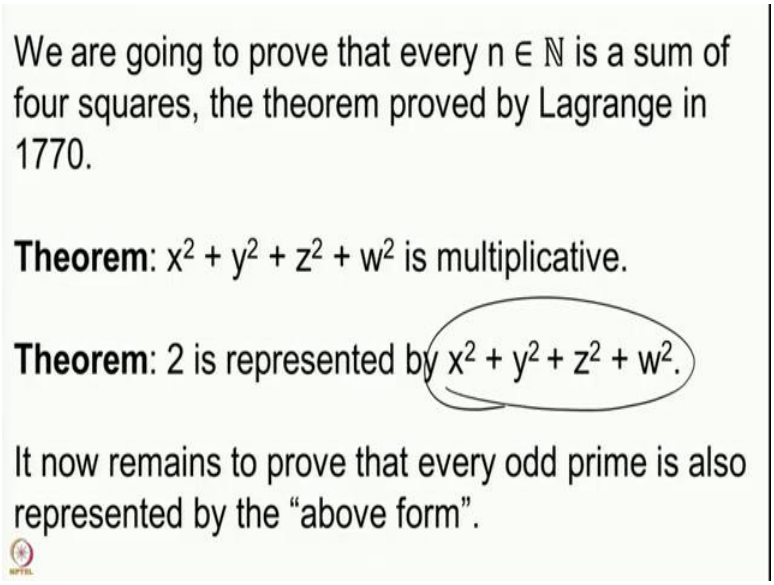
A Basic Course In Number Theory
Professor Shripad Garge
Department of Mathematics
Indian Institute of Technology, Bombay
Lecture - 47
Sums of squares - III

Welcome back, we are now going to prove Lagrange's theorem, which says that every positive integer can be written as a sum of at most 4 squares. This is the result that we want to prove. We have made some progress in the last lecture already, we proved some statements towards this proof. So, let us see those statements.

(Refer Slide Time: 0:43)



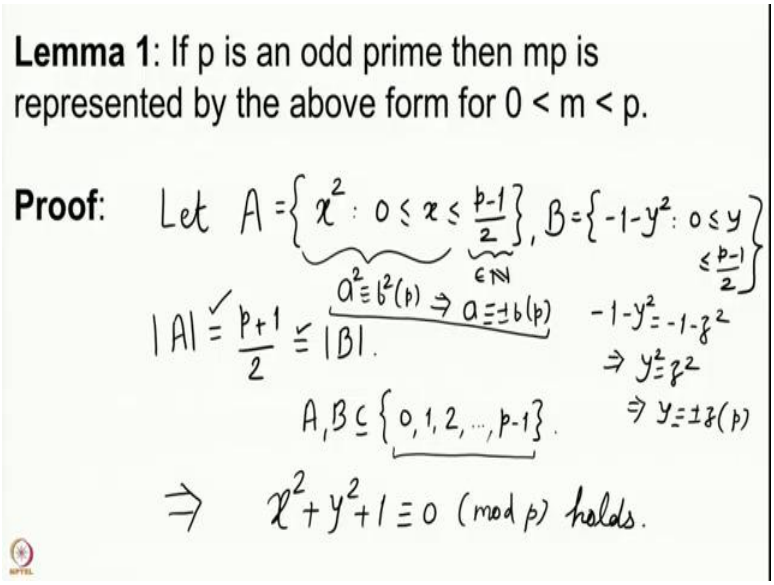
The first one that we have proved is that this form is multiplicative. That means, if we take 2 numbers, which are sums of 4 squares, so suppose we have a square plus b square plus c square plus d square and we multiply to this by x square plus y square plus z square plus w square, then this is again a sum of 4 squares, alpha square plus beta square plus gamma square plus delta square.

I remarked in the last lecture that this is of course true for sums of 2 squares which we have proved and it is also true for sums of 4 squares, but this is not true for sums of 3 squares. This is related to the remarkable discovery of quaternions by the British mathematician Hamilton. But I will let you search on internet for that, and we will proceed with this. So, after proving that this form is multiplicative, we only need to prove that every prime can be written in this way, that every prime is represented by this form, that is the only thing we need to prove.

So, we start with the first prime p equal to 2. And we also saw in the last lecture itself, that 2 is represented by this form. You can write 2 as 1 square plus 1 square plus 0 square plus 0 square and now the only thing that remains to prove is that every odd prime is also represented by the above form.

