Mathematical Logic
Prof. Arindama Singh
Department of Mathematics
Indian Institute of Technology, Madras

Lecture - 42
Goedel's Incompleteness Theorems

By the time of 1900, there are many paradoxes coming from many sources, including Russell's paradox, Buralliforti paradox; so many paradoxes are there. Then there was the doubt whether whatever we are doing, are having some inconsistency in between them. Say I am working in natural numbers, does there exist one inconsistent or unsatisfiable statement, which we are not aware of? What did it present; it is deducible inside the natural number system; so this doubt came.

Then in 1900 when Hilbert gave the programs for the whole century to work out; there is 23 problems; among those there were two concerning to this. One was this, whether natural number system is consistent or not; it was pertaining to this question. And the other question was decidability of Diophantine equations. These two problems were taken up by logicians; they were easier for them to handle. And Goedel solved this problem of consistency or inconsistency in natural number system. And Turing solved the other problem that first order logic is undecidable. Therefore, Diophantine equations also cannot be solved. But he has not finished it. It took another 35 years to show that there is no algorithm to solve the Diophantine equations. This was the story behind it.

Then what happened in 1931 when Goedel proved these theorems, it is told that; there is a story; at that time Hilbert was drinking, and he just threw the glass and smashed it being angry; because he has started a big program for proving the consistency of arithmetic. And that was called the Hilbert's program. So, now Hilbert's program is settled. That was the impact and it was really considered as a pinnacle of achievement. It is an intellectual achievement of 20^{th} century; that is how the results were considered. It is difficult really to complete it in one or two lectures that we have. But then this will be the last lecture. So, it is a very fitting lecture, for, that we must discuss about it.

The problem is whether to find completeness in the natural number systems and whether to prove consistency in the natural number systems. We have so many conjunjectures along with paradoxes. We have conjectures also.

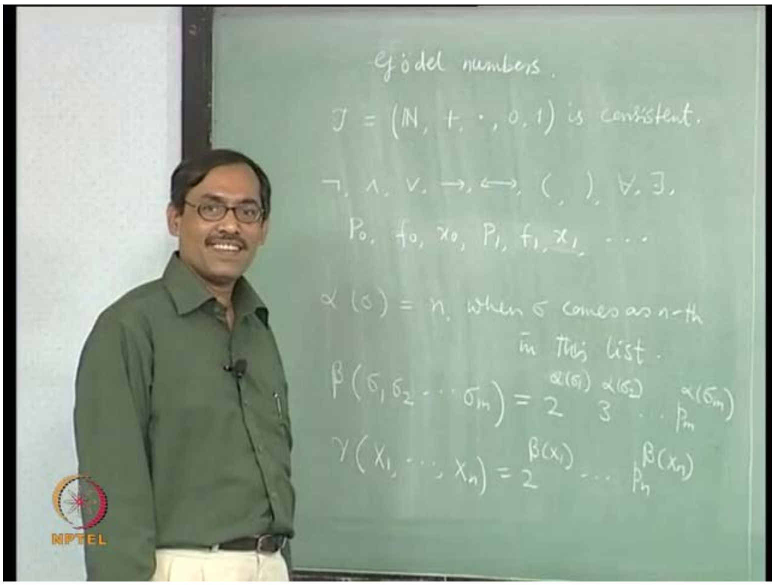
Now also there are some conjectures in natural numbers, which are remaining. Like our Goldbach conjecture, which tells that every even number can be expressed as sum of two primes. Every even number bigger than two. Except 2 all of them can be expressed. So these are certain things, which really bother mathematicians. Then Goedel really proved that well, there will be some conjectures which will remain unproved. That was basically the incompleteness theorem, that there will be true sentences in natural number systems, which will have no proofs inside the natural number system. And whatever theory you have described for the natural number system. But it is not so exactly, because there is some constraint, like natural number system with only operation as addition is a decidable theory. So, anything you conjecture can be proved within that. But once multiplication is there, it is no more decidable. Then there comes the question of this completeness.

We will see slowly; how this progresses. I should give a disclaimer in the beginning that the formal proof, anyway, we will not be able to do within even five lectures; it needs some more machinery. But then I will give an outline and it is almost formal. So, every part I will tell you, where its formality is lacking and how it can be done.

Gödel's proof really had two ingenious apparatus. One is, he says, that within the natural number system you can express something is provable or not. That is a big thing. Something is provable; you can say some predicate of natural numbers is true or false; that we can express it. Second thing is, he has achieved some fixed points; in the sense that suppose, you have one unary predicate. Let us say not exactly predicates, formulas having one free variable, they can be considered as unary predicates. Then suppose you consider B of x, or say B of y, as one of the predicates; then he says that there exists a sentence X in the natural number system so that X if and only if B of X is provable. This is what is called Diagonalization Lemma, which he proves. These two are the main things what he does.

