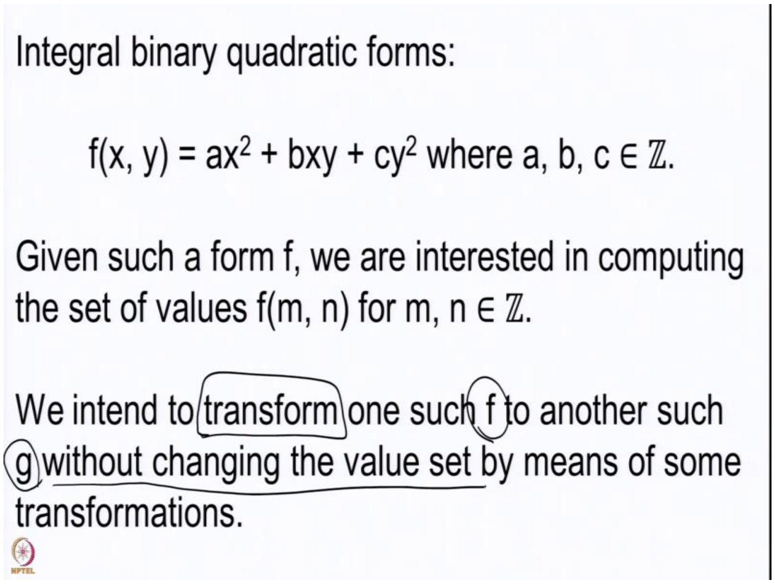
A Basic Course in Number Theory
Professor Shripad Garge
Department of Mathematics
Indian Institute of Technology, Bombay
Lecture 39
Equivalence of binary quadratic forms

Welcome back. We started a new theme in our last lecture. We are studying binary quadratic forms over integers. So, these are also called integral binary quadratic forms. So, here by form we mean a homogeneous polynomial, binary form would tell you that it is a polynomial in two variables. So, we have a homogeneous polynomial in two variables and by quadratic, we mean that it is of degree 2. So, it is a degree 2 homogeneous polynomial. Therefore it will have monomials x square x y and y square. There are two variables.

So, the variables are x and y and the degrees are two so you can have x square you can have y square and you can have x y and then you have the linear combinations of these three monomials giving you a homogeneous degree 2 polynomial in two variables. But where do the coefficients a, b, c come from? So, you have a form a x square plus b x y plus c y square where do the a, b, c come from, they come from integers? Therefore these are what are called integral binary quadratic forms.

(Refer Slide Time: 01:45)



Integral binary quadratic forms:

$$f(x, y) = ax^2 + bxy + cy^2 \text{ where } a, b, c \in \mathbb{Z}.$$

Given such a form f, we are interested in computing the set of values f(m, n) for m, n ∈ $\mathbb{Z}$.

We intend to transform one such f to another such g without changing the value set by means of some transformations.

We saw this definition in our last lecture. So, here a, b, c are integers. We can take them from a more general set of numbers. So, I told you in the last lecture that we should be able to put values for x and y from the same set where a, b, c are taken from. And therefore you should be able to have the property that the set has addition and multiplication. And since we do not want to write x y and y x separately. It should better be commutative with respect to multiplication.

So, such a thing is called a commutative ring. So, you may have the set of complex numbers. You may have the set of real numbers. You may have the set of rational numbers and there are many rings it in between these rings. So, the a, b, c can take values coming from these rings. There are some more general set of rings but we are going to restrict ourselves to integral such forms.

So, in what follows throughout this theme, I may sometime write form this to be a quadratic form or I may just say a form but we always understand that we are going to consider only the integral binary quadratic forms. Now the main question in this theory is to find the value set of such a form. So, whenever we are given such a form f x y we want to know what is the set of values f m, n where m and n are integers.

