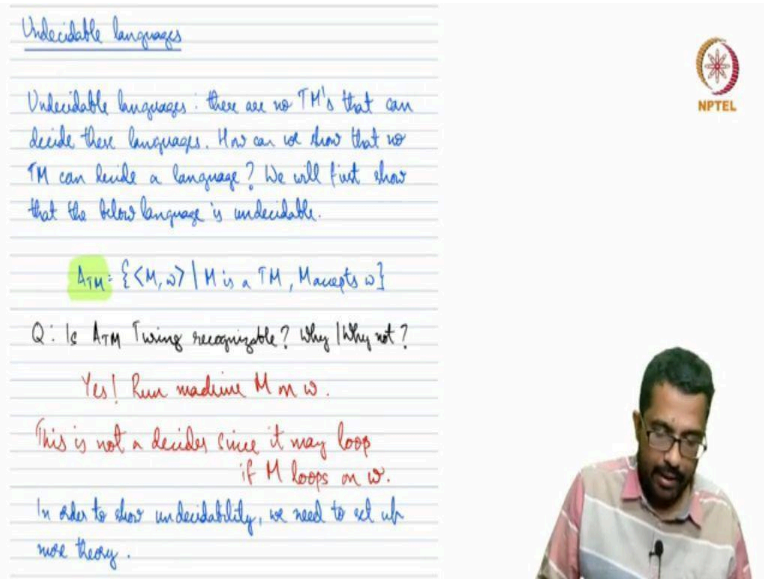
Theory of Computation

Professor Subrahmanyam Kalyanasundaram
Department of Computer Science and Engineering
Indian Institute of Technology, Hyderabad

Lecture 40

Countability of Sets

(Refer Slide Time: 00:16)



Hello and welcome to lecture 35 of the course theory of computation. In the previous couple of lectures, we saw decidable languages arising out of regular languages and context-free languages. So, now we will try to understand undecidable languages. So, we will try to see like undecidable languages; but before we can prove that these languages are undecidable, we will have to build some theories. So, if you think about it, to show that something is undecidable is not clear how to show it. So, basically you have to say that there are so many approaches in which you could possibly decide a language; and you have ruled out each and every one of these approaches.

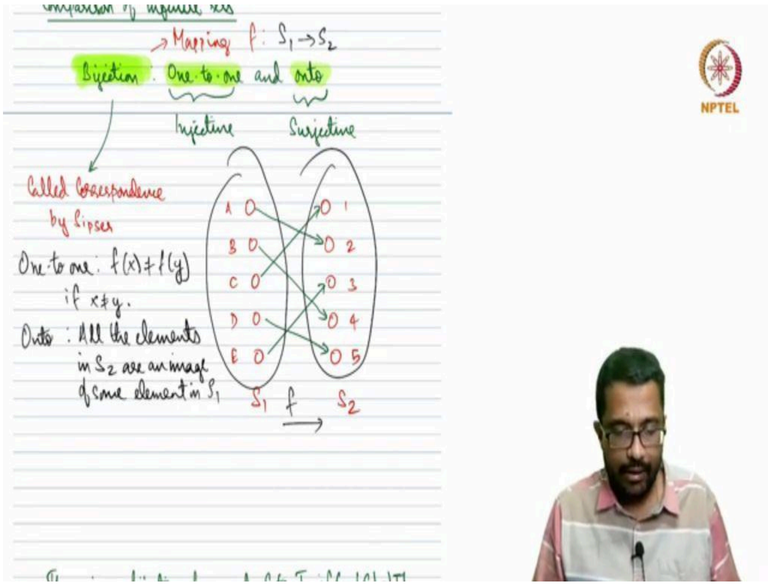
So, how do you possibly show that no Turing machine is able to decide this language? So that is really a challenge, and it is not so straightforward. So, in order to get there, we will have to actually build some theories. So, the language that we will show is undecidable is this acceptance problem of a Turing Machine ATM. So, given a Turing Machine M and a string W, the question is, does M accept W or not? So, this it is not, this is this happens to be somewhat tricky to build a decider for; and will show that this is in fact undecidable. So, before getting into decidability, is this Turing recognizable?

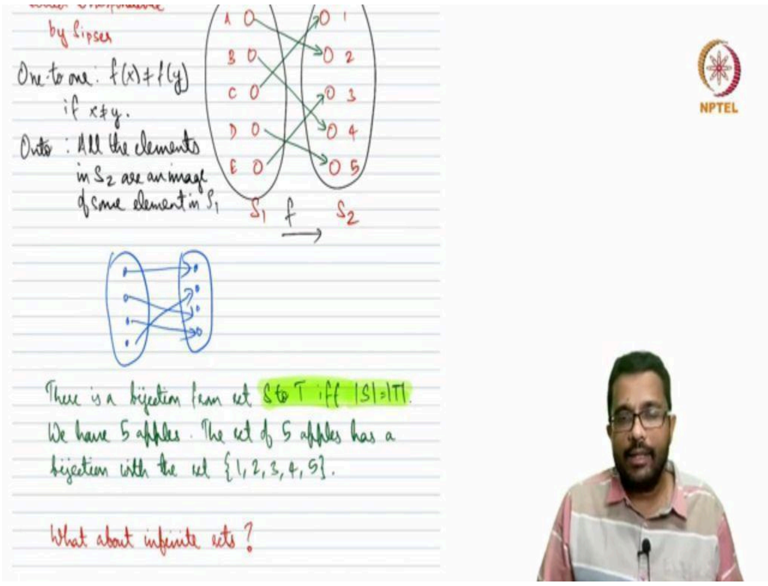
$A_{TM} = \{< M, W > | M \, is \, a \, TM \, , M_{accepts} \, W\}$

So, if you think over it for a bit, you will see that it is Turing recognizable. Because what you can simply do is run or you can, or you can the recognizer will just simulate the machine M on the string w; so run the machine M on w. Now, if M accepts w this will you accept; if M rejects, you reject. But, the problem is that this is not a decider. This is because, if M a decider has to accept if M accepts w, and reject if M does not accept W. But, if M loops on W. if M does not reject W and M does not accept either if M loops on W; there is no way this machine will get to an outcome.