So, B of X means he has to really convert every sentence as a number; then B of n, not really B of X, that is what he is going to do. We will see slowly how it progresses. The first apparatus will start with, which is called the Goedel numbers. What he does here, is just coding everything of first order logic, or even second order, as a natural number; that is what he wants to do. Before this he starts with one assumption that suppose we have one theory of arithmetic, theory of natural numbers, which is having also multiplication in it. Some reasonable theory he assumes; we are not formalizing, again, like your Peano's axioms or something. Then we have a theory say N, and we have plus and we have into; we may have 0 and one also there. Suppose we take a theory having this minimum things, then there can be other operations, predicates, functions, you can define from these things.

(Refer Slide Time: 05:50)

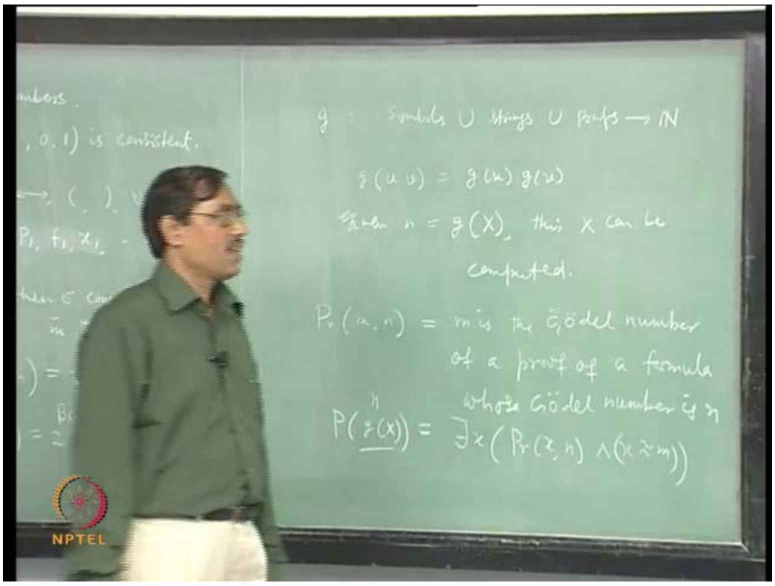


Suppose this theory is consistent. See, the main thing is if it is inconsistent then everything will follow, so it is not interesting. Let us start that we have one theory of arithmetic, which is consistent. Let us call it T; that is the assumption; from there only everything is starting. So, what he wants to do is, you give any formulas, any statements or anything in the first order logic, or even second order where, for all over the predicates are there; but he only goes for the first order formulation of arithmetic, where you are not using for all P and others, but that can be included in the theory. Now what he does is, he encodes them as natural numbers; that is the first thing. So, how to give numbers to the symbols? He just gives a scheme. There are many schemes now, but we will go to the original, whatever he has done, by using the prime factorization there. What he does, you just list all the symbols, may be these things will be using, and there can be some predicates, some functions symbols and variables. So, you may assume that predicates are finite in number, but even if you do not assume there is a way to give the numbers.

So, first is P 0 next f 0, P 1, f 1, P 2, f 2,... If you have variables then you can post them in between say P 0, f 0, x 0, and so on. Make a list of all these things. So, P's can be infinite in number, f’s can be infinite, x's can be infinite in number, predicates function symbols and variables; then what happens? Suppose you take the n-th symbol here; so it is a list, an ordered list. Now define alpha of symbol equal to n, sigma comes as n-th in this list, so somehow numbers have been fixed. Now, the thing is given any number you can go back to this symbol also.

Then what happens, he introduces beta of string of symbols. So, not only symbols, we need strings, because we want to formulate or form the formulas. We need strings. So he writes it as say, 2 to the power sigma, n of sigma 1, which is alpha of sigma 1, 3 to the power alpha of sigma, pm to the power alpha of sigma m, where, 2 to pm, these are prime numbers in that order; so pm is the m-th prime. Now, you see that property still holds. Then, if you are giving some numbers here, you can do its prime factorization, which is unique. Then find out what are the symbols; sigma 1 to sigma m can be found out. It is constructive; it is computable. So, beta itself is computable and its inverse also is computable. Given a number factorize it; get the symbols.