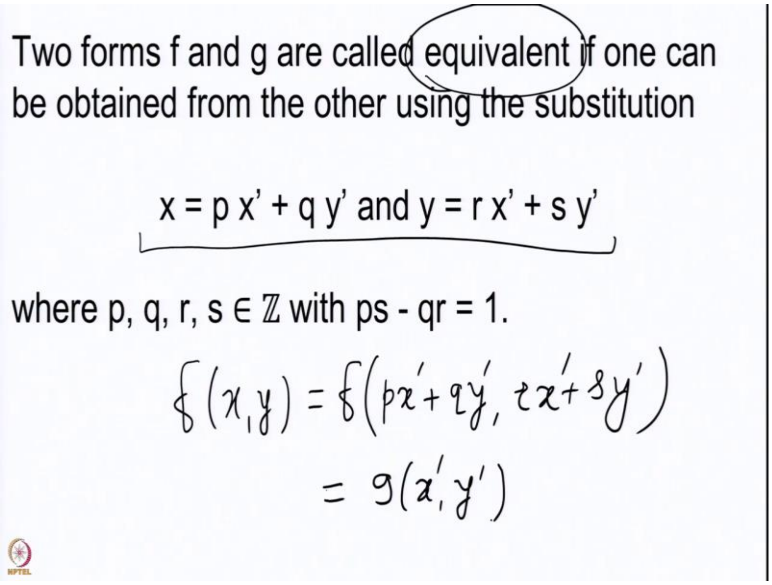
So, we are going to take all possible values of integers for m and n we put these values in our form. So, we are going to get a m square plus b m n plus c n square. We are looking at the set of these numbers and we want to determine we want to say which are how do the values look like. And once we have some information about how these values look like, we would like to know whether all the numbers of that form can be represented by our form f.

We will soon see an example. So, this problem will be clear very soon. But what we want to do while studying this also is to have some transformations. So, we would like to transform one form f to another such form g. So, we start with an integral binary quadratic form call it f and we would like to go to another integral binary quadratic form g through a way. So, we will describe this way and we would also like to describe this way with the property that the value sets do not change.

So, this is what we mean by without changing the value set. The problem while for, in determining the value set is as follows. You may have a form but the form may have very big integers a, b, c as coefficients and therefore whenever you want to compute the values say directly or you may want to obtain some, you may want to obtain some behavior of the value set then it may be very difficult to obtain that.

So, we would like to transform any such given form to a somewhat simpler form by some transformations and we would like to have these transformations to be invertible. If you are going from f to g, you should be able to go back from g to f by a similar another transformation. Perhaps the same transformation or perhaps another transformation. So, this is a relation that we would like to define between the integral binary quadratic forms. And we have already defined this relation towards the end of our last lecture.

(Refer Slide Time: 06:09)

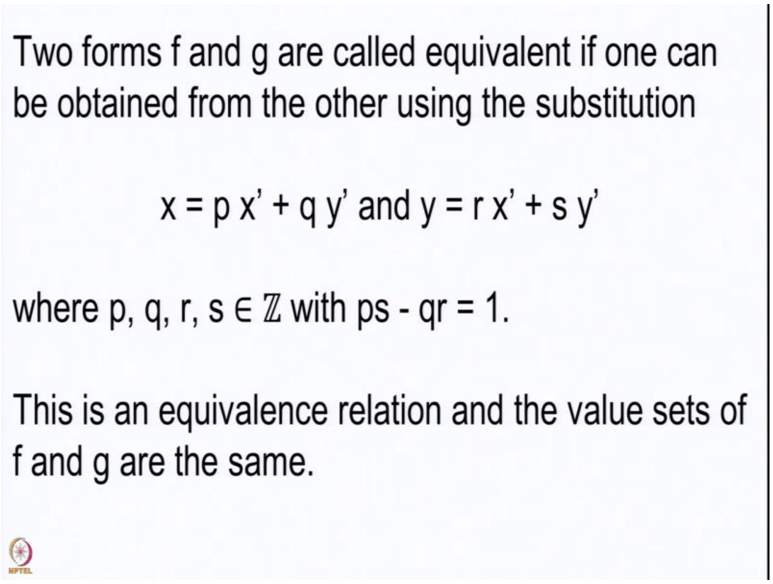


Two forms f and g are called equivalent if one can be obtained from the other using the substitution

$$\underbrace{x = p x' + q y' \text{ and } y = r x' + s y'}$$

where p, q, r, s ∈ $\mathbb{Z}$ with ps - qr = 1.