It will also continue to loop; because what we are doing is simply running M on W, and then just reporting whatever the outcome is. So, if M loops on W, this is not going to reject. But, a decider for ATM has to reject $< M, W >$, if M does not accept W. But, this is not a decider since it may loop if M loops on W; so, that is why it is not a decider. However, it is Turing recognizable. Now, as I said before, to actually show undecidability, we need to set up a bit more theory; then let us get to that.

(Refer Slide Time: 03:49)





So, the first point that I have to cover is that of bijection; and a Sipser calls it correspondence. It is not really that harder concept; it is something that many of you may have already learned in your school or maybe in your school, possibly. So, bijection is a function, it is a mapping from one set A to another set B. So, it is a mapping from mapping from one set to another, mapping from let say mapping from A to B; and it has two properties. One is that it is one-on-one, sorry, one-to-one; and the second is that it is onto.

So, by that what I mean is a mapping from let say, here let us say I have two sets; S1 here and S2 here. So, S1 is the set on the right side, and S2 is the set on the left side; so this is S1 and this is S2. Now, for every element in S1, I want to identify a map; so this is the map f. So, every element in S1, I want to identify an image on S2. So here, A has an image, B has an image, C has another image. So, what do I mean by one-to-one and what do I mean by onto. One-to-one means no two elements from S1 can have the same image. So, if you see here, that is indeed the case.

A like you take any two elements of the s of the left side, they do not map to the same right side element; so, this is one-to-one. Onto simply means that all the right side elements are covered; there is no element in the right side which is not an image. So, maybe I will just define it very briefly. So, one to one means $f(x)$ is not equal to $f(y)$, if x is not equal to y. So, the only way two elements can have the same image; or if the two elements are indeed the same, or let us say this in a twisted way of saying it. So, basically two distinct elements x and y if x is not equal to y, their images are also going to be different. And onto means so maybe over here, I will say f of S1 to S2. $f(S1)\rightarrow S2$

All the elements in S2 are an image of some element in S1; meaning every element in S2 has a pre-image. So, you pick any y from S2; there is some x in S1 such that $f(x) = y$. So, which means that there are no elements here which are not a, which are not an image of an element in S1; so, this is what onto means. And one-to-one and onto together means it is a bijection; and it is actually it is actually a nice set of properties. You, so we have two sets here S1 and S2 like that I have pictorially drawn.

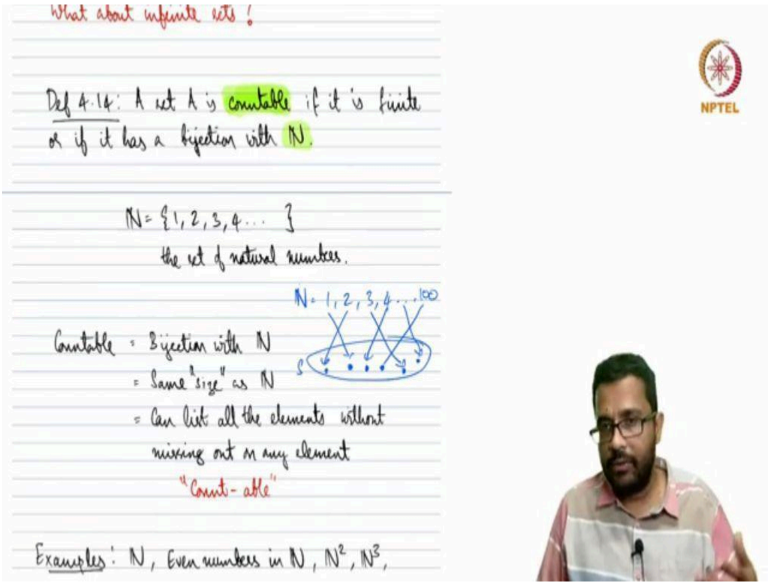
So, notice that if I had S2 had one more element; now we cannot have a bijection between S1 and S2. This is because there will be an unequal number of elements in S1 and S2. And however, which way you try the one element in S2 will not be covered. Also notice that suppose there was one more element in S1, even then it will not be the, you cannot get a bijection; because there will be one extra element in S1. And consequently, you will have to have two elements in S1 map to the same element in S2; so, the one-to-one property will have to be violated. So, what I am saying is that you can have a bijection from S1 to S2, if and only if it is the same size.

And this is very easy to see for finite sets; so let me just write this here. We can have, maybe, in fact I do not have to write it; I have already written it already. There is a bijection from S to T, if and only if the size of S is equal to size of T; size of S is equal to size of T and this is pretty clear. So, maybe you can just give another example, let us try some one more thing. So, let us say four elements here, four elements here, you can have a bijection. So, let's say one, this is a bijection; and you cannot have a bijection if the sizes of the sets are different. So for instance, I have written one more example here is if you have a set of five apples; this set can have a bijection with the set 1, 2, 3, 4, 5.

So, you can map one apple to number 1, one other apple to number 2, and so on; so, this is the point of bijection. So, bijection means it is a mapping from one set to another with the properties one-to-one and onto. And we can only have it if the cardinalities of the sets are the same. Now, this is all easy to see when we are dealing with finitely sized sets.

