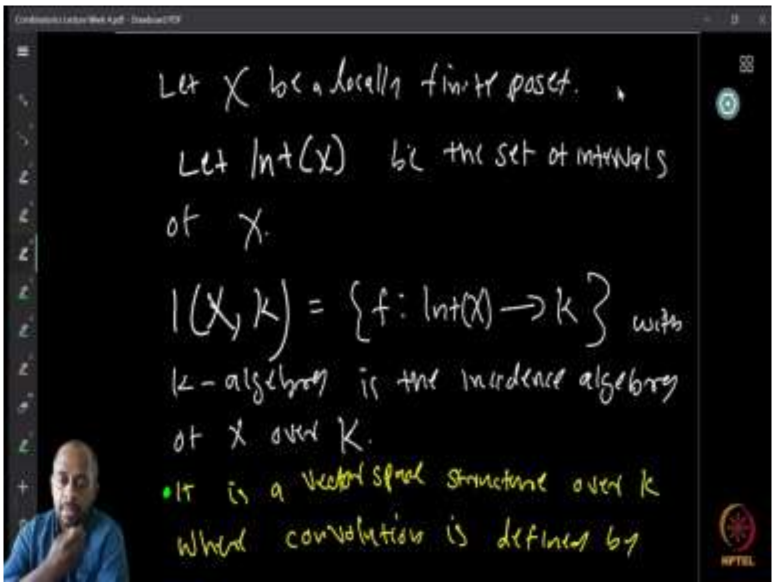
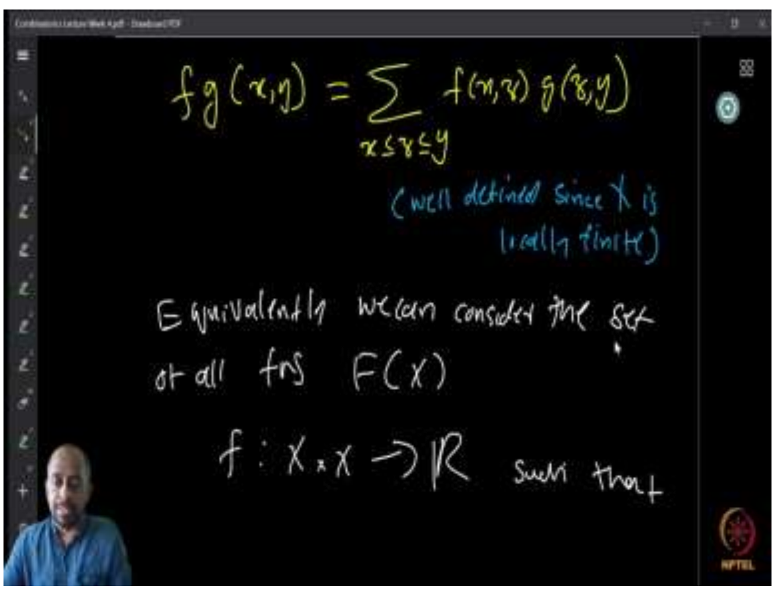
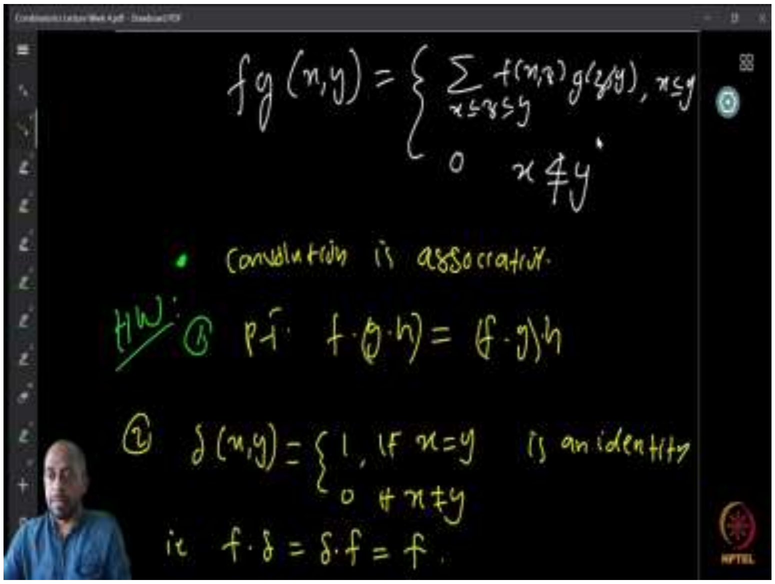
Combinatorics
Professor Doctor Narayanan N
Department of Mathematics
Indian Institute of Technology, Madras
Mobius Inversion Formula