$$f(x,y) = f(px'+qy', rx'+sy')$$
$$= g(x', y')$$



Two forms f and g are called equivalent if one can be obtained from the other using the substitution

$$x = p\,x' + q\,y' \text{ and } y = r\,x' + s\,y'$$

where $p, q, r, s \in \mathbb{Z}$ with $ps - qr = 1$.

This is an equivalence relation and the value sets of f and g are the same.

Let me quickly recall that for you. So, we say that two forms, so of course we mean the binary integral quadratic forms. They are called equivalent. If you can obtain one of them from the other using this substitution, so you have your form f x y and here you substitute the values. So, instead of x you will have p x prime plus q y prime and instead of y you will write r x prime plus s y prime.

So, this will be another form in new variables x prime and y prime and this new form a different integral binary form will be said to be equivalent to our earlier binary form f. Now I made this comment in the last lecture and I make it again that this is an equivalence relation. This is why we call we use the word equivalent normally mathematicians are very careful while using words if some relation is not evidently equivalence and equivalence relation, then we give some another word for that.

We say that this is related to g or we would use some other another word for these things, it would be difficult to give an example now. But here we are saying, so even from the definition you may notice that we have not explicitly said whether we are obtaining f from g or whether we are obtaining g from f. We are simply saying that f and g are equivalent if one of them can be obtained from the other using the substitution.

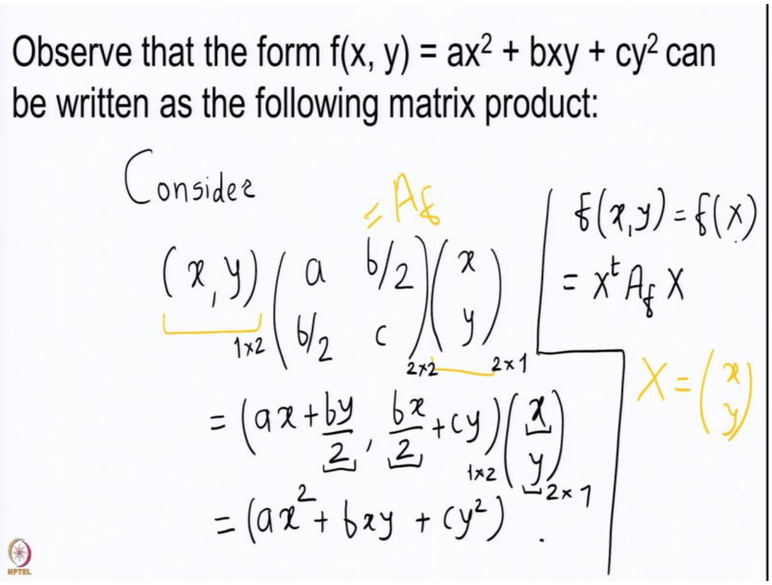
So, if you can obtain f from g, we call them equivalent or if you can obtain g from f then also we call them equivalent. So, by very construction this definition is symmetric whenever f is equivalent to g, g is also equivalent to f. But we can also restrict our definition by saying that we are obtaining g from f as we have done in the example that you put the values of x and y in the formula for f then you get another homogeneous polynomial by simplification in the new variables x prime and y prime.

And when you say x prime y prime just as variables as, so you may rename them as x and y once you have completed the coefficients and therefore you get a new possibly different integral binary quadratic form, then we call that g. So, you may obtain g from f or you may obtain f from g by definition both are allowed, but let us say that we are going to only obtain g from f. And then we will see that this is an equivalence relation.

Further it is also important that the value sets remain the same. So, this is the very fundamental definition in this theory of integral binary quadratic forms. The definition of equivalence. So, we have some transformation. We have a substitution which is actually a transformation. If you know some group theory you will see that there is an action of the group SL to Z on the set of binary quadratic forms over integers.