So, now, we will not keep referring to this form again and again. This is the form that we are looking at, we are not considering any other quadratic form until we complete Lagrange's theorem's proof. So, we will just keep calling this form as the above form or simply the form in the next statements. So, let us see this basic lemma.

(Refer Slide Time: 3:00)



Suppose, p is an odd prime, then m p is represented by the above form for some 0 less than m less than p. What we want to prove is that p is represented by this form. So, we would eventually prove that the m is equal to 1. We have that there is m equal to 1 such that mp is represented by the above form.

But we will first obtain this slightly weaker result that a multiple of p, moreover, that multiple is by a number which is less than p. So, we have a multiple of p which is less than p square that is represented by the above form, which is the sum of 4 squares. So, let us see the proof of this statement.

We consider 2 sets, we consider 2 sets here, our first set is called the set A, which has numbers x square where x varies from 0 to p minus 1 by 2. Remember now that p is an odd prime, so this is a natural number it is in N. So, we have total, the cardinality of A is 1, 2, 3 up to p minus 1 by 2 plus 1 because we have to add 0 to that. So, we have exactly p plus 1 by

2 numbers. And similarly here we have minus 1 minus y square again y goes from the same 0 to p minus 1 by 2.

So, we have that both the sets A and B contains exactly p plus 1 by 2 numbers. Observe here one basic thing that we have used is that if you take any x from 0 to p minus 1 by 2, and any other x prime, then their squares are never equal. In fact, if you have a square congruent to b square modulo p, then a has to be plus or minus b modulo p. This follows from Euclid's lemma, we have proved the Euclid's lemma, which says that if p divides product m n of 2 integers, then p has to divide m or n.

So, if you have a square congruent to b square modulo p, then p will divide a square minus b square which is a plus b into a minus b. So, p will have to divide a plus b or p has to divide a minus b and that gives you that a has to be congruent to b or to minus b. So, since we have taken the numbers here, which are of the form that for any number x, the negative of x is not counted modulo p because negative of x would go from p plus 1 by 2 to p.

Except for 0, we have not taken any other number, which is its own negative and of course, 0 square is not going to be equal to non 0 square modulo p. So, that tells you that the cardinality of the set a is p plus 1 by 2 and similarly, the cardinality of the b because if you have, so what we have used here is that a square congruent to b square mod p implies a equal to plus or minus b mod p. So, this simple statement enables us to prove this equality.

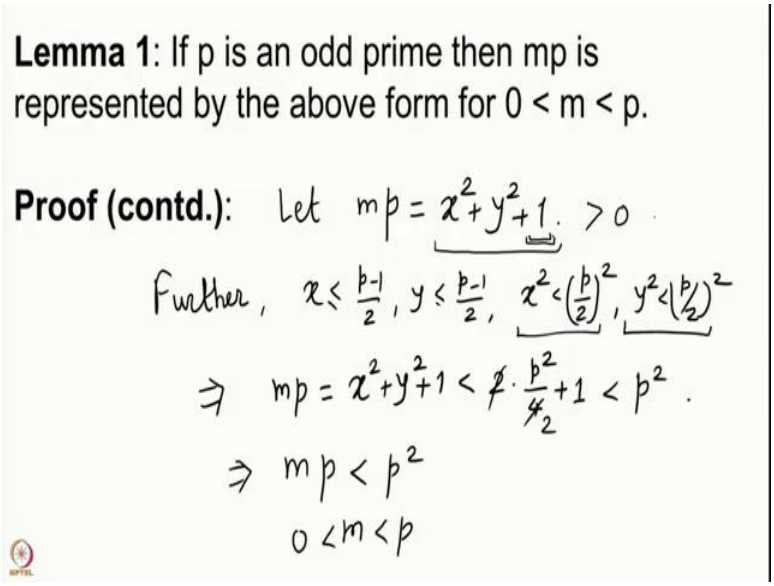
And here if you have minus 1 minus if you have minus 1 minus y square equal to minus 1 minus z square, this would imply that y square equal to z square and then again y equal to plus or minus z mod p. So, this would imply equality and so you have that this holds modulo p, which is again a contradiction. So, this enables us to prove this equality. So, now, of course, both a and b are contained in when you look at their modulo p, they are contained in 0, 1, 2, 3 and so on up to p minus 1 modulo p, and both have p plus 1 by 2 elements.

