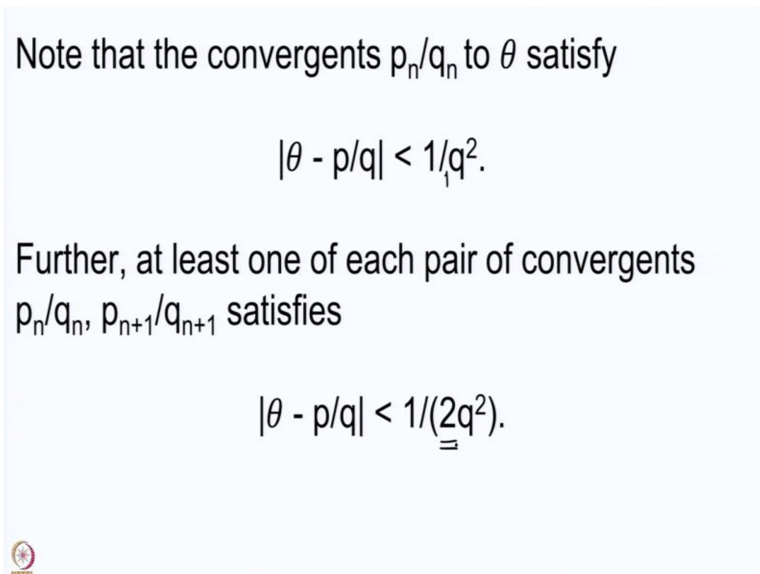
A Basic Course In Number Theory
Professor Shripad Garge
Department of Mathematics
Indian Institute of Technology, Bombay
Lecture - 56
Convergents Are The Best Approximations - 1

Welcome back. We are discussing how well we can approximate a given real number by means of rational numbers. And we saw towards the end of our last lecture that the convergents for the natural continued fraction expansion that we have for our real number will satisfy the property that they are going to give us not us good approximations but better approximations in some sense. So let us quickly go through the results that we have proved in the last lecture.

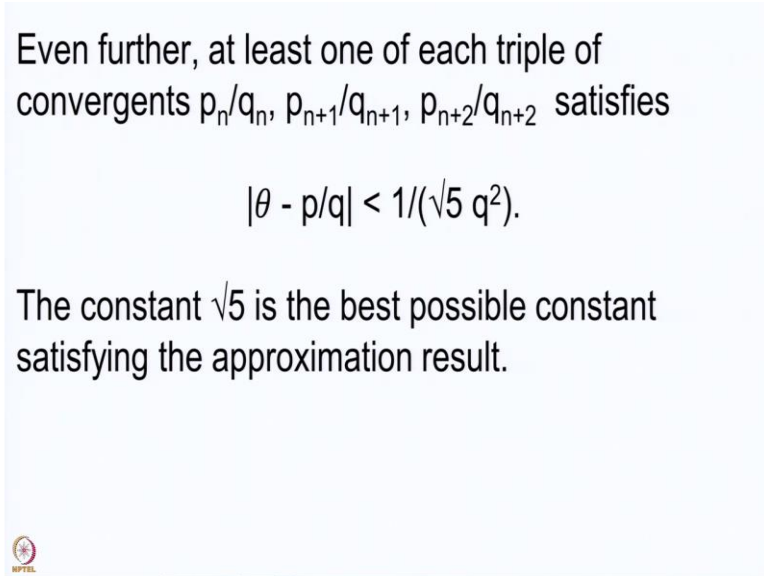
(Refer Slide Time: 00:52)

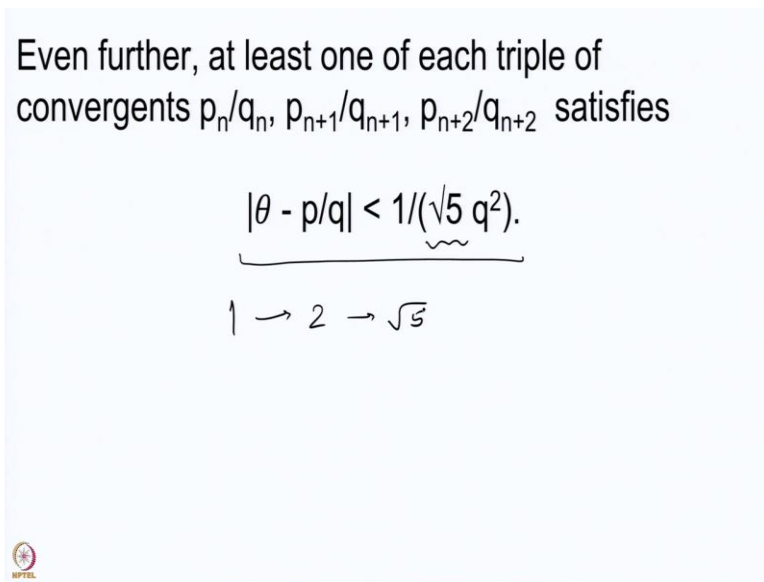


What we have is that the convergents certainly satisfy the inequality that theta minus p by q is less than one upon q square, so this is something that was anyway clear from the construction that we had for the continued fraction expansion for a given theta.

So this was anywhere there but we saw that we could improve this further to see that at least one from each pair of consecutive convergents will satisfy that theta minus p by q is less than 1 upon 2 times q square. So the constant which was 1 here has now become 2. The constant which was 1 here has now become 2. So this is an improvement.

(Refer Slide Time: 02:04)





Further, we actually saw that this is also not the final result. If you consider one of consecutive triples of convergents, then you even get a better result, which is that 2 can be further improved to root 5 q square.