Welcome back to this course on Combinatorics. In the last week's lecture, we took a very short survey of the notions related to partially ordered sets and we hope to use it this week. So, what we are going to do today is to look at a technical Mobius inversion. So this gives us inversion formulas that can be used as we did in the case of inclusion exclusion other topics.

(Refer Slide Time: 01:03)







So to recall we said that a partially ordered set is locally finite if all its intervals are finite. Now let us look at the set of all intervals of a Poset X and denote it by Int(X). Now consider the set of all functions from Int(X) to some field let us say K. So this set is denoted by I(X, K) and together with the algebra that comes from the field the K-algebra this is called the incidence algebra of the partially ordered set X over the field K.

We do not really need this details but just to make it complete I will just mention it so that this I(X, K) is a vector space over K where the convolution is defined by or the product is defined by the formula $fg(x,y) = \sum_{x \leq z \leq y} f(x,z)g(z,y)$

Now, this summation is well defined because we assume that X is locally finite which means that the set of all such z is going to be finite so the sum is finite. So, therefore this product is well defined. Now equivalently we can also instead of looking at this incidence algebra we can just look at the set of all functions F(X) for our purpose.

And the set of functions F(X) which takes X × X to ℝ such that,

$$fg(x,y) = \begin{cases} \sum_{x \leq z \leq y} f(x,z)g(z,y), & x \leq y \\ 0, & otherwise \end{cases}$$

So we are just adding zero here for the other functions where this interval is not. So therefore they are not comparable.

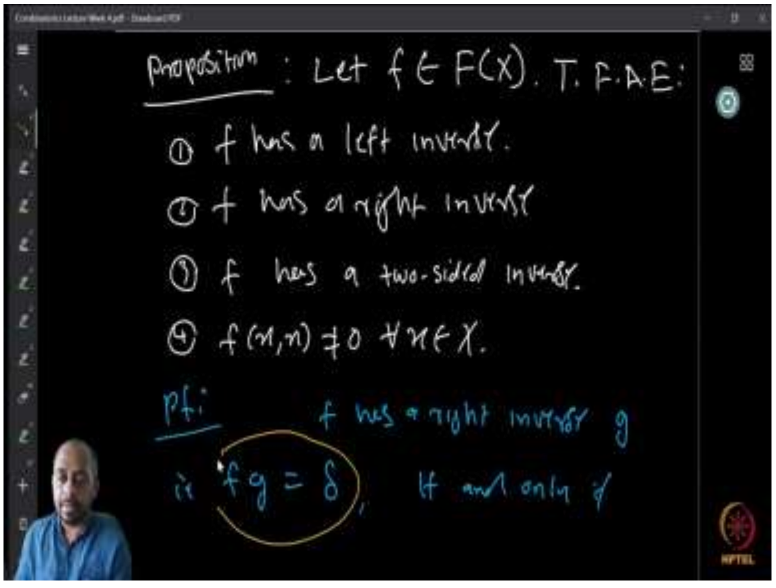
We do not have the intervals, so for all those elements we are defining it to be to be zero. So, these two are essentially the same now if you if you look at this you can show that the convolution, the way we have defined convolution it is associative. So, I want you to do this as a homework so prove that $f.(g.h) = (f.g).h$ that proves the associativity for f, g and h in our F(X) or in Int(X).