I will let you think about whether it is a left action or right action, but this is what we have. And so this I may also sometime use the word transformation because here we have a the SL to Z is coming because we have p s minus q r equal to 1. So, that is why this is the formula for the determinant of a 2 by 2 matrix and that is precisely the matrix SL to Z that is acting on the set of integral binary quadratic forms.

(Refer Slide Time: 10: 39)



Observe that the form $f(x, y) = ax^2 + bxy + cy^2$ can be written as the following matrix product:

Considee

$$\underbrace{(x, y)}_{1\times 2}\underbrace{\overset{= A_f}{\begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix}}}_{2\times 2}\underbrace{\begin{pmatrix} x \\ y \end{pmatrix}}_{2\times 1}$$

$$= \underbrace{\left(ax + \frac{by}{2}, \frac{bx}{2} + cy\right)}_{1\times 2}\underbrace{\begin{pmatrix} x \\ y \end{pmatrix}}_{2\times 1}$$

$$= (ax^2 + bxy + cy^2).$$

$$f(x, y) = f(X)$$

$$= X^t A_f X$$

$$X = \begin{pmatrix} x \\ y \end{pmatrix}$$

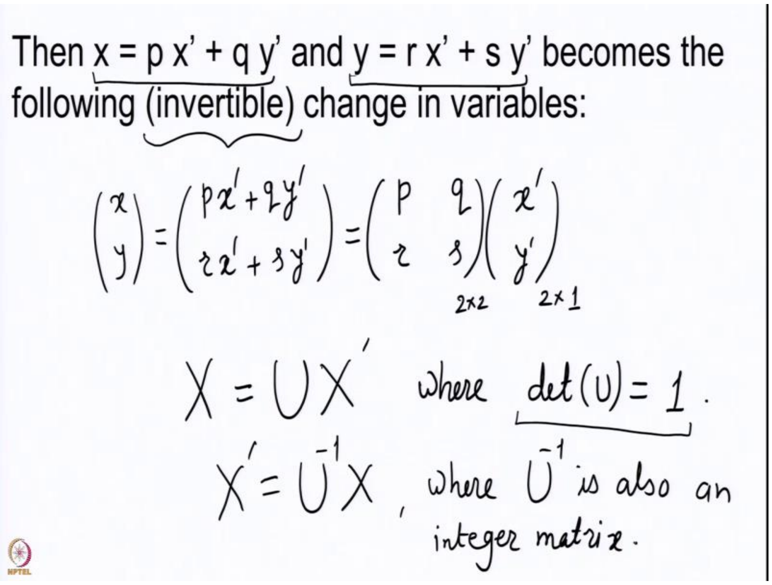
So let us, let me give you some gist of this matrix notation, which makes our things very simple. So, to begin with let us try to understand how the form f itself can be written as the matrix product. So, we consider x, y we are taking it to be a row vector. And here I have this 2 by 2 matrix. Note that this is a symmetric matrix and then I multiply to that by a column vector. So, the vector which we have here the row vector and the column vectors, these are transposes of each other.

That is one thing that we notice. Let us try to compute this value. So, we first obtain the left side multiplication and that would be a x plus b y by 2 comma b x by 2 plus c y and now we multiply to it by the column vector x y and we see that this is earlier. So, this was a 1 by 2 matrix. This is a 2 by 2 matrix. And this is a 2 by 1 matrix, in the end we are going to get a 1 by 1 matrix. So, this is 1 by 2 and this is 2 by 1, we will get a 1 by 1 matrix, which is a x square.

Plus this will give me b x y by 2 and this also gives me b x y by 2. So, I just have b x y. And finally we have c y square. So, we have written our integral binary form in the following way, equal to x transpose a f x. So, what is our capital x, capital x is the column vector x, y and af is this 2 by 2 matrix. So, this 2 by 2 matrix is obtained by putting the coefficients for x square and the coefficient for y square on the diagonal in that order and the coefficient of x y is divided by 2.