So, there has to be a common element in them. This is because, if they both had distinct elements, then their union would have p plus 1 elements. Because if you have p plus 1 by 2 elements in set A, p plus 1 by 2 elements in set B, and if all these numbers are distinct, which would mean that none of the elements in A is in B and none of the elements in B is in A, then the Union would have cardinality p plus 1, which is not, which is going to give us contradiction, because here we have exactly p elements.

So, there has to be a common element in the set A and B and this implies that there is some x square here, which is equal to some minus 1 minus y square. So, x square plus y square plus 1 congruent to 0 mod p holds. So, we have proved that x square plus y square plus 1, if you just treat x and y as variables, and you consider this equation x square plus y square plus 1 congruent to 0 mod p then this equation has a solution over z by p z over the set of all congruence classes modulo p.

We have produced one solution for this. So, x square plus y square plus 1 is a multiple of p, because x and y are integers, which are of certain, which are in certain range, and therefore this particular sum of 3 squares is now a multiple of p.

(Refer Slide Time: 10:08)



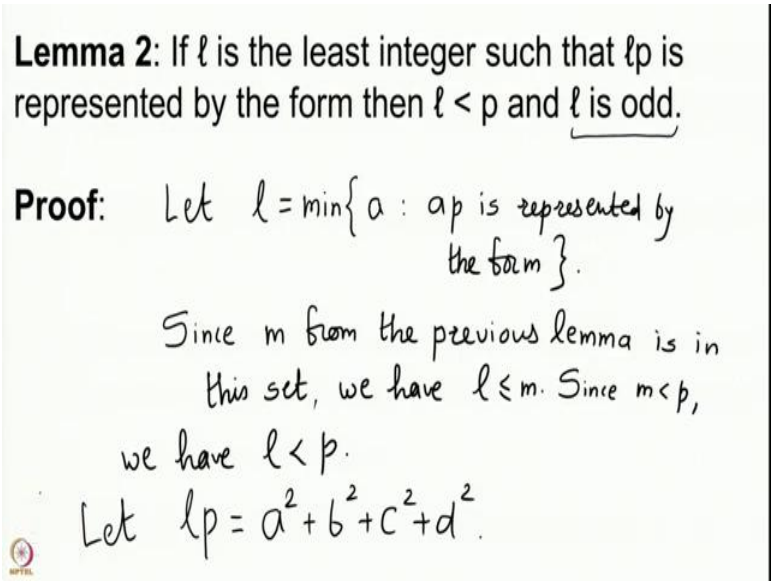
Let us call this sum as mp, further, x is less than p minus 1 by 2, y is less than or equal to p minus 1 by 2. So, they are squares. In any case, x square is less than p by 2 square y square is less than p by 2 square, so m p, which is x square plus y square plus 1 is less than 2 times p square by 4 plus 1, because we have that x square is less than p square by 4, y square is less than p square by 4. So, the sum will be less than 2 times p square by 4 and you have 1, and if you cancel this 2, you just get p square by 2, which is easily seen to be less than p square.

Remember p is an odd prime and therefore, 1 is always less than p square by 2. So, that would tell you that m p is less than p square and hence p has to be less than m. Since you have mp to be sum of 2 squares x square plus y square plus 1, therefore mp has to be a number which is bigger than 0, m p cannot be equal to 0. So, you have that. So, we get here m less than p, but we will also get that m is bigger than 0, because this quantity here is

certainly bigger than 0. You have 1 here, which tells you that m p is bigger than 0 and so m cannot be 0.

So, we have proved this lemma, we wanted to prove actually that p is represented by this form, but what we have proved is a weaker form of that statement, which is that a multiple of p strictly less than p square is represented by our form x square plus y square plus z square plus w square. So, this lemma 1 is proved, we now go to one more lemma and then we proceed towards proving Lagrange's theorem.

(Refer Slide Time: 13:02)



So, this is second lemma, which says that if l is the least integer such that lp is represented by the form, then l should satisfy 2 conditions, first of all l has to be less than p and secondly, that l is odd. So, what we are looking at is the following thing. So, let l be the least, what we also call as min of a, where a p is represented by the form. So, l is taken to be the least multiple, least integer such that the multiple of p by this integer is represented by this form.