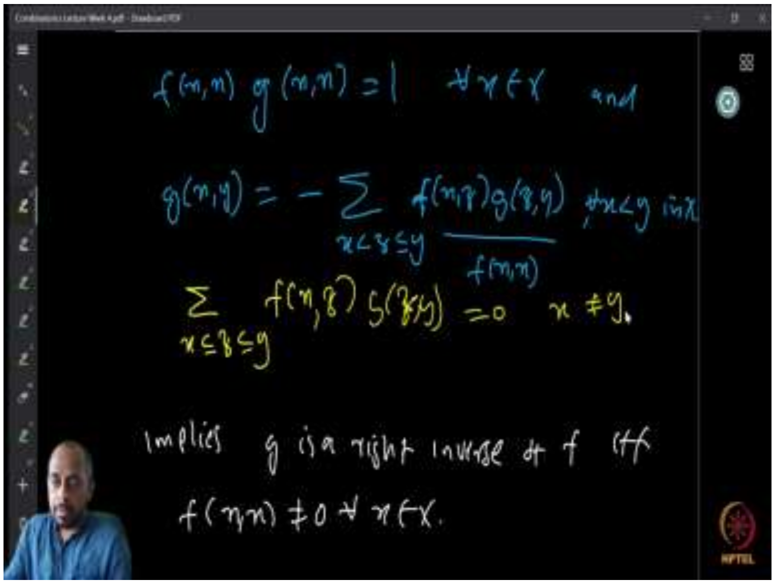
Now again without mentioning it I just assumed the field here to be $\mathbb{R}$, we can assume it to be $\mathbb{C}$ also or in fact any field but for our purposes it will be sufficient to assume it is either $\mathbb{R}$ or $\mathbb{C}$. Now I define another function δ such that

$$\delta(x,y) = \begin{cases} 1, & if\ x = y \\ 0, & if\ x \neq y \end{cases}$$

This is also defined over the poset and this function δ acts as an identity for the convolution. So that is a claim, so again you can prove that by using the definition of convolution $f.\delta = \delta.f = f$. So, therefore it proves that delta function is the identity function.

(Refer Slide Time: 06:00)





Now here is the proposition take any function f in F(X). So then the proposition says that the following statements are all equivalent. The first says that f has a left inverse; second statement says that f is a right inverse, third statement says that f was a two-sided inverse. And finally we say that f(x,x) is a non-zero for every x in the poset X, so all these statements are equivalent, so that is the proposition.

So, how do you prove this? Well let us look at one of the cases, let f was right inverse called g. Now if f has a right inverse g then that means that the convolution product fg is δ by the definition of right inverse and this is true if and only if. So when is fg is equal to delta? Which means that fg(x, x) is equal to $\delta(x, x)$, which is 1 by the definition δ.

So therefore f(x,x)g(x,x)=1 for every x in X, then

$$g(x,y) = -\sum_{x \le z \le y} \frac{f(x,z)g(z,y)}{f(x,x)}, \ \ for\ all\ x < y\ in\ X$$

Now why is this true so we are saying that fg = δ if and only if this is the case or g is the right inverse you finally if and only if this is the case.

Now why is this? Well, just look at the definition of the convolution product, so if you look at the convolution product we have fg is defined to be $\sum_{x \le z \le y} f(x,z)g(z,y)$. Now this is $fg(x,y)$ by definition of the convolution product.

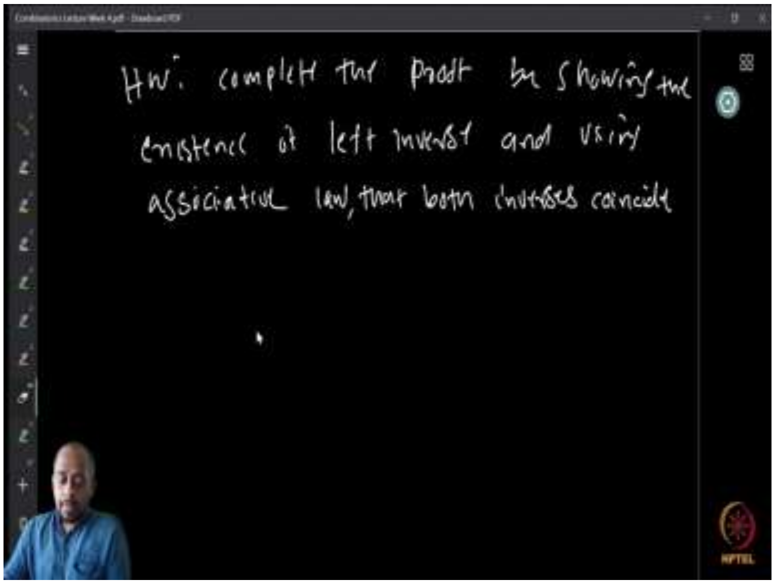
So if you have fg is δ then $fg(x,y) = \delta(x,y)$. But $\delta(x,y)$ is equal to equal to 0 if x is different from y. And therefore if x is different from y this is equal to 0. So in this summation the first term is $f(x,x)$, when $z = x$.