(Refer Slide Time: 11:49)



Now, we will go to give numbers for proofs also. Any formula would be a string of symbols. Now, we will go for a list of formulas. Suppose you take a proof that will look like X 1 up to X n. That, you give 2 to the power beta of X 1, pn to the power beta of X n. So, take any proof. Now it is a finite sequence of formulas and each formula has a beta. Again, take the prime to the power betas. This is also a computable function; and you can see if you take these numbers prime factorizations; take the prime factorization, get the indices, get the strings again, factorize them, get the symbols, so you can reconstruct. Now, what he does, he calls these entire, not he, I am calling; so, will just unify these entire alpha, betas and gamma as g. So, g will be a function from all the symbols we are using; union of all the strings, which includes your formulas and union all the proofs, to natural numbers. That is how this function is defined, these alpha, beta, and gamma all whichever is applicable for whichever it is. The other way you could have started with g as alpha then extend it to beta then extend it to gamma that is your g.

Now this g has some properties. Like, g is a function from these, that is a first thing; for every symbol, every formula, every proof it is defined. Then if you have g u v, u is a string, v is a string of whatever, it is proofs or formulas or symbols, that will be equal to g of u time g of v because prime factorization is used; 2 to the power something and so on. So, multiplication is there; multiplication is assumed here inside it. So, we are not going beyond arithmetic; and it is constructible. And next way is, from any number given if it is g of something then that thing can be computed. So, given n equal to g of some entity, this X can be computed.

So, now what he asks us to think is that given any such entity you think of that entity along with that number. Whenever we say something is true of that number, you might intuitively think it is true of that thing of which this is the Goedel number. So, next thing what we will do is use this Goedel numbers to define a particular type of predicates. Let us call it say Pr m comma n, in the set of natural numbers. It is really a relation; the relation is being translated to our language as Pr of m n. What we will do for this is, m is the Goedel number of a proof of a formula whose Goedel number is n. It says m is the proof of n, if you can suppress that g into proof itself, m is the proof of n. So, m is the Goedel number of a proof, n is the Goedel number of a formula and that proof is the proof of that formula. That is what this predicate is. Then what you do, define P of say g of x as there exists x, such that Pr of x m and x equal to n. So, this says there exists a proof; it is over the first variable.

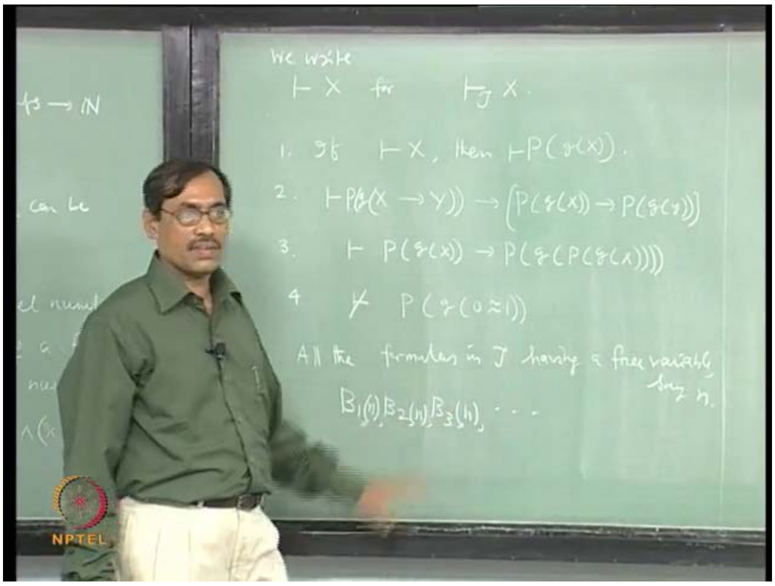
So, there exists a proof whose Goedel number is m and that is the proof of a formula whose Goedel number is n. It simply says that X is provable, so this g X is really equal to n; you could have written as n here instead of g X. One of the variables we have quantified; the other variable remains. It should be P of n; so P of n equal to there exists this; it says that formula X is provable; there exists a proof of X, that is what it says. It means formula X is provable.

Student: Sir, if number is provable?

You have defined; we have defined the Goedel number of a proof. Each formula 2 to the power Goedel number of that formula, 3 to the power Goedel number of this second formula, pm to the power Goedel number of the m-th formula. So, Goedel number of a proof is defined therefore. This predicate says it is a proof of X; X is provable rather. It is not exactly telling that. It is telling P of n means a formula X whose, Goedel number of which proof exists and Goedel number of that X is n that is what it says, intuitively X is provable. So, instead of writing just n, we will go on writing g of X; it will tell us what that X of which we are telling that a proof exists. This is the first order; we have to go around P of g of X.