And then we put it in the off diagonal entries. So, this is the way we obtain the value of the integral binary quadratic form in terms of a matrix product. So, f x y or you may also call it f of capital X f of capital X is X transpose a f x. This is how we have obtained the binary form and its value and now let us see how this helps us in doing this computations.

(Refer Slide Time: 13:55)



Then $\underbrace{x = p x' + q y'}$ and $\underbrace{y = r x' + s y'}$ becomes the following $\underbrace{\text{(invertible)}}$ change in variables:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} px' + qy' \\ rx' + sy' \end{pmatrix} = \underset{2\times 2}{\begin{pmatrix} p & q \\ r & s \end{pmatrix}} \underset{2\times 1}{\begin{pmatrix} x' \\ y' \end{pmatrix}}$$

$X = UX'$ where $\underline{\det(U) = 1}$.

$X' = U^{-1}X$, where U^{-1} is also an integer matrix.

So, once again this transformation X going to this and why going to this can be written as x y which can be written as p x prime plus q y prime and here we have r x prime plus s y prime. So, you would have recognized here that this is nothing but the multiplication of this 2 by 2 matrix with this 2 by 1 matrix. So, we have p x prime plus q y prime and r x prime plus s y prime. So, we have capital X to be capital U into capital X prime.

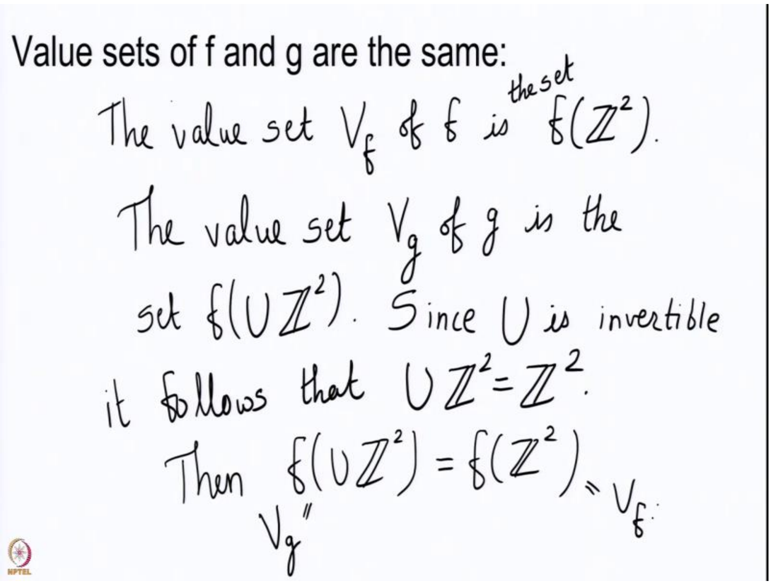
Where we have that the determinant of U is 1. So, the matrix of x y has now changed to matrix x prime because of this transformation capital U. So, we initially applied our form f to X ,now we are changing we are writing X as U into X prime. So, we are going to get a new form and let us see how that new form looks like before we go further. Let me also tell you why this change of matrix.

Before we go further let me also tell you why this change of variables is invertible. This is precisely because the determinant of the matrix U is 1. So, we have a 2 by 2 matrix with integer entries if its determinant is 1 or minus 1, then its inverse is also going to have integer values, thus entries in the inverse are also integral. So, whenever we have X equal to U X prime we have X prime to be U inverse X and here U inverse is also an integer matrix.

So, the entries of U prime might be alpha, beta, gamma, delta and you would have the same property that alpha delta minus beta gamma, the determinant of U inverse is equal to 1. So, we have the same way when you go from X to X Prime by applying this Matrix U of determinant 1. You can go come back from X Prime to X by applying U inverse, which also will then have determinant 1 and because the determinant of U is 1, U inverse will also have integral entries.