And then $g(x,y)$, so you take that sum that keep that m on the left side and move everything to the right, so what is the remaining thing that is $\sum_{x \leq z \leq y} f(x,z)g(z,y)$So this is what you will have on the right side with the minus sign. Now, so on the left side we had $f(x,x)g(x,y)$, so you take that to the right side you will get $g(x,y)$ to be this way.

So, therefore, if g is an inverse of f, then g should be defined, so if it should be defined then f(x, x) cannot be zero. I mean the f(x,x) is not zero because f(x, x) into g(x, x) is equal to1. And therefore we can do this division and therefore we get g(x, y) this way. So therefore if there is a right inverse then we already know that f(x, x) is different from zero and g(x, y) is in this form.

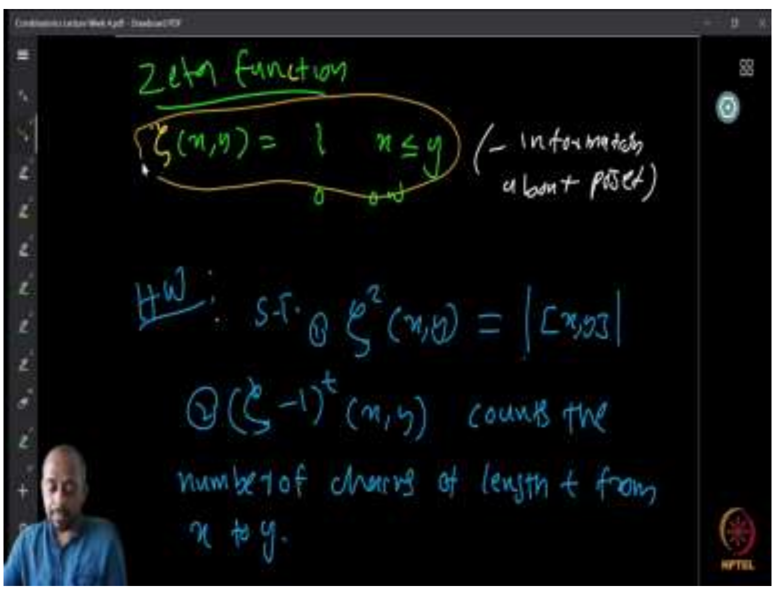
So similarly you can show the other properties, if f has a left inverse then also similar idea works and then using the associativity property we can also show that it has a two-sided inverse. So all these are identical to the assumption that f(x, x) is different from zero because in that case we can of course define g this way.

(Refer Slide Time: 11:46)



So, I want you to do all these computation as a as a homework, not computation, formally writing this as a homework. So complete the proof by showing the existence of left inverse and then using the associative law, so that this inverses coincide, left inverse and right inverse coincide.

(Refer Slide Time: 12:18)



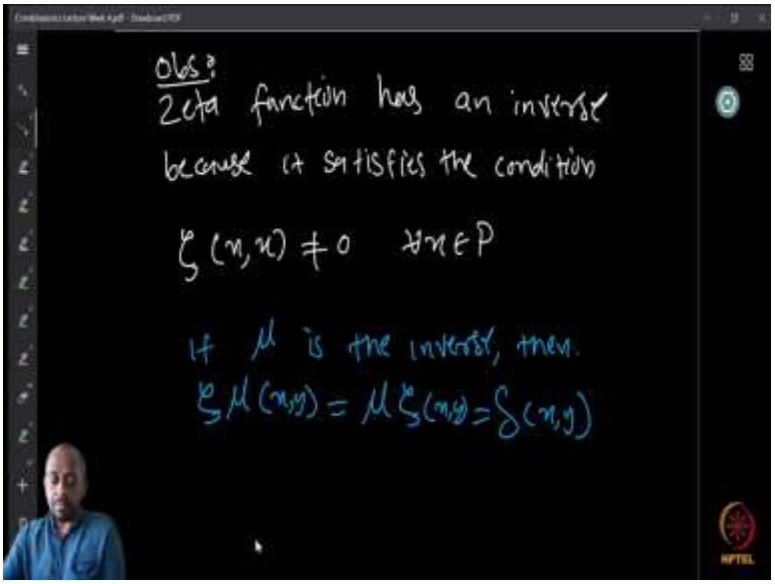
Now we define one more function. So, we already defined a delta function so we are going to define another function zeta function. So the zeta function of a poset is defined as follows