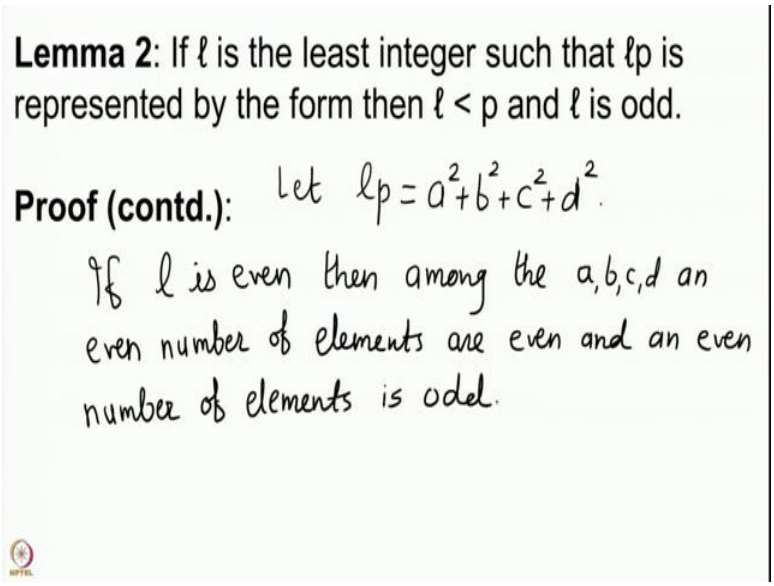
Now, in the previous lemma we have already proved that there is a multiple of p, we call that m times p which is represented by this form this is what we have proved in lemma 1 and we also noticed there that the m that we had in the lemma 1 is less than p. So, since, m from the previous lemma is in this set, we have an l to be less than or equal to m, because you are taking minimum of all integers whose multiples by p are represented by this form.

So, it may happen that your l might be equal to m, but it is also possible that l might be less than m. So, in any case, you have that l is less than or equal to m and we have already proved that m is less than p. So, we have l is less than p. So, we have a multiple of p by a number l

which is less than p such that lp is sum of 4 squares and l is least such integer. We now want to prove that l is odd.

So, let us see what we have. So, suppose we write lp as a square plus b square plus c square plus d square. Suppose, this is a representation for lp by the form which we are looking at. Now, the integers a, b, c, d are some integers they can be even or they can be odd. If your l is an even integer.

(Refer Slide Time: 16:51)

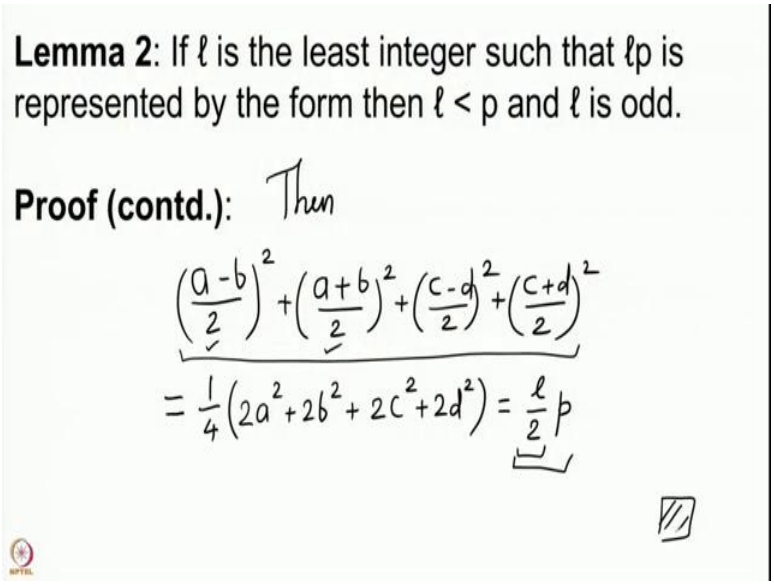


So, if, I will use the next slide for this, let lp be equal to a square plus b square plus c square plus d square if l is even then among the a, b, c, d an even number of elements are even and an even number of elements is odd. So, we have these 4 elements a b c d. Suppose, that a is odd and b c d are even if that is the case, then b square plus c square plus d square are all even, their sum is even. And you have a to be odd, so a square plus this even thing will actually give you an odd number, which will mean that lp is odd, then l cannot be even.

So, similarly, whenever you have an odd number of odd elements among the a, b, c, you have that their sum, their squares will sum to an odd number and the remaining numbers are anyway even so the total sum is odd, which will mean that l has to be odd. So, if you happen to have l to be even, then two of them are odd, four of them are odd or none of them are odd. So, the number of odd elements among a b c d is either 0 or 2 or 4.