So, this is the way we quickly see that when you go from X to X Prime the change of variables is invertible. In fact, the relation among the quadratic forms is simply given by the change of variables and now it is easy to see for you that when you have two matrices of determinant 1 their product is also going to have determinant 1. So, when you go from X to X prime and X prime to X double prime, you are actually going from X to X double prime by means of a matrix whose entries are integers and whose determinant remains to be 1.

(Refer Slide Time: 17:44)



Value sets of f and g are the same:

The value set V_f of f is the set $f(\mathbb{Z}^2)$.

The value set V_g of g is the set $f(U\mathbb{Z}^2)$. Since U is invertible it follows that $U\mathbb{Z}^2 = \mathbb{Z}^2$.

Then $\underset{V_g''}{f(U\mathbb{Z}^2)} = f(\mathbb{Z}^2) = V_f$.

This quickly tells you that it has to be an equivalence relation. But let us see further. First of all that the value sets are the same. This is observed because we have our change of variables to be invertible. So, the value sets the values Vf of f is we apply f to Z cross Z. So, you are going to take m comma n where both m and m come from integers and you apply this binary integral quadratic form to the pair m comma n and then whatever the values of this form that you get the whole set this is the set.

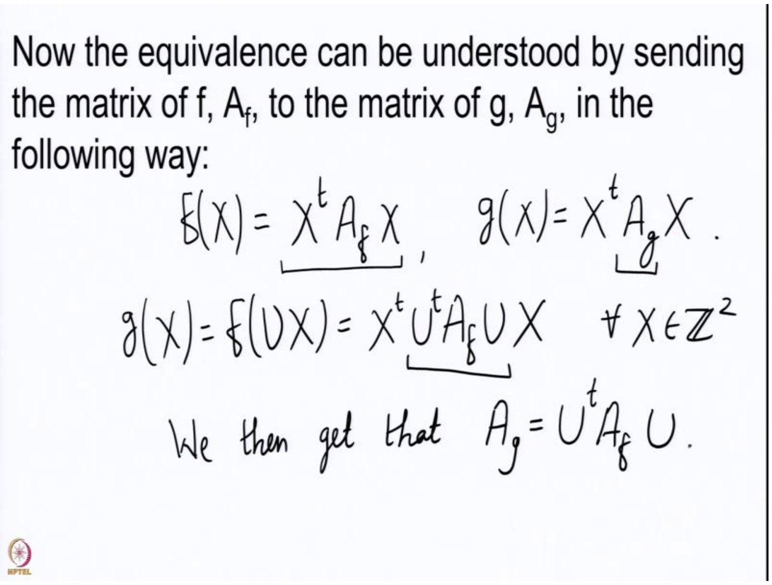
Whenever you have a function from a set x to a set y and you take a subset of x say w then f of w is a subset of y and it is the set of all values f of small w where small w belongs to capital W. So, here you are taking f of all m comma n as m and n vary over integers. So, this is the value set and the values set Vg of g is the set f of U Z 2. This is because we will apply the transformation U to the vector capital X and then we obtain the value of the quadratic form f on this, which is really the form g applied to the x and y.

So, this is the value set but since you is invertible it follows that U Z 2 is Z 2. So, let us think about this, we have applied U to the column Vector say m comma n coming from Z 2 and therefore after applying U to m n v, get some one more column Vector. So, we get an element in Z2 again. So, the multiplication by U on the left gives you a map from Z 2 to Z 2 and so U Z 2 is sitting inside Z 2, but U is also invertible.

So, by applying U inverse to some element in the image, we have found a vector such that when you apply U to this Vector you get the desired Vector. Therefore this U Z 2 is equal to Z 2 the only thing we need here is that the map U is on 2. Whenever the map U is on 2, U of Z 2 is going to be equal to Z 2. So, once you have that U Z 2 is Z 2, then we get that f of U Z 2 which is really the set Vg is same as f of Z 2.

So, this is Vg and this is Vf. So, both the value sets remain the same whenever we apply these transformations both the value sets are the same. So, we are not changing the values by applying these Transformations, but let us also understand how these transformation look like when we, you go from the matrix Af which was the matrix for the quadratic form f to the matrix Ag for the quadratic form g.