$$\zeta(x,y) = \begin{cases} 1, & x \le y \\ 0, & otherwise \end{cases}$$

Now this zeta function contains a lot of information about the poset we will maybe use it sometime but at the moment I want you to look at this and then try to observe some properties about it. And here are a couple of properties that you can try to prove the first is to show that if you have the zeta function then $\zeta^2(x,y) = |[x,y]|$.

So, this should be quite easy to prove. Then you can also prove that $(\zeta - 1)^t(x,y)$ counts the number of chains of length t from x to y. So look at these two function ζ^2 and $(\zeta - 1)$ one and then look at the properties and then show that these properties are indeed okay so this is a homework.

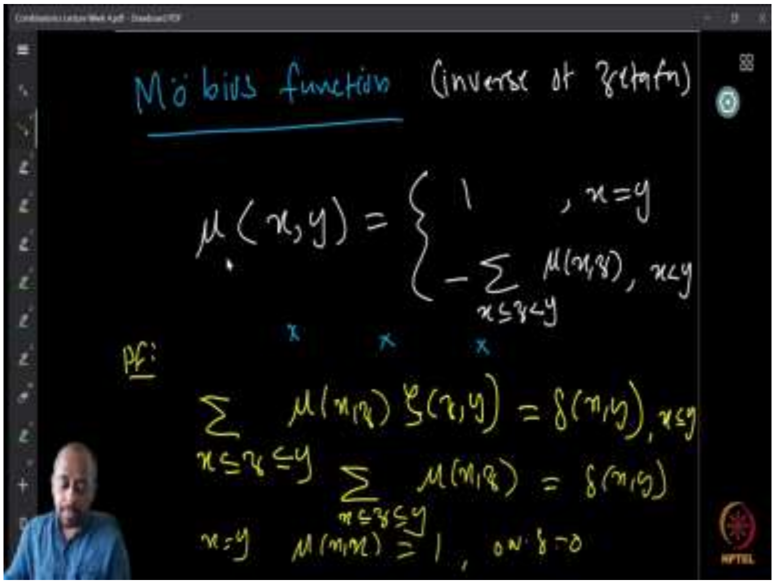
(Refer Slide Time: 14:06)