And we can rearrange them in such a way that a b are of the same parity and c d are of the same parity. That means a and b are both odd or both even and similarly, c and d are both odd and both even. So, if you have this, then what do we get?

(Refer Slide Time: 19:38)



Then a minus b by 2 square plus a plus b by 2 square plus c minus d by 2 square plus c plus d by 2 square. Remember that a and b are of the same parity, either both are odd or both are even, and therefore their difference and there are some are both even numbers. So, these 4 numbers are all integers. So, we sum these up.

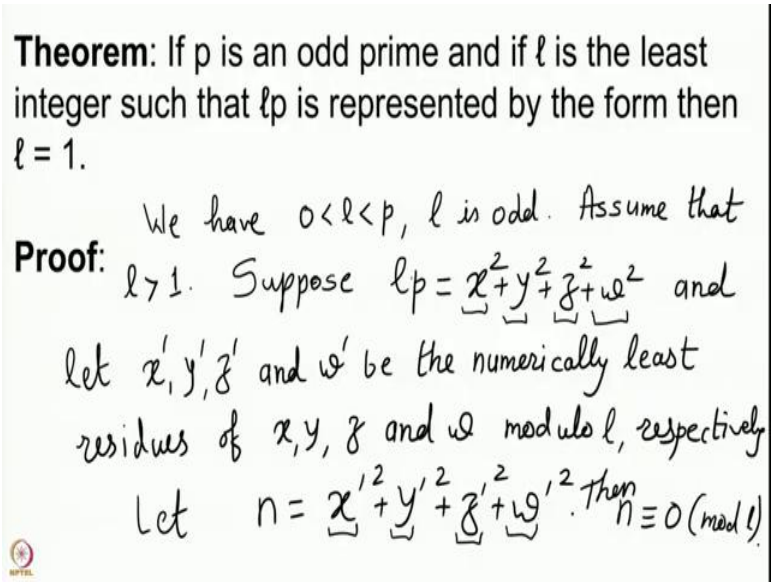
Here we get let us take the 1 by 2 square common that will come on the outside and we have a minus b whole square plus a plus b whole square, this will give us a square, this gives us a square so you have 2 a square plus 2 b square and the minus 2 a b that you get here will get cancelled with the plus 2 a b that you get here. Similarly, here we have 2 c square plus 2 d square, ultimately we get that this is l by 2 into p.

Because the 2's that you have here will get cancelled with the 2 that appears in the denominator. So, you have a square plus b square plus c square plus d square divided by 2. And so you have a multiple of p, which is smaller than l, because l was even and l was positive. So, we have now a multiple of p, which is less than l and this multiple into p is also sum of 4 squares, we have actually these 4 integers, whose squares sum to this particular number.

So, we therefore get that l cannot be least, this is a contradiction to our assumption, that l is an even integer. So, this contradiction says that l is even, this assumption cannot be true, and therefore l has to be odd. So, we are taking the least positive integer where lp is represented by this form, then l is an odd integer strictly less than p. With these 2 lemmas, now let us go and prove Lagrange's theorem, which is that essentially Lagrange's theorem because we are

going to prove that every odd prime is represented by the form, then Lagrange's theorem will follow very soon after proving this result.

(Refer Slide Time: 22:35)



So, this is the integer that we want, this is the theorem that we want to prove that if p is an odd prime, and if l is the least positive integer with the property that lp is represented by the form, then l has to be 1. So, we will start by assuming that l is not 1 and get a contradiction. We have already noted that l is here, l is odd, these are 2 very important results that we have obtained and they are going to be used. Now, consider, so assume that l is bigger than 1.

Now, I must recall for you one concept which we have introduced some lectures back, this is the concept of a numerically least residue. So, whenever we are given any integer n and we are going modulo another integer say l, then we will look at residue of n modulo l and typically this residue is chosen from the elements 0, 1, 2, 3 up to l minus 1, this is where we look for the residue. But I told you that from the point of view of multiplication, addition, etc, it is more convenient to take this residue from a set which consists of both positive and negative elements.