Now, we want to extend this notion to infinite sets; so, we want to extend this notion to infinite sets. So, we want to say that even for infinite sets; let us say A and B are infinite set size, infinite sized sets. We want to say that A and B are of the same size if there is a bijection from A to B, even though they are infinite sets, infinite size sets. So, we want to extend this notion to infinite sets.

(Refer Slide Time: 10:38)



And that is where the definition of countability comes into picture. So, we say that a set A is countable, if there are two possibilities; one is that it is finite. And two is that it has a bijection with the natural set of natural numbers n. So, a set A is countable if it is finite, or if it has a bijection with a set of natural numbers; the set of natural numbers is simply 1, 2, 3, 4 up to infinity. 1, 2 starting from 1, so the word is countable.

So, why do we say countable? Because suppose it has a bijection within n. So, which means this is, I will draw it horizontally, so this is n; 1, 2, 3, 4 and so on. And there is another set, let say some set S, we want to say S is countable. Meaning there is a bijection from N to S, which means 1 is assigned to some element here, 2 is assigned to some element here, 3 is assigned to some element here, 4 is assigned to some element here and so on.

And every element in S is mapped to some element in N or N is mapped to some element in S. So in other words, we can think of basically we are calling one element the first element. So, this element is the first element, this element is the second element, this element is the third element, and this element is a fourth element and so on.

So, basically we are in a way we are labelling, we are kind of labelling elements in S1 by 1. So, this is the first element, the second element, this is the third element, this fourth element and so on. And every element in S, so you consider any element, let say you consider this element in S; it must be mapped to some number here. So, maybe this is the hundredth element maybe.

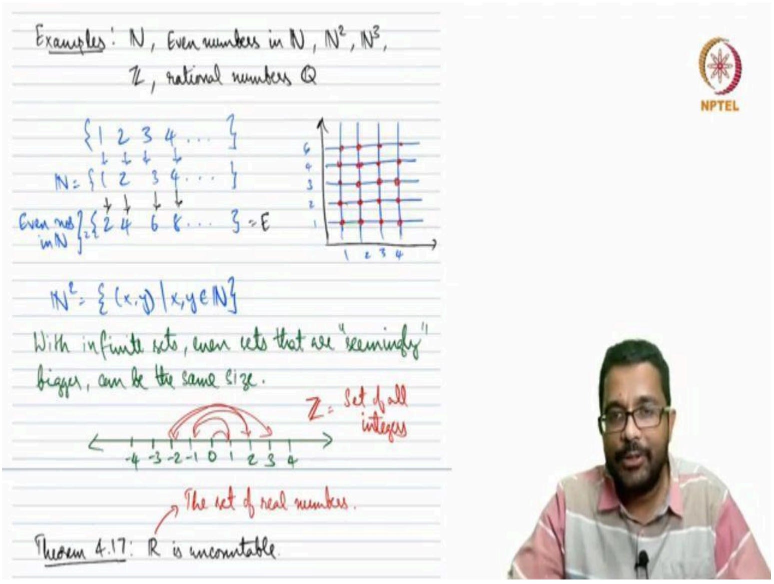
So, every element in S has to have some; it has to feature in this listing somewhere. If it is, if it does not feature that means it does not have a pre-image; and that means it is not a bijection, it will not be onto. So, every element features in this listing somewhere. So, basically you can take this S and you can count elements one by one, this is number 1, this is number 2, this is number 3, the number 4 and so on; this is what I mean by countable. At initially, on the first glance, it may seem that this is a fairly trivial thing. Even if it is an infinite sized set, I should be able to set move one element at a time, and say this is the first element, this is the second element, this is the third element and so on.

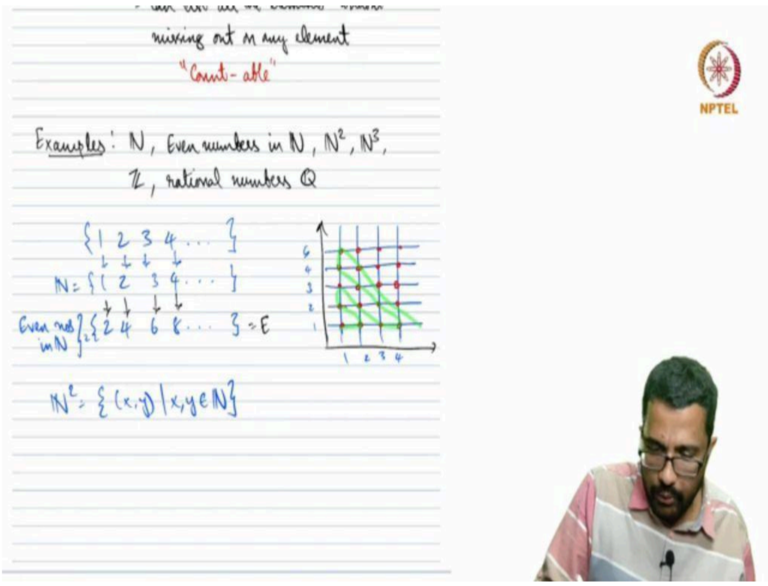
But, it turns out that it is not so it is not so common. In fact, most of the sets that we have that we know are going to be uncountable. So, being comfortable, even amongst infinite sets, is a very very special property; so we will see why it is a special property. So, the reason, so I am hopefully at this point, is clear why it is called countable.