Before that, when you say that it has a proof, it has a proof where? In our system T. So, always that is our assumption; it has a proof means proof in our system T. We will also write something like: suppose you write the turnstyle symbol for X is provable, X has a proof. It means, it has proof in T; so that means we will write like this. We write this X for proof in T; we will just make this abbreviation.

(Refer Slide Time: 17:59)



Now, let us see what are the properties of this P, the provability predicate. Let us call it provability predicate. What are the properties? The first is, if X has a proof then P of g of X also has a proof. Next, we can say suppose X implies Y P of g of this; this implies P of g of X implies P of g of Y. Once you have a proof of X implies Y and you also have a proof of X, then we have a proof of Y; it is like a modus ponens; from there it is coming. Third is, P of g of X implies P of g of P of g of X. Which means, if X is provable, has a proof, then X is provable is provable; it also has a proof. It is easy to see from this existential formula, but intuitively also it is clear.

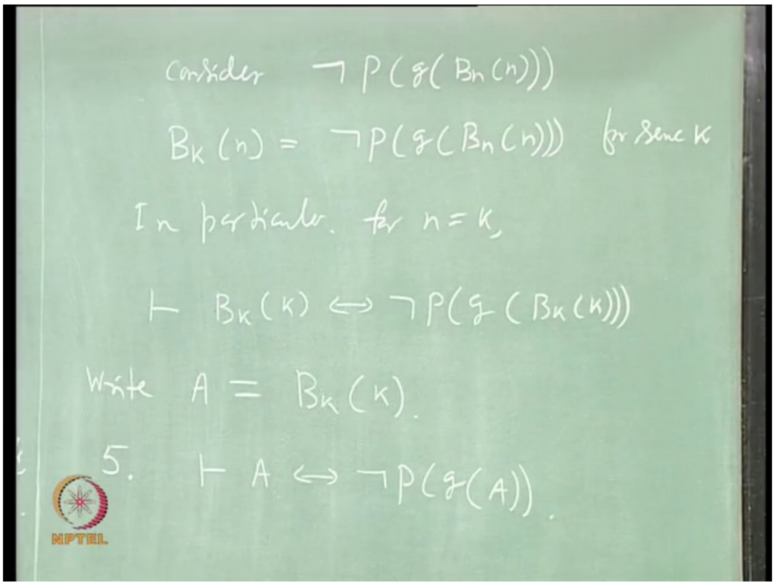
Student: Sir, what is intuitively?

First one, if X has a proof then X is provable; as a proof that is what; basically intuitively means that. Formally, it will come from this, if we just manipulate the symbols. Now, there is one more thing we need. We are concerned with consistency. We can formulate consistency with this provability predicate. For example, at least formulate inconsistency. Let us see what happens. See, 0 is equal to 1 is an inconsistent statement. We will say that inconsistency is written as P of g of 0 equal to 1; inconsistency, 0 equal to 1 is provable; and this has a proof. We will take this formulation as a formulation of inconsistency: 0 is never equal to 1, in natural numbers.

When you say 0 equal to 1 is provable, it is inconsistency of natural number system; that is the interpretation. When you say it is consistent; our assumption is that natural number system is consistent; we should have this. It is not the case that 0 equal to 1 is provable, has a proof; this is the assumption of consistency now. Slowly, we will go for other properties of this provability predicate and something else. Here, I am digressing a bit towards diagonalization, which is a difficult thing to do. I will elaborate it in different way using some paradox.