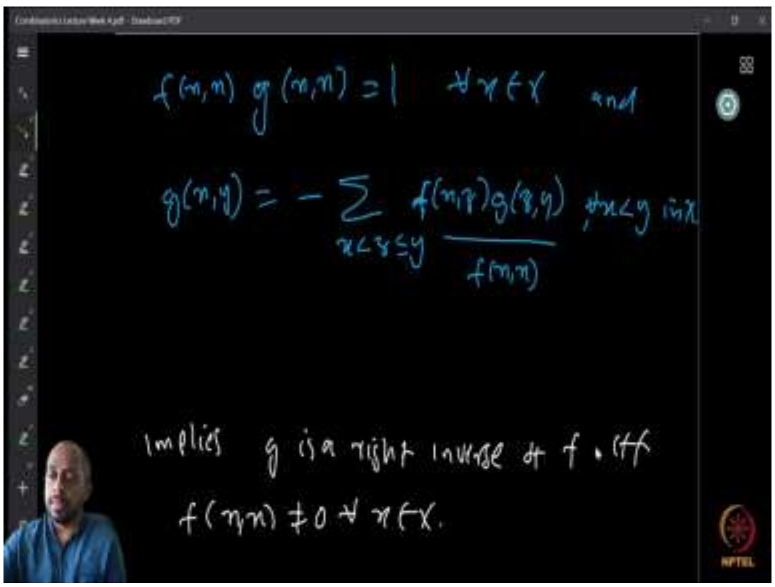


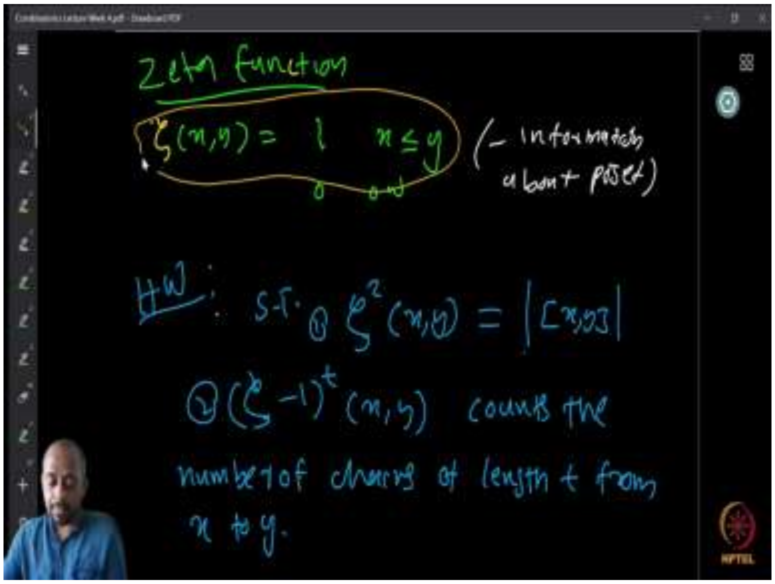
Now one observation is that, the zeta function must have an inverse because by the earlier property where we prove that if the function f(x) is different from 0, then it is invertible. So here we have defined $\zeta(x,x)$ is because x is comparable to itself, $\zeta(x,x) = 1$, so therefore it is different from 0 and therefore the function has an inverse.

Now we define the inverse of zeta function and call it to be mu (μ). So by the definition of inverse $\zeta\mu(x,y) = \mu\zeta(x,y) = \delta(x,y)$, so delta is the identity.

(Refer Slide Time: 15:13)







Now we call μ the Mobius function. So, the inverse of zeta function is called Mobius function and we can see that

$$\eta(x,y)=\begin{cases}1, & if \quad x=y\\ -\sum_{x\le z\le y}\mu(x,z), & if\ x<y\end{cases}$$

Now why is this so? If you look at the definition of the inverse here, so inverse was looking something like $g(x,y)=-\sum_{x\le z\le y}\frac{f(x,z)g(z,y)}{f(x,x)}$.

Now similarly, using that definition we can try to see why this should be the case for the Mobius function. So how do you prove this? So, to prove this, well, we can just look at the definition of the convolution product and the fact that μ is the inverse of ζ.

So, what is $\mu\zeta(x,y)$ is $\delta(x,y)$? So let us look at what is $\mu\zeta(x,y)$? So,

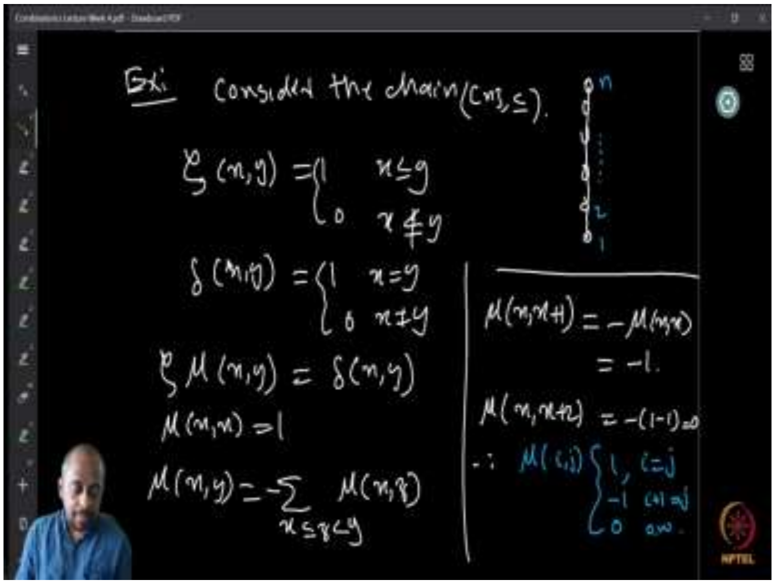
$$\mu\zeta(x,y)=\sum_{x\le z\le y}\mu(x,z)\zeta(z,y)=\delta(x,y), \qquad whenever\ x\le y$$