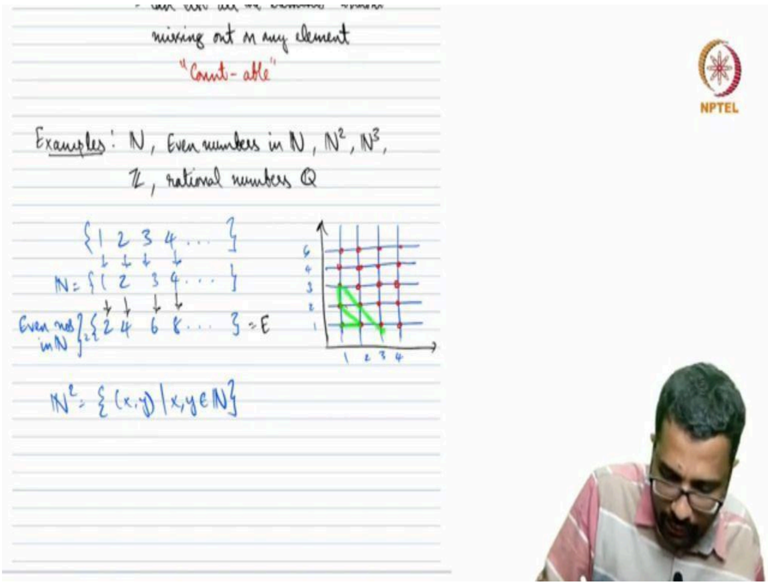
Basically, you have an infinite set and we can count it. This is the first element, this is the second element, this is the third element and so on; so that none of the elements of the set are missed out. All the elements appear at some point in the ordering. So, if you want to list out a set is countable, if you can if it has a bijection with n; or we can say you can list out the elements of the set one by one, so that no element is missed out.

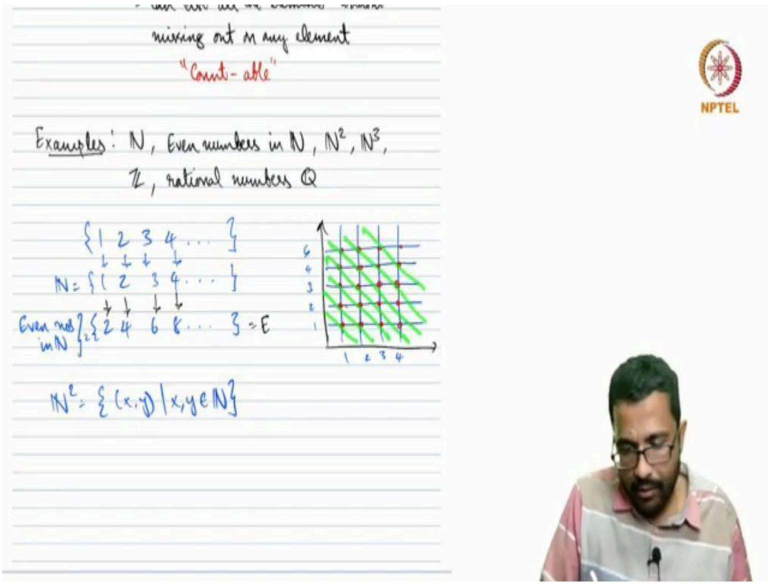
So, it is like you are counting, this is the first, this is second, this is third and so on. This is what I mean by the word countable.

(Refer Slide Time: 14:24)









So, example: so N itself is countable. So, if it is not clear, so far; so any set has a bijection to itself. So, this is N; I can have an identity mapping. So, the first element is 1, the second element is 2, the third element is 3 and so on. So, N is obviously N has a bijection to itself, you map every element to itself; so all elements are covered. Now, consider a slightly different thing. Consider the set of even numbers in N; it is 2, 4, 6, 8. So, on first glance, we may think that we cannot have a bijection; because clearly this set, so this is N. This is l even numbers in N; this is only half the numbers in N. We are only collecting the even numbers in N.

So, you may think that you cannot have a bijection; but that is wrong because I can have a bijection like this. I can map an element 1 to 2 and element 2 to 4, element 3 to 6, 4 to 8 and so on. So basically, every element i is mapped to two times i. So, 2 is the first element, 4 is the second element, 6 is the third element and so on. And I can list down all the elements in the second set; maybe I will call it E. So, all the elements in E are mapped like this; and it is clearly one-to-one; no two elements x and y in N get the same image. And it is also onto; no element in E is missed out. So, this is something that is surprising with infinite size sets.

So, even though we thought we were dealing with a set that is half the size of N. It happens to be of the same size as N. Same size as N, in the sense that there is a bijection with N; and that is going to be our definition for two sets having the same size. Clearly it works for finite sets. But, what we are going to say is that even for infinite sets, we will say that two sets are the same size, if there is a bijection from bijection between them. So, E and N have the same size, even though E is like half the elements of N; so this is something that happens with infinite sets.

So, maybe now the next thing is consider n squared; so, by n squared, I mean like the grid; so this is 1, 2, 3, 4. So, I am talking about all the pairs so by n squared, I meant all the pairs x, y, where x and y are in N. So, I can talk about the grid kind of thing like that what we have here; these are the grid points are the elements of N squared. So, now this looks like even more, it looks much bigger than N; because the number of points in N squared seems to be the square of the number of points in N. Because, it is like it is a Cartesian product with itself; and Cartesian product usually multiplies the size. In fact, so one may think that N squared is of much bigger size, but it is not. We can actually track the elements of N square in this following way.