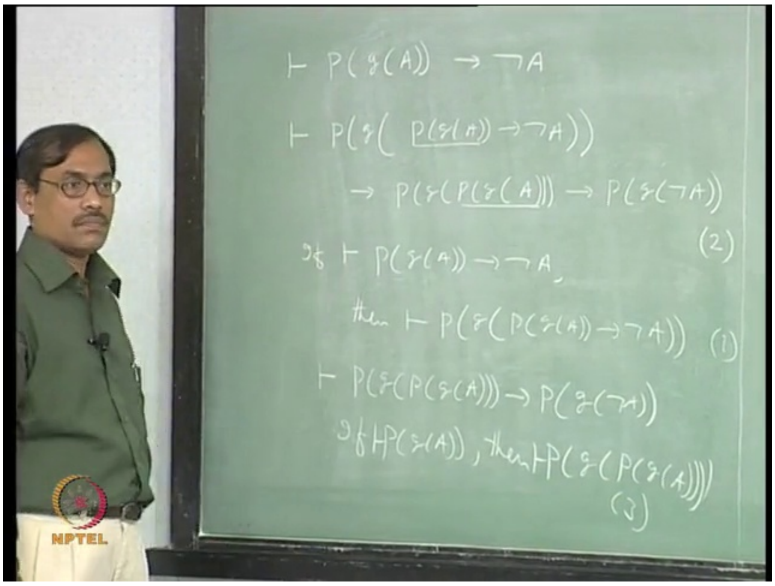
Let us see now. Look at all the formulas in T having a free variable, one free variable exactly, say in N. They all are predicates over natural numbers with one free variable. Now, that set is countable; all the formulas are countable. So, these also are countable. Then we can have an enumeration of it. Let us enumerate. They are say B 1, B 2, B 3, and so on, with one argument each, and so on, this is your enumeration. Now, consider this predicate not P of g of B n f of n. This is a predicate having one free variable. Now, this is one of these predicates in our list; which one we do not know. Suppose it is B k. So, you have B k of n for some k, which k, I do not know; for some k this happens. Now here also B k of n equal to this. Here, n is a free variable. Then you have P implies P as a theorem. So, P if and only if P also a theorem. Then you can take universal closure, for every n. For every n B k n if and only if this side is provable, because that is equal to that if and only if P.

(Refer Slide Time: 23:02)



Now, once it is for every n, in particular, I can take a universal specification, take n equal to k. So, for n equal to k, I have a proof of B k of k if and only if not P of g of B k of k. Let us give some symbol instead of always writing B k of k. I write A equal to that. So, I have now the fifth property, which says if and only if not P of g of A is provable. Now, you have to play with this A.

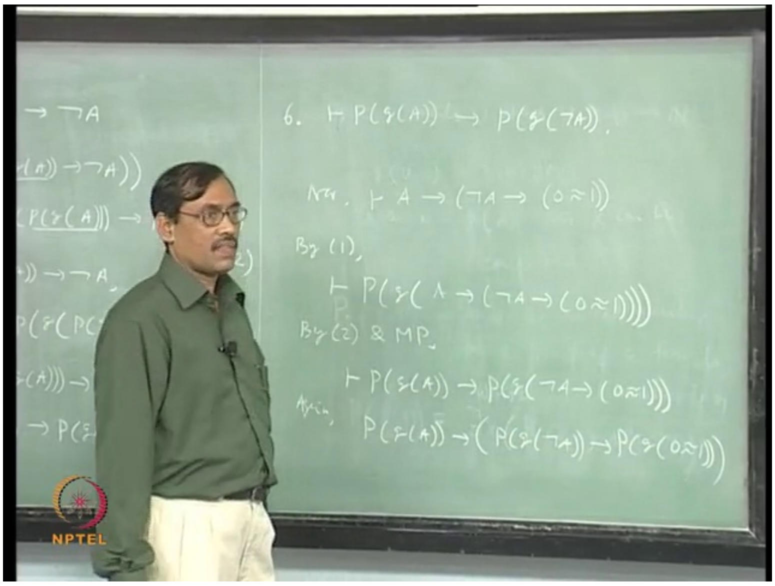
(Refer Slide Time: 25:41)



See, what happens; use contraposition; it says P of g of A implies not A is provable; one part of it; this not of this implies not of this; which is equivalent to A implies not P of g of A. So, I get P of g of A implies not A. Then I have P of g of implies P of g of implies. Look at property two of P. Property two says P of g of X implies Y implies P of g goes there. So, that is what I have done. This thing is already provable; so it takes P of g of here. So, P of g of A comes here; and this one is here. Now I have already P of g of A implies not A, so I take P of g of, apply my second property and get this. I use property one. If this is a theorem, then P of g of whole thing is a theorem. Now use this one. P of g of this and this one modus ponens. That gives P of g of P of g of A implies P of g of not A. Again use 1. 1 again says if P of g of A then P of g of this is 3; not 1 use property 3. Now, these are the same if X then Y and you have Y implies Z therefore, if X then Z.

There is another way. I want only P of g of A. So I can say P of g of a implies P of g of g of A, again 3, use both of the things and hypothetical syllogism. So, you can write that way; that will be easier to see g of A implies P of g of P of g of a by 3. Apply Hypothetical syllogism by that, we get P of g of A implies P of g of not A.

