R_1 \neq R_2$. This is nothing but set inclusion and is clearly a partial order on $\mathcal{C}$. Every chain has an upper bound, namely, the union of relations in the chain. So $\mathcal{C}$ has a maximal element. Take one such. If it is not a linear order, say two elements are not comparable, then the first part of the proof leads to an extension contradicting the maximality. ∎

Let us say that two subsets S and T of the real line R are 2-translation equivalent if $S = S^1 \cup S^2$ and $T = T^1 \cup T^2$ such that $S^1 \cap S^2 = \emptyset$, $T^1 \cap T^2 = \emptyset$ and T^i is a translate of S^i for $i = 1, 2$. Recall that a set B is translate of A means that there is a number x such that $B = A + x = \{a + x : a \in A\}$. We are talking about only subsets of R.

Theorem 12: There is a decomposition $[0, 1) = A_1 \cup A_2 \cup A_3 \cup \cdots$ into disjoint sets such that for any i and j, the sets A_i and A_j are 2-translation equivalent.

Proof: For points of $[0, 1)$, define $x \sim y$ if $x - y$ is rational. This is an equivalence relation. Each equivalence class is a countable set. Let S be a choice set for the collection of equivalence classes. For each rational $0 \leq r < 1$, let S_r denote the set of all numbers $x + r$ modulo one. We claim that these sets make up all of $[0, 1)$. Indeed let $0 \leq x < 1$ and $y \in S$ be the element chosen from the equivalence class of x Let $r = x - y$ if $x \geq y$ and $r = y - x$ if $y > x$. Then r is rational number. Since $0 \leq x < 1$ and $0 \leq y < 1$ we conclude that $0 \leq r < 1$. In the first case $x = y + r \in S_r$ and in the second case $x = y - r = y + (1 - r)$ modulo 1 and is hence in S_{1-r}.

We claim that these sets are disjoint. In fact if $x \in S_p$ and $x \in S_q$, then there are points $y \in S$ and $z \in S$ such that $x = y + p$ or $x = y + p - 1$ and $x = z + q$ or $z + q - 1$ In any case $y - z$ is a rational number. Thus y and z are in the same equivalence class. Since we have only one element from each class, we conclude that $p = q$.

Thus we have a countable decomposition of $[0, 1)$ as stated. We now show they are 2-translation equivalent. consider S_p and S_q, $0 \leq p < q < 1$. Then $S_q = S_p + (q - p)$ modulo 1, where $0 < q - p < 1$. Let $a = q - p$. Put $A_1 = S_p \cap [0, 1 - a)$ and $A_2 = S_p \cap [1 - a, 1)$. Similarly, $B_1 = S_q \cap [a, 1)$ and $B_2 = S_q \cap [0, a)$. Then $A_1 \cup A_2$ is a decomposition of S_p; $B_1 \cup B_2$ is a decomposition of S_q; $B_1 = A_1 + a$ and $B_2 = A_2 - a$. This completes the proof.

Theorem 13: There is a subset of $[0, 1)$ which is not Lebesgue measurable.

Proof: Take any one of the above sets S_p. Since Lebesgue measure is translation invariant, all the sets should have same measure, if they are measurable. This is impossible.

Here is an interesting consequence of the Axiom of choice, called Banach-Tarski Paradox. This is based on a beautiful and profound construction of Hausdorff. This leads to a new theory.

Theorem 14: A foot ball can be cut into finitely many pieces and the pieces can be rearranged to get two foot balls each of the same size as the original one.

What does this mean mathematically? Let surface of the unit ball in three dimensions be denoted by $S = \{(x, y, z) \in R^3 : x^2 + y^2 + z^2 = 1\}$. We can decompose $S = A \cup B$ into two disjoint sets such that $S \sim A$ and $S \sim B$.

What does $P \sim Q$ mean? We can partition $P = P_1 \cup P_2 \cup \cdots \cup P_k$ and $Q = Q_1 \cup Q_2 \cup \cdots \cup Q_k$ and construct rotations $\{\rho_i, 1 \leq i \leq k\}$ of R^3 such that $\rho_i(P_i) = Q_i$. In other words, cutting P and rotating the pieces we get exactly Q.

Here are some references.

W. Sierpinski: Set Theory.

W. Sierpinski: Congruence of sets
(Lectures given at Lucknow University).

F Hausdorff: Set Theory.

Rubin, H and Rubin, J: Equivalents of Axiom of Choice.

****