(Refer Slide Time: 21:50)



Now the equivalence can be understood by sending the matrix of f, A_f, to the matrix of g, A_g, in the following way:

$$f(X) = \underbrace{X^t A_f X}_{}, \quad g(X) = X^t \underbrace{A_g}_{} X .$$

$$g(X) = f(UX) = X^t \underbrace{U^t A_f U}_{} X \quad \forall X \in \mathbb{Z}^2$$

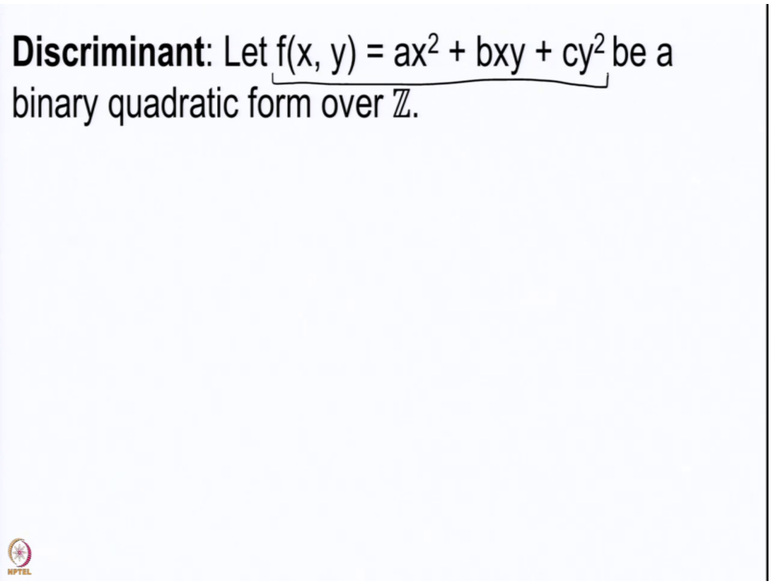
We then get that $A_g = U^t A_f U$.

Let us also try to understand that. So, we have understood that f of x y. So, f of capital X is nothing but X transpose A f X. And similarly when you apply g to X you get X transpose Ag X. These are the definitions of the matrices Af and Ag and now we are going to get these values that g capital X is nothing but f of UX and so this is obtained by this formula. So, we are going to apply the formula to UX. The transpose will give me X transpose U transpose Af UX. So, we have this quantity here and we have this quantity here since these are same for every X in Z 2, we then get that Ag is U transpose Af U.

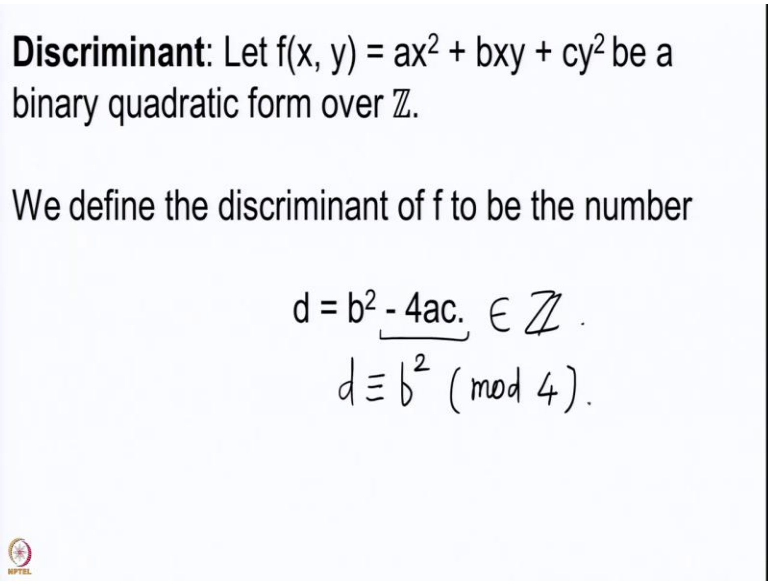
So, we have a matrix of determinant 1, of course with integer values such that the matrix Ag is related to the matrix Af by the formula that Ag is U transpose Af U. And now it is easy to say once again that this is an equivalence relation in a different way. Because if you, if you just put f on both sides you want to relate Af to f you take U to be the identity matrix. If you have Ag in terms of Af and you want to write Af in terms of Ag you will take the inverse of U.