(Refer Slide Time: 29:58)



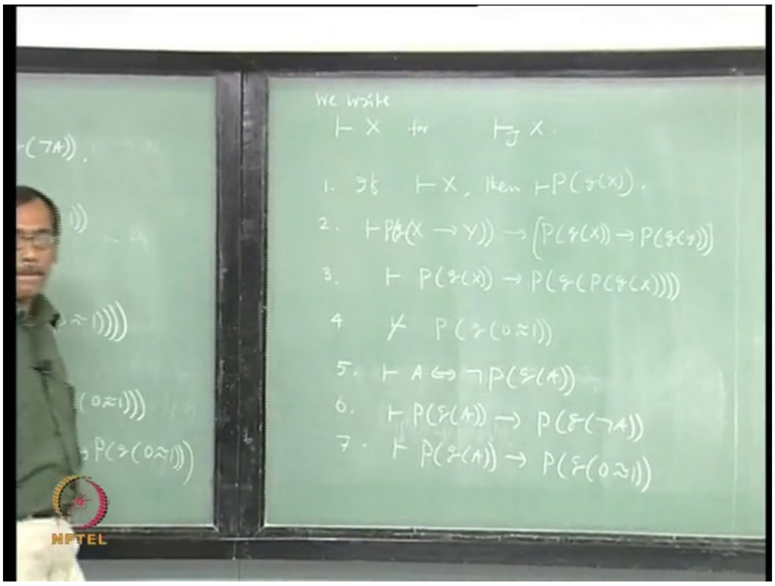
Call this as the sixth property. We are going slowly, developing one by one now. Let us take another, say, A implies not A implies anything, this is the theorem in FL, really. A implies not A implies B, anything you can write; so we start with 0 equal to 1. Now take P of g of that, apply 1. By 1, P of g of this whole thing. Now what happens, use 2. Use 2 and modus ponens; that gives; and modus ponens gives P of g of A implies P of g of not A implies 0 equal to 1.

You have to check each step. This is in the form P of g of X implies Y. So, by 2 and modus ponens you will get P of g of X implies P of g of Y; clear? Once more we can use, because this is also in the form X implies Y. So, if we use once more, then it will come to, once again it gives P of g of A implies P of g of not A implies P of g of 0 equal to 1; just P of g of goes through implication that is what it says.

So, you get this implies P of g of this implies P of g of this. Let us keep on writing what we have got and sixth was. Then we have here P of g of A implies P of g of not A implies P of g of 0 equal to 1. Now look at 5 and 6; this is in the form X implies Y; this is in the form of X implies Y implies Z; your axiom 2 says X implies Y implies X implies Z, X implies Y implies Z implies X implies Y implies X implies Z. By modus ponens you will get X implies Y implies X implies Z and X implies Y is here; so you get X implies Z.

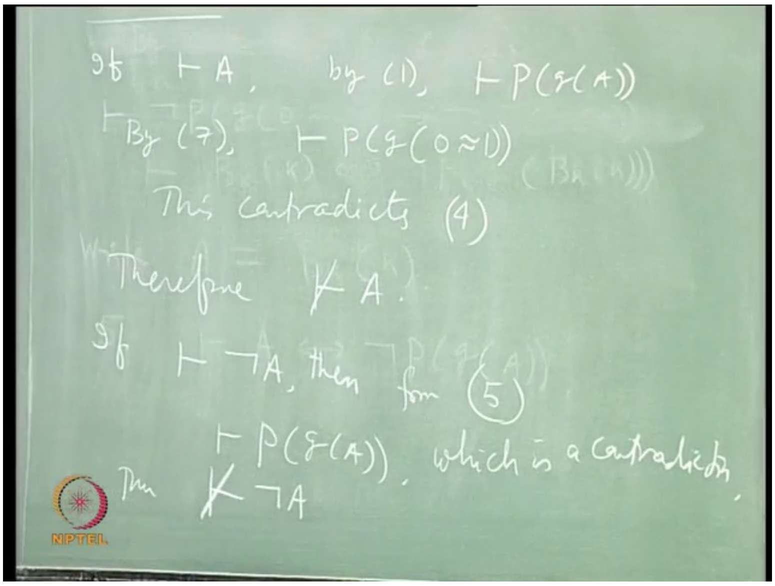
So, using axiom 2, which is distribution of implication and modus ponens twice we have this 7, which says P of g of a implies P of g of 0 equal to 1.

(Refer Slide Time: 34:51)



From this you should get X implies Y implies Z so that gives X implies Y implies X implies Z you have X implies Y as 6. So, you get X implies Z, X is P of g of a Z is P of g of 0 equal to 1, this is what you get. So, by contra-position from 7, we get not of P of g of 0 equal to 1 implies not of P of g of A. Well, there is another shortcut. Instead of contra-position, we can do that, let us see.