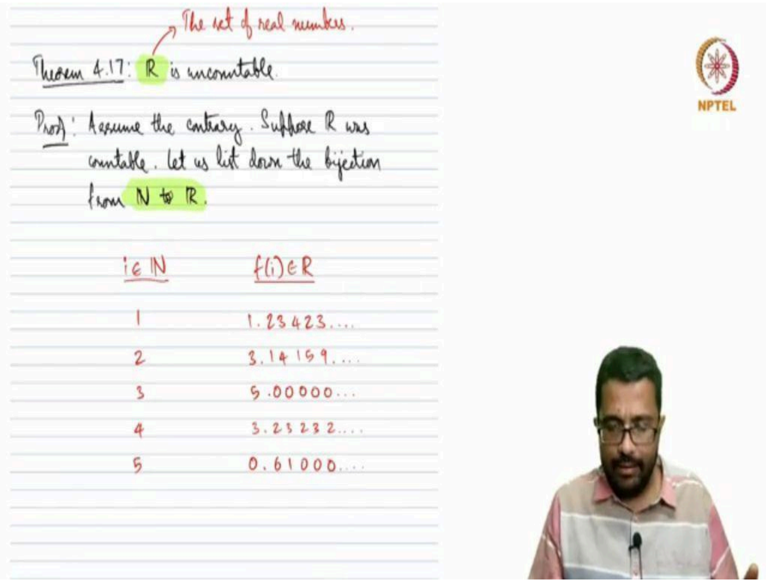
So, you start with 1 or you start with 1-1; then you go to, then you go to 1-2, then you go to 2-2, then you go to sorry. Let me just say it 1-1, then you go to 2-1, then you go to 1-2, then you go to 2-2. Did I do it? No, I did not do it; so you can do this kind of thing. So, you can start with 1-1, then you go to 2-1, then you go to 1-2, then you go to 1-3, then 2-2, then. So, basically what you are doing is you first take all the pairs which add up to 2, which is this 1-1, then take all the pairs add up adding up to 3, which is 1-2 and 2-1. Then, take all the pairs adding up to 4, which is 1-3, 2-2 and 3-1. Pairs adding up to 5, which is like 1-4, 3-2, 2-3 and so on; and pairs adding up to 6 and so on.

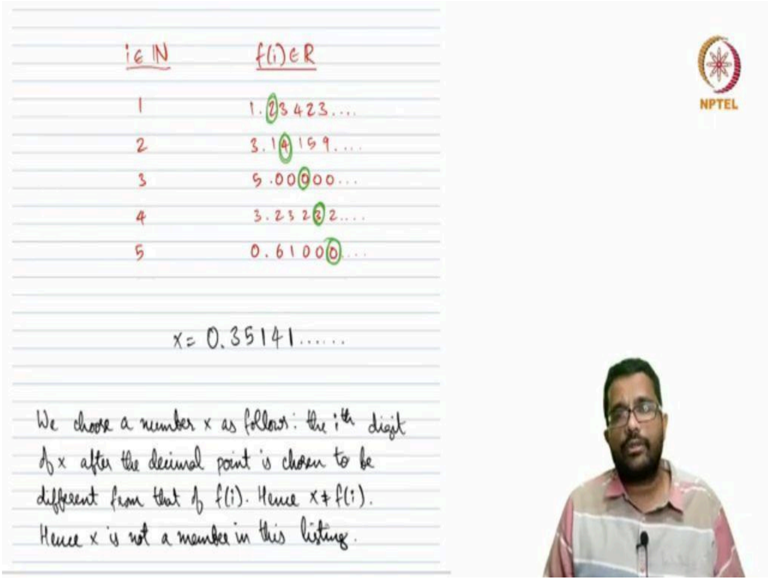
So, basically you are covering from the closest point closest to the origin in ways like this. So, this way all the points are covered and no two points get the same mapping; hence it is one-to-one and onto. So, with infinite sets, even sets that are seemingly this important, seemingly bigger can be the same size; this cannot happen with finite sets. Then, the next thing I want to talk about is integers; let say Z is the set of integers. So, this is the set of integers; so you have 0, 1, minus 1, 2, minus 2 and so on. And in fact, there is a clue as to why it is countable just in the way that I am writing it.

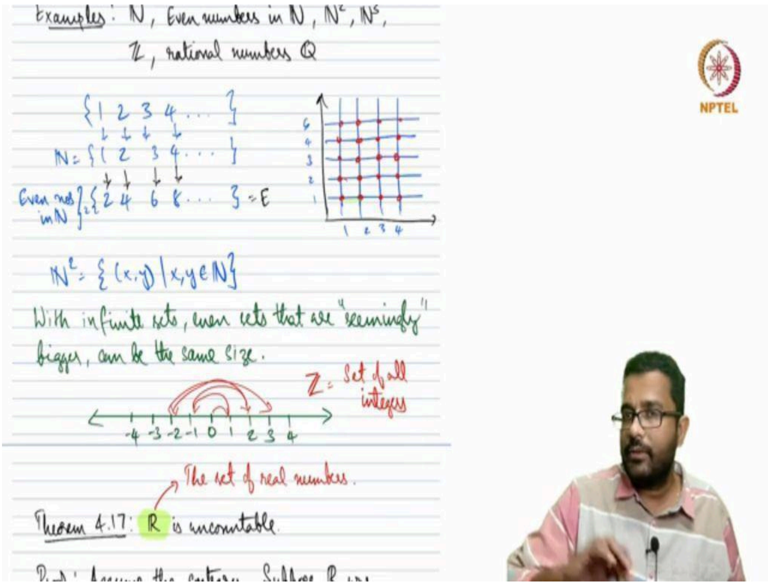
You could start from, you could, you just want to check like map everything. You start from 0, then you go to 1, then you go to minus 1, then you go to 2, then you go to minus 2. So, you can start from the centre and kind of spread out, 3, minus 3, 4, minus 4 and so on. So, all the positive as well as negative numbers are all positive as well as negative integers are all covered. So, Z is also the set of integers that is also countable; so Z is a set of integers, set of all integers; not just natural numbers, positive as well as negative. And in fact, the set of all rational numbers is also countable. I am not going to explain, but it is similar to the N square thing; because rational numbers can be thought of as like a ratio of two integers p by q, p and q.