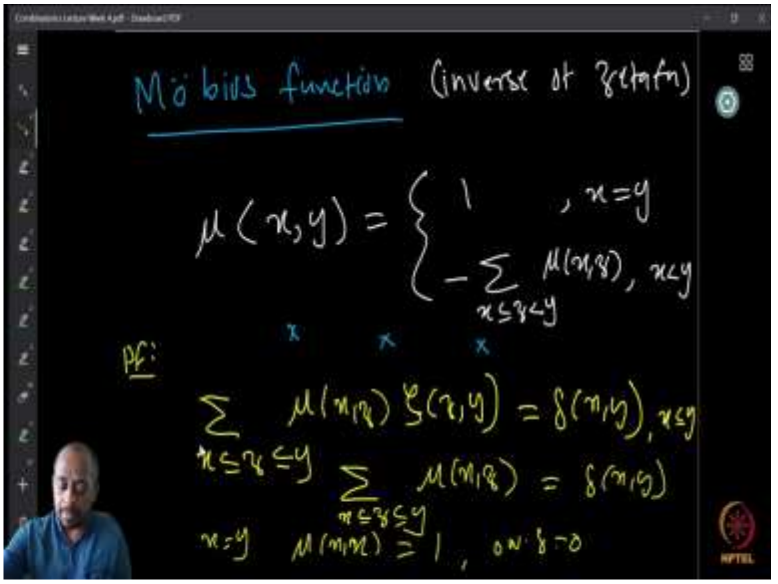
Now zeta function so has a nice property zeta function is equal to one when x is less than or equal to y and zero otherwise. So in this interval what we know is that zeta function is one whenever it is looking at the value z because z it is between x and y, it is in the interval, so therefore we know that this summation can be written as by us because zeta is one whenever we are counting and otherwise it is 0.

So, therefore this is actually equal to $\sum_{x \le z \le y} \mu(x, z) = \delta(x, y)$, this is true. Now, if x is equal to y then then what we get is that like there is only one time so $\mu(x, x)$ is equal to 1. Now otherwise because it is $\delta(x, x)$, otherwise $\delta(x, y)$ is equal to 0.

Now, if $\delta = 0$, we can move one of the terms to the left side, for example, $\mu(x, y)$ we move to the left, so therefore we will get this identity, because we just took the last term $\mu(x, y)$ from here and therefore that is because the summation is 0 we get $\mu(x, y)$ is equal to minus summation. So that is the proof that this is indeed the inverse of zeta of xy. So we have the Mobius function which is the inverse of zeta.

(Refer Slide Time: 19:36)





So let us look at some example and try to compute one or two cases of the Mobius functions. How to compute one Mobius function? So let us look at the simplest example possible like consider the chain, let us say the set of integers 1 to n ordered by the usual less than or equal to order, usual order over the natural numbers.

So this this is our poset representing, the chain is represented like this. Now, what is this, we already know already but we can just write it. So $\zeta(x, y)$ is equal to 1 if x is not equal to y and 0 otherwise and $\delta(x, y)$ is equal to 1 if x is equal to y and otherwise 0. Now by a definition of the convolution and the definition that zeta and mu are inverses of each other, $\zeta\mu(x, y) = \mu\zeta(x, y)$.

Now, $\mu(x, x)$ is equal to 1, we already we already established that it is actually 1 and $\mu(x, y)$ has this formula $-\sum_{x \leq z \leq y} \mu(x, y)$. So, let us try to use this and try to compute μ in this case. So, what is $\mu(x, x)$ that is 1, we already know. So let us look at what is the next, $\mu(x, x + 1)$ the next element.