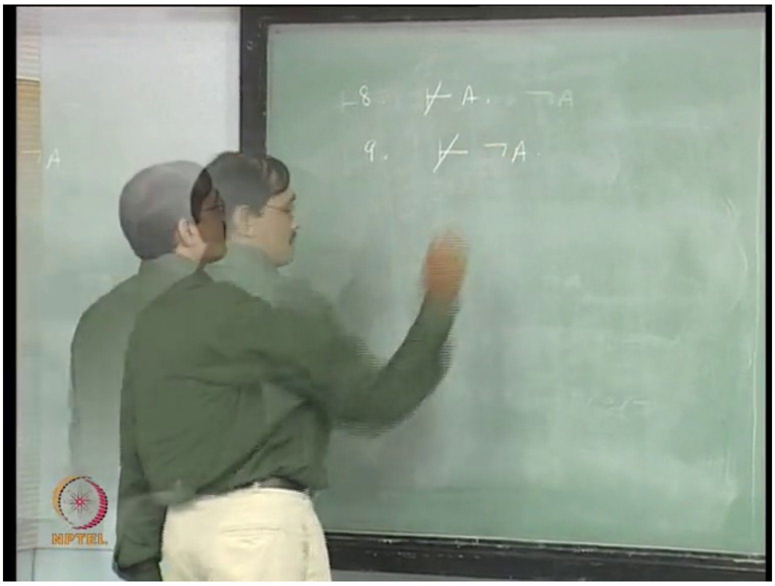
(Refer Slide Time: 35:43)



Now, let us keep this 7. From 7 we will be doing something. If A is having a proof, by 1, P of g of A is having a proof. Now, by 7, it says P of g of 0 equal to 1, this contradict 4; 4 is it is not provable, but you are telling it is provable because of our extra assumption entails A; A is a theorem; that is why. Therefore, what we conclude A is not a theorem not not A.

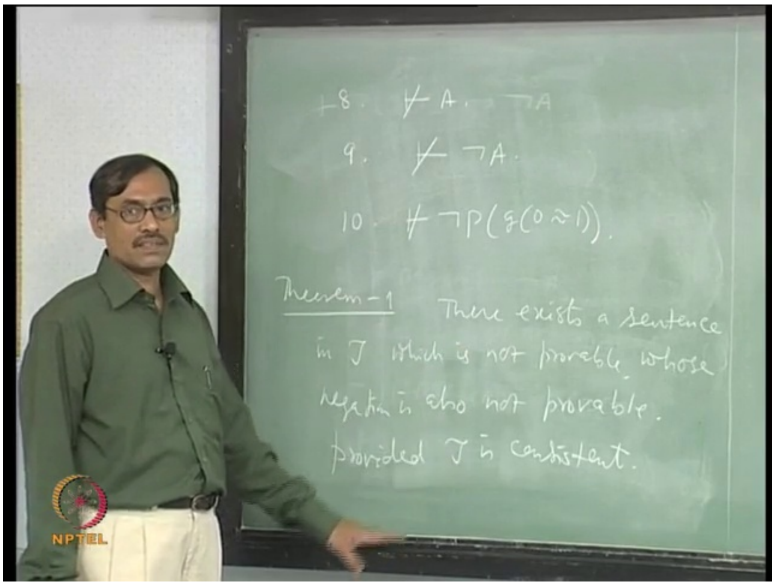
A is a theorem giving a contradiction so A is not a theorem. Now, if not A is a theorem then from 5, P of g of A is a theorem. Not A implies P of g of A, contraposition of the other side, not of P of g of A implies A. Therefore, not A implies A of g of A; one side of that. That gives P of g of A is a theorem. But we have already seen if P of g of A is a theorem there will be a contradiction, which is a contradiction. Therefore, not A is not a theorem. That is our 9th.

(Refer Slide Time: 38:25)



I take next line. If not of P of g of 0 equal to 1 is a theorem. Suppose, we start with not of P of g of 0 equal to 1 as a theorem then look at 7. Contrapositions of 7 says, by 7, we have contraposition. And contraposition, we have not of P of g of 0 equal to 1 implies not of P of g of A. So, if this is a theorem then not of P of g of A is a theorem. Then by 5, A is a theorem which again will contradict, which contradicts; let us write 8. So, what we get, it is not a theorem. That is our 10th one; which says and the last one. Proof is over.