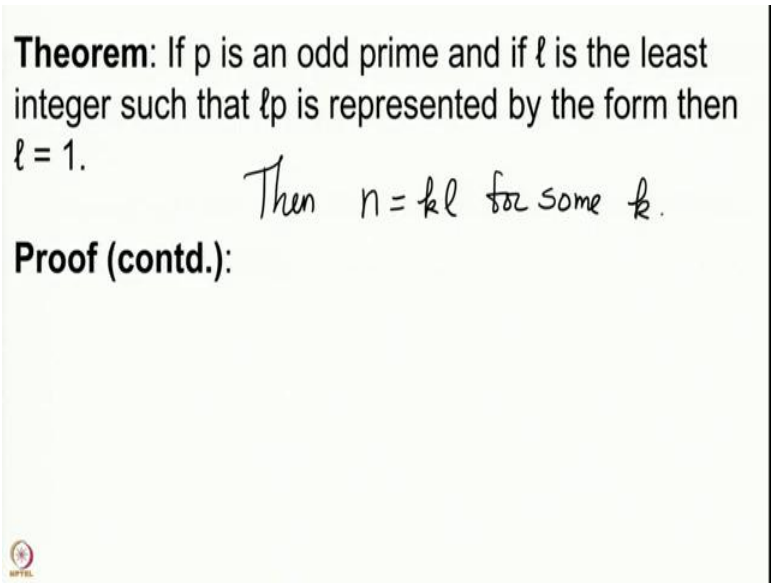
So, this is to be taken from minus l by 2 to l by 2 and we do not include minus l by 2 because modular l both these numbers are same. Here of course, our l is odd, so we have no such problem of considering minus l by 2 and l by 2 those are not integers. So, we are looking at numerically least residues, that means we are looking at the residue which is in the set minus l by 2 to l by 2.

So, we will start with our representation of lp, this is our representation of lp by the form and let x prime, y prime, z prime and w prime be the numerically least residues of x, y, z and w modulo l respectively. What do I mean by respectively, respectively would mean that x prime is the numerically least residue of x, y prime is the numerically least residue of y, z prime is the numerically least residue of z, and w prime is the numerically least residue of w, this is what we mean. x prime, y prime, z prime, w prime are some integers after all, let n be the sum of their squares.

Now, x is congruent to x prime, y is congruent to y prime, z is congruent to z prime and w is congruent to w prime modulo l. So, as far as you are looking at the congruence modulo l, x prime is x, y prime is y, z prime is z and w prime is w and so their squares will also be congruent, and then the sums of the squares will also be congruent. So, modulo n modulo l, n is congruent to the earlier sum, which is l times p, and therefore, n is congruent to 0 mod l.

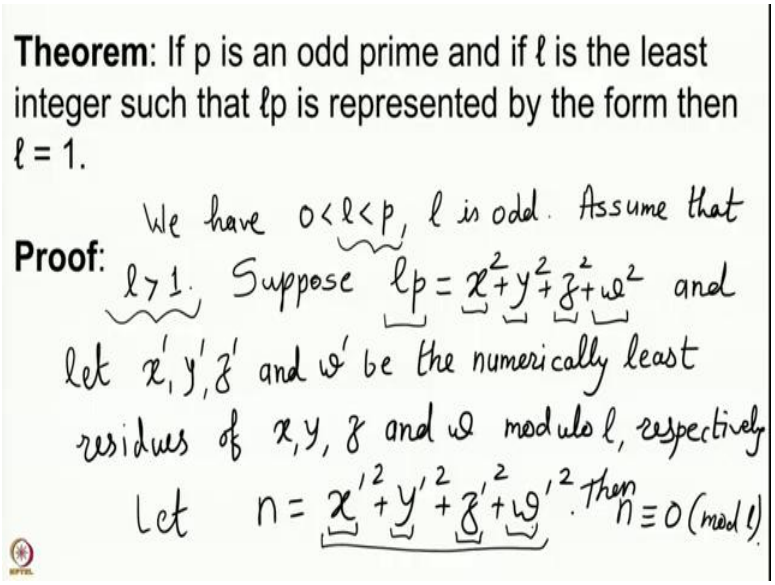
Because we have that each of these is congruent to each of these. So, their squares are congruent, and therefore their sums are also congruent. So, n is congruent to lp modulo l, which means that n is 0 mod l. So, n is a multiple of l.