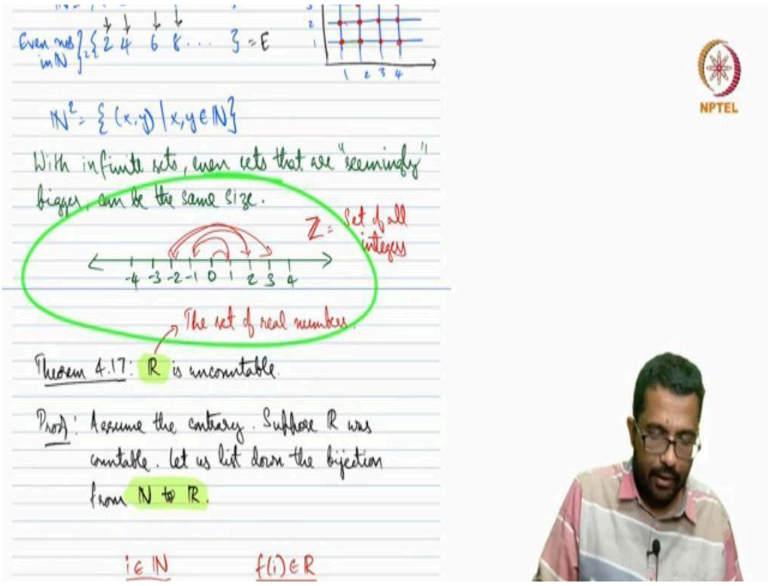
And then you can enumerate p and q as like this as a grid or something, and then you can count them one by one. So, natural numbers are countable, even numbers in N are countable, even though it is seemingly smaller; N squared is countable, even though it is seemingly bigger; the set of all integers is countable, and the set of all rational numbers is countable. So then, I already said that being countable is a sort of a special property. But so far, whatever we saw has all been countable. So, one may I naturally ask, show me something that is not countable.

(Refer Slide Time: 22:49)




i ∈ N   f(i) ∈ R
1   1.23423....
2   3.14159....
3   5.00000...
4   3.23232....
5   0.61000...
x = 0.35141......
We choose a number x as follows: the i th digit of x after the decimal point is chosen to be different from that of f(i). Hence x ≠ f(i).
Hence x is not a member in this listing.
NPTEL


Examples: N, Even numbers in N, N², N³,
Z, rational numbers Q
{1 2 3 4 ... }
N = {1 2 3 4 ... }
Even nos in N {2 4 6 8 ... } = E
1 2 3 4
N² = {(x,y) | x,y ∈ N}
With infinite sets, even sets that are "seemingly" bigger, can be the same size.
Z = Set of all integers
-4 -3 -2 -1 0 1 2 3 4
The set of real numbers.
Theorem 4.17: R is uncountable.
NPTEL



So, the set of all real numbers, the set of all numbers on the real line, this is uncountable. And it is not a very difficult thing to see. But, it is a very clever idea; it is a, so just think about it. You will have like I said about decidability, like to show that something is undecidable, you have to rule out all possible Turing machines. To show that something is uncountable, you will have to say that whatever way you try to list it, you try to map; this is the first element, this is the second element, this is the third element. Whatever way you try to list down the elements of R, you cannot list them down like that; some somehow or the other some element will be missed out.

For other sets that we saw, like N square or Q or Z; we were able to list down all the elements of that set without missing out on any of them. But, for R to show that R is uncountable, R is a set of real numbers. You have to show that whichever way you try to list down the elements of R, you cannot capture all of them. So, the way the proof goes is the following; so you assume that the statement is false. Suppose R is countable. So, then that means that if R is countable, there is a bijection. So from the bijection, we will get a contradiction. So, let there will be a bijection from N to R, so from the natural numbers to real numbers.

So, here we list down the bijection; let say the first number is 1.23423. So, I am putting multiple decimal; it could be in general, it could be an infinite. It could be an infinite thing after the infinite sequence after the decimal point; let say 2 is 3.14159. So, that is kind of the number pi; 3 is 5. The third number is 5, just the number 5; so I just put 0000. Suppose, we can list down the bijection in some manner; so, now from this we will derive a contradiction. Suppose, R was countable, what does it mean? It means that you can list it down like this. Now, we will show that whatever be the listing; so this is some arbitrary listing. We can get some number that is missed out by this listing. So, when we showed that something was countable, like Z was countable.

So, when we showed that Z was countable, we are actually capturing every number in Z, starting from 0, 1, minus 1, 2, minus 2. All the (times) if you ask me, where does this 100 come from? I can tell you 100 is going to come at this position; whereas here, we are not that is not going to happen. We can show that numbers are going to be missed out, so let us see how. So, let us construct a number x; so the speciality of the number x is this. So, what I will do is that, so the first x is going to be 0 point something. And it is going to be something between 0 and 1; and what it will do is this. So, look at the first number after the decimal point of the first number, first digit after the decimal point of F1, which is 2.