So, you have the reflexibility, you have the symmetry and now you want to get the transitivity, which is also very easy because product of two matrices with integer values of determinant 1 is again, an integral matrix whose determinant is equal to 1. So, what we have seen so far is that the transformations applied to the binary integral quadratic forms keep the same values and further this is an equivalence relation. Let me quickly define the discriminant of a binary form. So, I also after having talked about Transformations, I talked about an invariant. So, here is one invariant. we start with a binary quadratic form over Z.

(Refer Slide Time: 24:57)



Discriminant: Let $\underbrace{f(x, y) = ax^2 + bxy + cy^2}$ be a binary quadratic form over $\mathbb{Z}$.

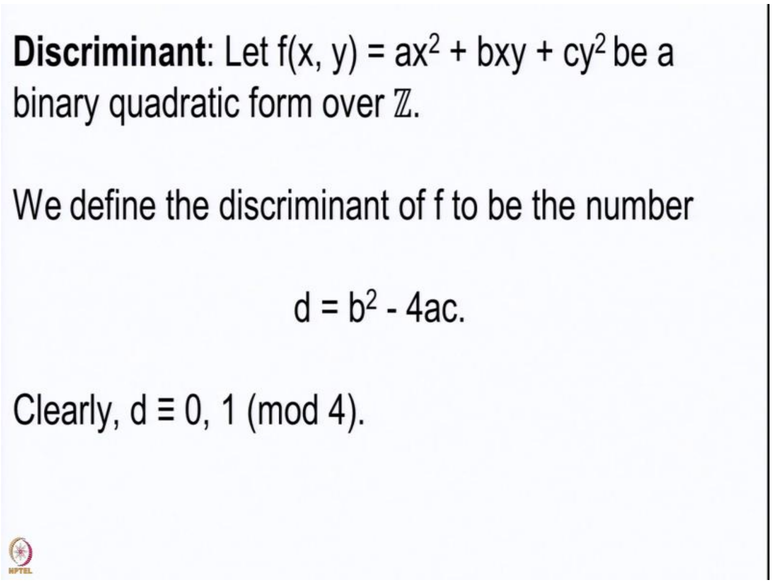


Discriminant: Let $f(x, y) = ax^2 + bxy + cy^2$ be a binary quadratic form over $\mathbb{Z}$.

We define the discriminant of f to be the number

$$d = b^2 \underbrace{- 4ac.} \in \mathbb{Z} .$$

$$d \equiv b^2 \pmod{4} .$$



Discriminant: Let $f(x, y) = ax^2 + bxy + cy^2$ be a binary quadratic form over $\mathbb{Z}$.

We define the discriminant of f to be the number

$$d = b^2 - 4ac.$$

Clearly, $d \equiv 0, 1 \pmod 4$.

We start with an integral binary form a x square plus b x y plus c y square and we define its discriminant to be this quantity, b square minus 4 a c. So, this is an integer. Moreover it has the property that module 4, this is a square. So, d is congruent to b square mod 4. Now, we have seen how squares look like modulo various integers and in particular module 4 this computation is very easy. We know that all squares modulo 4 are 0 or 1 and therefore it follows that the discriminant of a binary integral quadratic form is always 0 or 1 mod 4.

Now there are some natural questions most importantly, what is the use of the discriminant? First of all, we should also see whether it is invariant under the change of variables that we have introduced. So, we will see that we will also see a small answer to a small question that if you have an integer d, which is 0 or 1 mod 4, is it discriminant of some form. So, we are going to see answers to these questions and many other related questions in the coming lectures. So, stay tuned. Thank you.