(Refer Slide Time: 27:34)



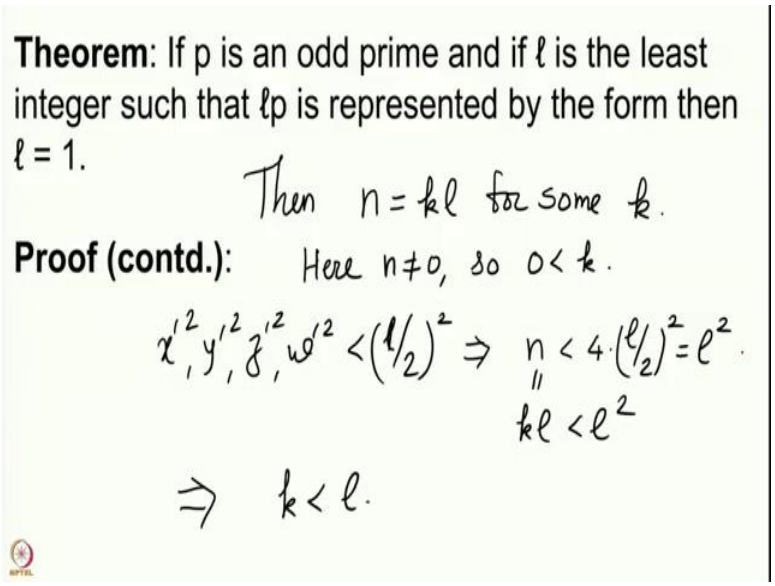
Then n is k l for some k. Can you have that n is 0? Is it possible that n is 0? Let us go back and see what it would mean that n is equal to 0.

(Refer Slide Time: 27:59)



First of all, note that n is a sum of 4 squares. If n was 0, it would mean that all these numbers x prime, y prime, z prime and w prime have to be 0. But these were the numerically least residues of x y z w, which will mean that each of the x y z w is 0 mod l. If x y z w are 0 mod l, then they are all divisible by l, then their squares are divisible by l square, the sum will be divisible by l square and it would mean that this number lp is also divisible by l square which would mean that l square divides lp, so l divides p. But we are assuming that l is not 1 and we also have that l is strictly less than p. So, these 2 conditions tell us that l cannot be a divisor of p.

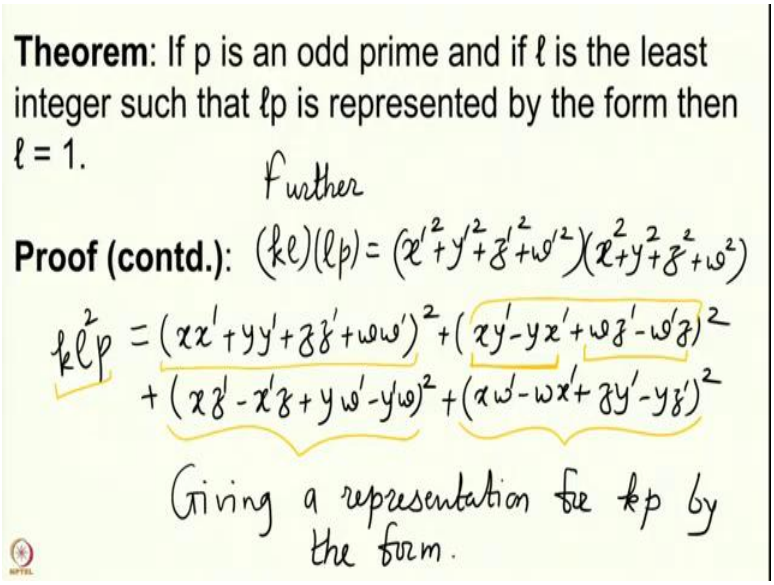
(Refer Slide Time: 29:00)



Therefore, first of all this n is not 0. So, the K is strictly positive. Further we have that each of the x prime square, y prime square, z prime square and w prime square is less than l by 2 square. Remember l is odd, so l by 2 is not an integer, so each of the x prime, y prime, z prime, w prime will have to be strictly less than l by 2 and so their squares are less than l by 2 square, which would mean that their sum is less than 4 times l by 2 whole square, which is l square.

If n which is k into l is less than l square, it would mean that k is less than l. So, now we have obtained a multiple of l which is k into l which is also represented by the form. So, you have k l is represented by the form, lp is represented by the form. Moreover, we know that the form is multiplicative, that means the product k l into lp has to be represented by the form.