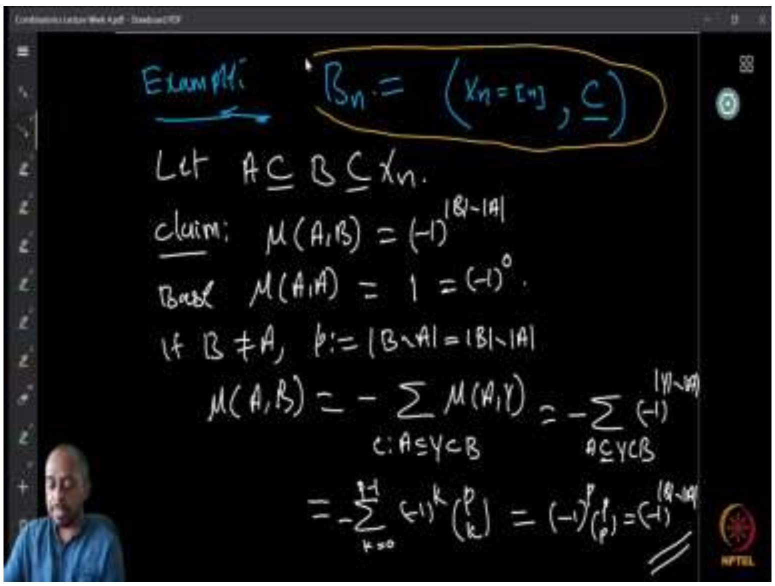
Now this by the summation $\mu(x, y) = -\sum_{x \leq z \leq y} \mu(x, z)$, here it is $-\sum \mu(x, x + 1)$. We know $\mu(x, x)$ because the last term is x plus one which is a strict inequality therefore we do not consider that. So because we have already taken it to the left side so therefore $\mu(x, x + 1) = -1$. Now what is $\mu(x, x + 2)$?

Well, first is the $\mu(x,x)$ which is 1, then the second term is $\mu(x,x+1)$ which is minus 1 and then that is it. $\mu(x,x+2)$ is 0. And if you if you continue this you can see that, because this sum is 0 this sum is zero and you will get all the subsequent entries to be also 0. So therefore we can see that

$$\mu(i,j) = \begin{cases} 1, \ i = j \\ -1, \ i+1 = j \\ 0, otherwise \end{cases}$$

So we computed the Mobius function of the chain.

(Refer Slide Time: 22:45)



One more example let us look at the poset B_n which is the poset of all the subsets of an n element set ordered by the inclusion. So consider A subset of B subset of X_n, so two sets A and B such that A subset of B this is also important. Now our claim is that $\mu(A,B) = (-1)^{\{|B|-|A|\}}$.

So we want to find out μ but we just start with this claim. Now, how can I start with a claim before finding it out? Well one way to do this is to start from the smallest and then try to work out a couple of examples, very small examples. And then try to see a pattern and then you can conjecture it, so that is that is something you should try maybe take this without looking at this expression that is given.

Try to come up with this prediction by looking at some example. So, one can do that and suppose we have done that and we have a prediction, so now how do you prove this? $\mu(A, B) = (-1)^{\{|B|-|A|\}}$? We can use induction here so for the base case we have $\mu(A, A)$ is equal to 1 by definition but that is equal to minus 1 raised to 0 cardinality of A minus cardinality of A is 0.

So therefore the base case satisfies the induction here. Now suppose B is different from A, if B is different from A then we have some elements in B\ A, because B is containing A, so look at cardinality of B\A and call that to be ‘p’. Now because of again the containment we know that this is also equal to cardinality of B minus cardinality of A.

Now, let us apply the induction hypothesis, so

$$\mu(A, B) = -\sum_{Y: A \subseteq Y \subset B} \mu(A, Y) = -\sum_{A \subseteq Y \subset B} (-1)^{\{|Y|-|A|\}}$$

$$= -\sum_{k=0}^{p-1} (-1)^k \binom{p}{k} = (-1)^p = (-1)^{\{|B|-|A|\}}$$

So that is it. So, we have showed that by induction this holds. So $\mu(A, B) = (-1)^{\{|B|-|A|\}}$ for the Bernoulli poset B_n.