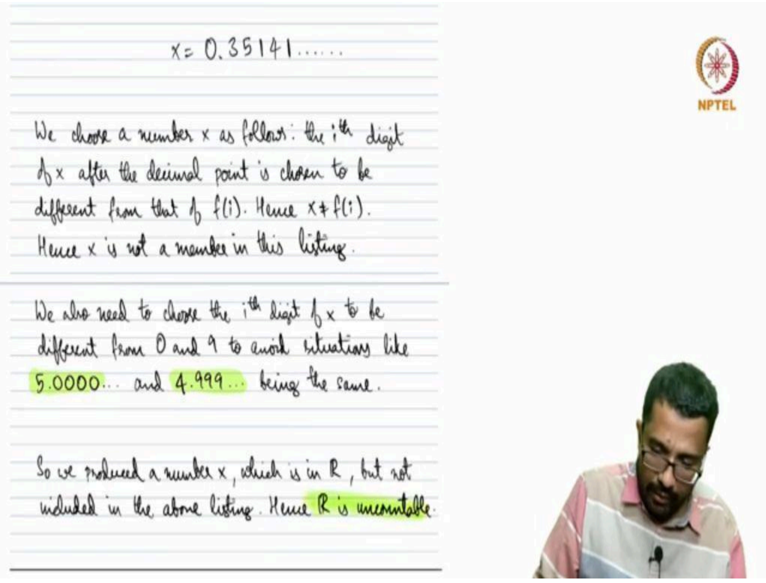
Look at the second digit after the decimal point of F2, look at the third digit after the decimal point of F3, fourth digit after the decimal point of F4, fifth digit after the decimal point of F5. So, what we will do is when writing down x, so I know the first digit of F1 is 2; so I will write 3 here. So, the goal is that x is going to be different from f1, because the first digit after the decimal point is different from the first digit after the decimal point of F1. Now the second digit of the second number is 4, so I will write 5. Third digit of the third number is 0; so, let me write 1.

Fourth digit of the fourth number is 3, so let me write 4. Fifth digit of the fifth number is 0, so let me write 1; so I have written something. So, now the way we constructed x, so the first bit after the decimal point is 3; now, that makes it different from F1. Because in F1, the first bit after the decimal point is the first digit after the decimal point is 2; so, x differs from F1. x differs from F2 in the second digit after the decimal point; it is 4 here and is 5 here. X differs from F3 in the third digit after the decimal point and so on. So the way it is constructed, we are making sure that x is not the first number in this listing, x is not the second number in this listing, x is not the third number in this listing and so on.

So, x are infinitely constructed, I will keep filling up. So, this is a theoretical listing that we have; or if R was countable, there is a listing. And now x is also, x also has infinite numbers, infinite digits after the decimal point. So, now you ask me is x the same as any number in this listing? No, because you take the 50th number in this listing. Now, x will differ from the 50th number in this listing, in the 50th digit after the decimal point. x will differ from the 100th number in this listing, in the 100th digit after the decimal point. So, x is going to be different from all the numbers in this listing. So, which means we have discovered an issue in this listing.

If this listing indeed covers all the real numbers, how come x is not captured? So, that is the contradiction. So, the fact that we are able to produce a number which is not in the listing is a contradiction.

(Refer Slide Time: 29:20)



2	3.1(4)159....
3	5.00(0)00...
4	3.232(2)2....
5	0.6100(0)...

$x = 0.35141......$

We choose a number x as follows: the i^{th} digit of x after the decimal point is chosen to be different from that of f(i). Hence $x \neq f(i)$. Hence x is not a member in this listing.

We also need to choose the i^{th} digit of x to be different from 0 and 9 to avoid situations like 5.0000... and 4.999... being the same.

→ Mapping $f: S_1 \to S_2$

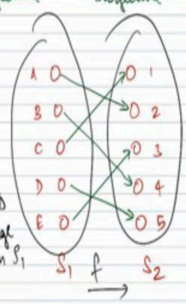
Bijection: One-to-one and onto

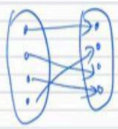
Injective Surjective

Called Correspondence by Sipser

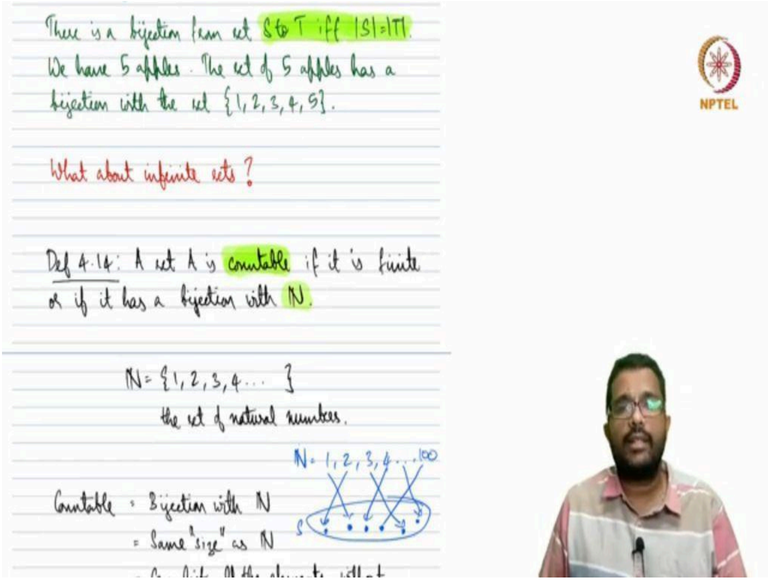
One to one: $f(x) \neq f(y)$ if $x \neq y$.

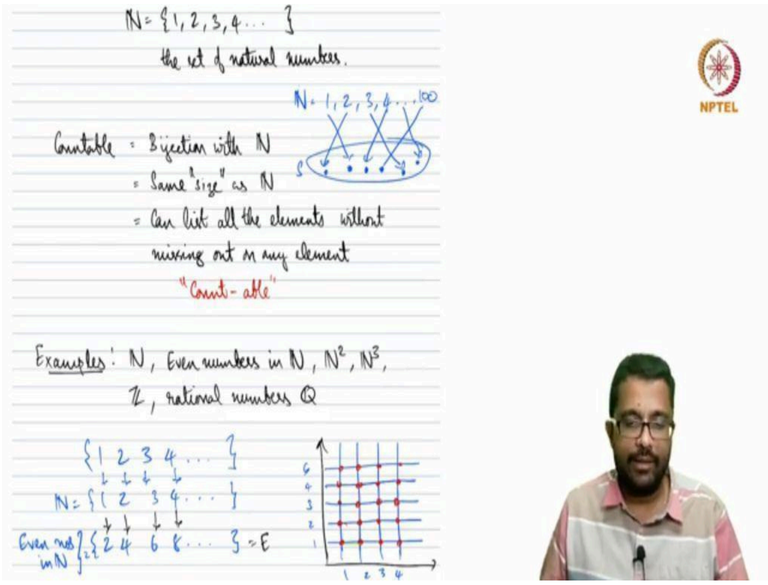
Onto: All the elements in S_2 are an image of some element in S_1