(Refer Slide Time: 30:24)



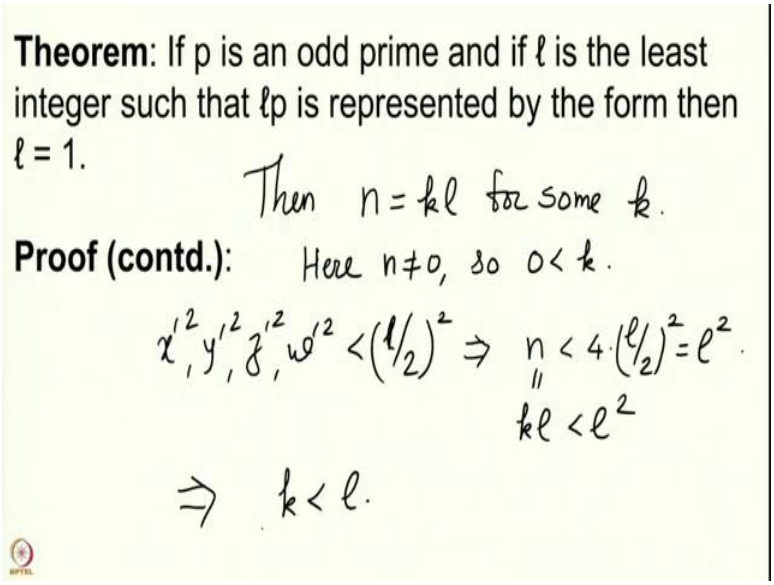
Further, k l into lp which is really x prime square plus y prime square plus z prime square plus w prime square into x square plus y square plus z square plus w square, has to be represented by the form again. And we will write this explicit formula, that we can write the product which is k l square p as a sum of 4 squares. Now, we make one very important observation. In all these sums of squares, here these two numbers see x is congruent to x prime modulo l and y is congruent to y prime modulo l.

So, this number which I have underlined here is 0 modulo l. Similarly, this is 0 modulo l. So, this whole number is 0 mod l square the square is 0 mod l square. Similarly, this is 0 mod l square, this is 0 mod l square, that means that these three terms are divisible by l square, we

have that here we have divisibility by l square, which would imply that this term should also be divisible by l square.

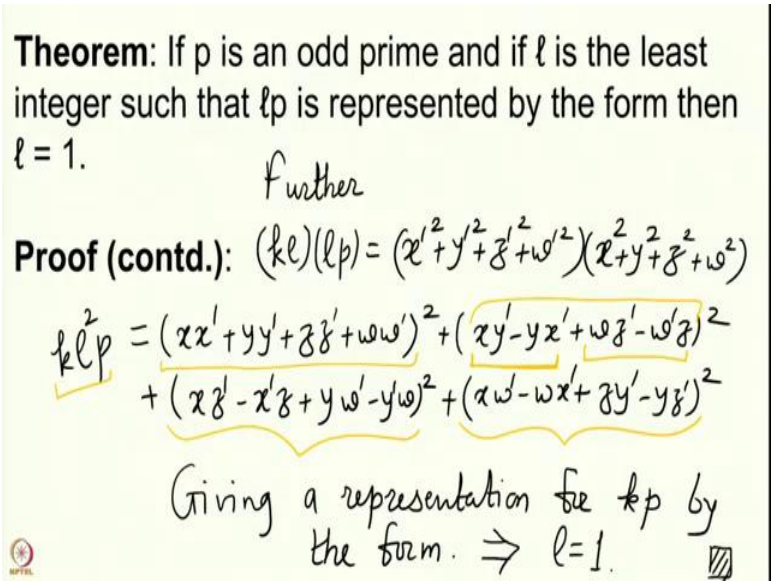
So, you can simply cancel l square from all the terms that you see here giving you a representation for k p by the form that we are talking about, which would mean that k p is also a sum of 4 squares. Here all these 4 squares that appear are each of them is divisible by l square, that means if you divide by l square, you get an integer square, each of the terms in the bracket is divisible by l. So, you have that each of the term divisible by l is an integer and therefore, you are going to get sums of 4 integers to be equal to k into p.

(Refer Slide Time: 33:55)



But k is less than l, this is something that we have already observed and so therefore, you have now a multiple of p by an integer less than l, which is also represented by the given form that is a contradiction because you had started with l being the least.

(Refer Slide Time: 34:19)



So, if your l was not equal to 1, we get a contradiction and therefore, l has to be equal to 1, which proves our result. So, what we have proved, let me just recall is that the form is multiplicative, two is represented by the form and every odd prime is represented by the form. Now, getting Lagrange's theorem from this is very easy. We will see it the first thing in our next lecture, but we have to stop here for the time constraint. See you soon. Thank you very much.