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.