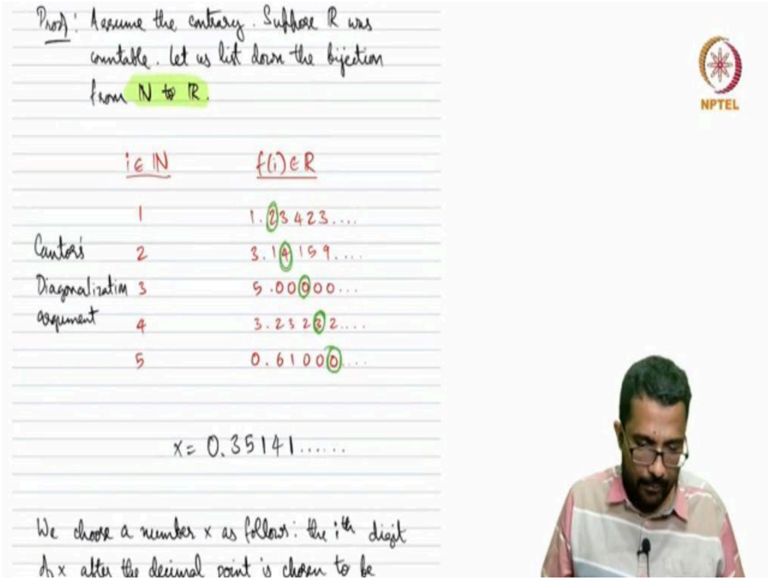




There is a bijection from set S to T iff $|S|=|T|$.


There is a bijection from set S to T iff |S|=|T|.
We have 5 apples. The set of 5 apples has a bijection with the set {1,2,3,4,5}.
What about infinite sets?
Def 4.14: A set A is countable if it is finite or if it has a bijection with N.
N = {1, 2, 3, 4 ... }
the set of natural numbers.
N = 1, 2, 3, 4 ... 100
S
Countable = Bijection with N
= Same "size" as N
NPTEL


N = {1, 2, 3, 4 ... }
the set of natural numbers.
N = 1, 2, 3, 4 ... 100
S
Countable = Bijection with N
= Same "size" as N
= Can list all the elements without missing out on any element
"Count-able"
Examples: N, Even numbers in N, N², N³,
Z, rational numbers Q
{1 2 3 4 ... }
N = {1 2 3 4 ... }
Even nos in N = {2 4 6 8 ... } = E
NPTEL



And one small thing to note is that sometimes we could have two numbers with different digits equal to the same number, let say 5.000 and 4.999. This actually denote the point 999 actually; 4.999 actually, it is equal to 5. So, when I am choosing the digits of x, I just take care to avoid 0 and 9 which is okay; because we have 10 digits, so which I have taken care here. So, the digits of x are not going to be 0 or 9, so that this situation does not arise. I do not want if I choose 999 or something in x, then it may accidentally be equal to some other number in the listing; I do not want that. I want x to be distinct from all the numbers in the listing. And so we choose every whenever we choose i^{th} digit of x, we choose something which is not 0 or 9; and also which is not the i^{th} digit of the i^{th} number.

So, we produced the number x, which is in; so x is clearly a real number. So, x has infinite digits after the decimal point, but it is a real number. But, it is a number that does not feature in this listing. So, which means it is, which means the listing is not thorough or listing is not proper. So, hence, it means that the assumption that R is countable. If R was countable, then all numbers would be listed. So, this means we produce a number that is not in the listing, which means R is not countable; hence, R is uncountable. So, I think we have already reached half an hour; so, I think I will cover the next topics in the next lecture.

But, just to quickly recap what we saw in this lecture, we developed, we defined what is bijection. Bijection is a mapping from a set to another set, which is both one-to-one and onto. And the reason for using bijection is that in finite sets, we can have a bijection from one set to another if and only if they are of the same size. Now, we want to extend this notion to infinite sets. So, a bijection is often if there is, if a set has a bijection with the set of natural numbers;

we call that set to be countable. And now, we saw that the set of all natural numbers; obviously, it is countable, because it has a self bijection.

Then, we saw that the set of even numbers, even natural numbers, set of N squared, where n is a set of natural numbers, the set of odd integers, set of all natural numbers; sorry, set of all rational numbers; these are all countable. However, the set of all real numbers is uncountable and that we use this clever argument, where we list down, we assume a listing, and list down all the numbers; and we choose x to be different from each of the numbers in the listing. So, one point that I missed to say earlier is that this technique, this proof technique is called Cantor's, George Cantor, c, a, n, t, o, r diagonalization argument; so this technique is called diagonalization.

So, it must be evident why it is called diagonalization because we have the listing, and then we choose x so that it differs from all the diagonal entries; the first entry in the first number, second entry in the second number, third entry in the third number and so on. We choose x to be different from all the numbers in the diagonal. So because of that, it is called the diagonalization argument. Hence, R the set of all real numbers is uncountable. And with that, we complete lecture number 35. And in lecture number 36, in the next lecture, we will continue to build towards the theory of undecidability.

We will try to build, we will first show that there are undecidable languages; then we will actually see which one it is. In fact, we already told which one it is. It is ATM acceptance problems in Turing machines. Anyway, that is all for lecture number 35. See you in next 36.