(Refer Slide Time: 40:15)

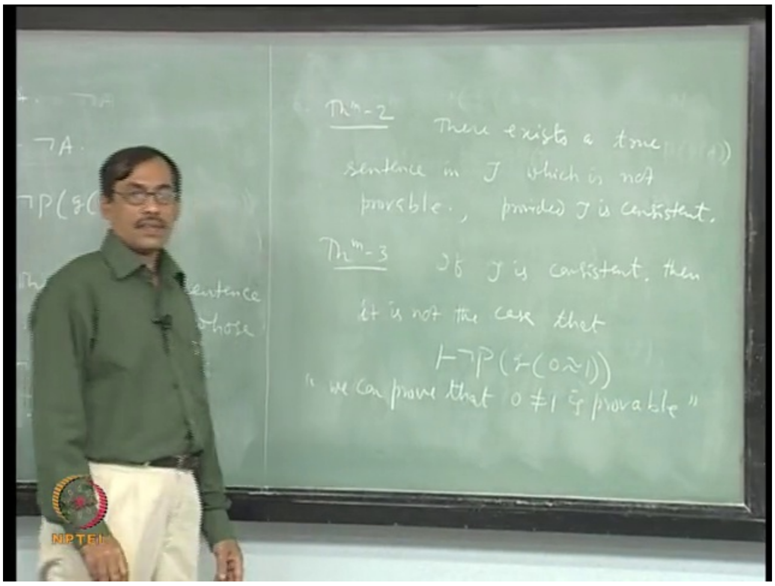


What you have proved? Let us see. These three are the main things we wanted. First, it says there exists a sentence in T, which is our consistent theory of natural numbers, which is not provable, whose negation is also not provable.

Student: But the sentence is right?

We do not know; we have no idea about A; A is that sentence here. So, it only says that provided T is consistent. In other words, we say every consistent reasonable theory of arithmetic is negation incomplete. That is the idea of negation completeness; at least one of them is a theorem, here none of them is a theorem. So, it is negation incomplete. That is your first incompleteness theorem; which is also written another way. We may write it also this way: there exists a true sentence in T, which is not provable. It is look like our earlier proof of x to the power y is rational, when x and y are irrational. See one of A or not A is true, either A is true or not A is true in the natural numbers, but none of them is provable.

(Refer Slide Time: 42:00)



So, there exist, well, a true sentence which is not provable. You can find out also which one is true and so on. But we are not going to deal with any truth here. You can say in that level that one of A or not A is true, but neither is provable. So, there exists a true sentence whichever is it, A or not A, which is not provable. That is basically the idea of negation incompleteness. Then we have the next one, which says that, well, all these things hold provided T is consistent.

So, the next one is if T is consistent, then it is not the case that, this is what your 10^{th} one tells. But what does it say? Can you read this? P of g of 0 equal to 1 is consistency; there is a proof of P of g equal to g of 0 equal to 1 is consistent. So, there is a proof of 0 equal to 1 is not provable, which is taken as the consistency. So, it says if T is consistent then the consistency of T cannot be proved.

See, one thing is 0 is not equal to 1, true, so 0 is not equal to 1 is provable. It is also provable that 0 is not equal to 1 is provable. Now let us look at inconsistency. 0 equal to 1 means inconsistency. If it is provable, provability of 0 equal to 1. Now it is not provable; so this expression is consistency, but not that much only; it is provable we can prove that 0 equal to 1 is not provable, that is an expression of consistency indirectly. So, here when you say we can prove that 0 is not equal to 1 is provable, it also expresses consistency, because there will be proof that one. If you have a different interpretation of consistency, it does not prove that; it only proves this first one. This theorem is interpreted this way. That if T is consistent, then its consistency cannot be proved by the mechanism of T; because all our proof apparatus is from T only. It might be provable by other mechanism. We do not know; not by this mechanism; that is what it says. So, after this Herman Weye said that god exists because mathematics is consistent, and devil exists because we cannot prove it. That is what it says, and there we should end.

So, today only we have seen these three theorems of Goedel. The two theorems. One is negation incompleteness of arithmetic; another is its consistency, we are not able to prove inside the system; both the theorems. When the first theorem says also something else; once you take a reasonable theory of arithmetic, it means you have plus, you have multiplication and then you have the induction axiom, with those you can do most of the natural number things. Once you have taken the induction axiom, it says it is no more in first-order; it is a second order theory, because you have to say for every unary predicate this happens. So, it also proves that second order logic is incomplete; whatever axiom system you make for it, it will remain incomplete. That gives some limit to our thought process itself. That is why it becomes so famous. It is something like limitation of human reasoning. We are not able to go beyond it. It looks, nothing after that we can go; anywhere; that is the importance of the theorem.