Now, this is definitely going to give us a sequence of rational numbers converging to any given real theta, with the property that theta minus p by q is less than 1 upon root 5 q square. And one may wonder, whether you can improve it further. We went from 1 to 2 to root 5 and the natural question would be, can you do even better?

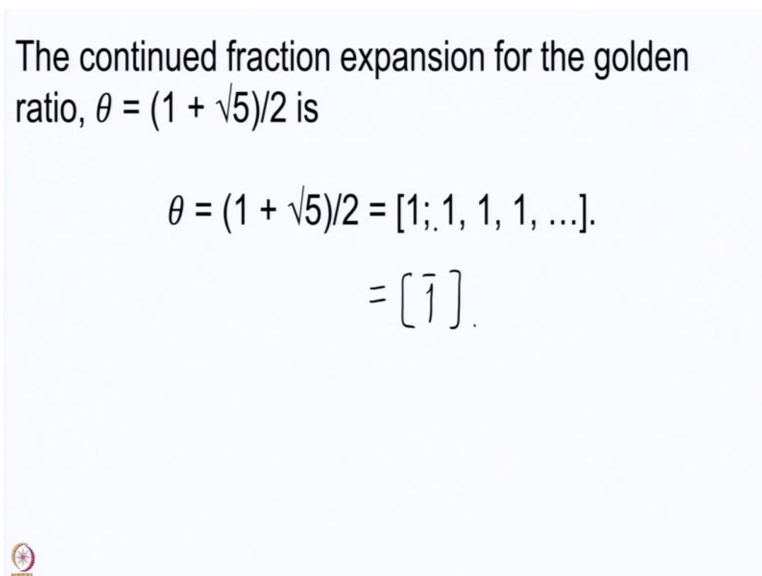
And the answer to that is that no, you cannot do better than root 5. Root 5 is the best possible constant satisfying this approximation result, which means that there are some certain real numbers for whom root 5 is the only one which will give you this property.

And then you may ask that we will remove those particular real numbers, then can you say that root 5 can be improved? Then the answer is yes, root 5 can be improved to root 8. Again, that turns out to be the best possible for there are some real numbers for whom root 8 is the only answer and it cannot be improved further.

If you remove them also, then root 8 can be further improved but ultimately, these sequences, these constants converge to the number 3. So the only thing that we have been able to improve is from 1 to 2, and then we are going towards 3 by removing some certain sets of real numbers.

But this is, in some sense, a good property of continued fractions, what we really want to say is that continued fractions give you, the convergents in the continued fraction expansion give you the best approximations. That is what we want to say. And we will go to prove that result today.

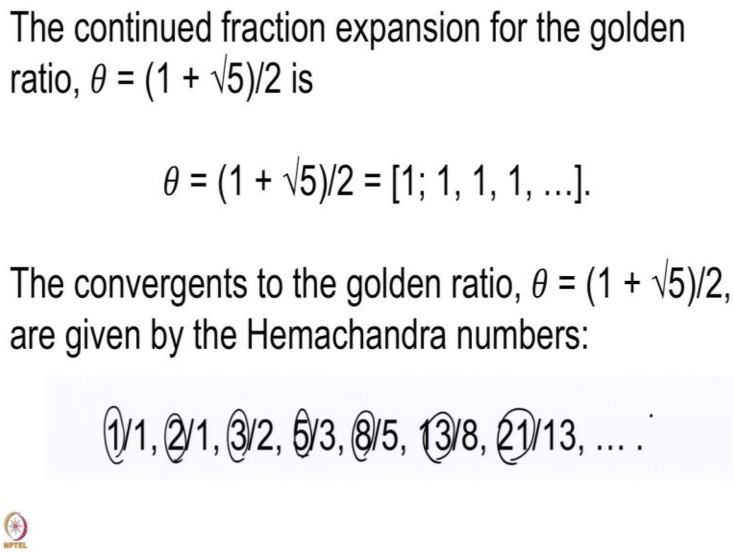
(Refer Slide Time: 04:08)



Before that, let us also recall that we have obtained this continued fraction expansion for the golden ratio, theta equal to one plus root 5 by 2. And we noted that this is periodic because it simply repeats after, from the first step onwards. So this is periodic and we can write it as one bar because that is the way we denote when we have periodic expansion. I will define as explicitly in a while.

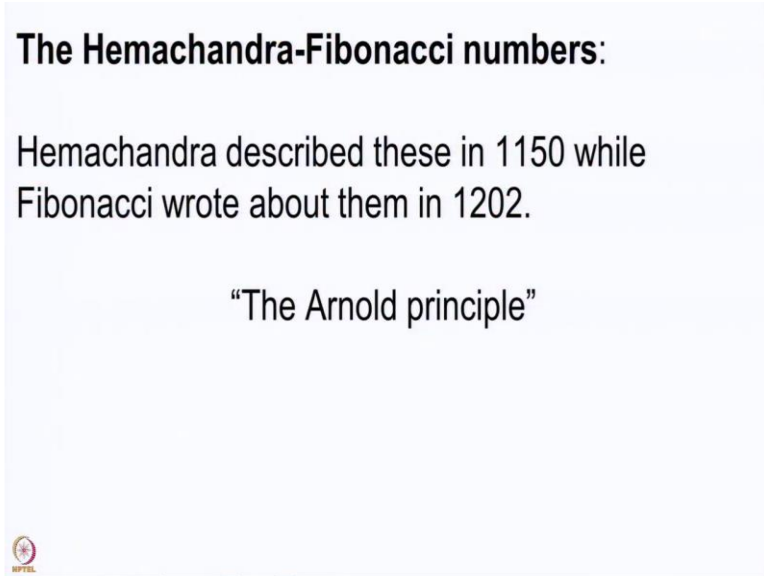
So there is, this is a very important number from many points of view but one point of view is that its convergents are some of the interesting numbers.

(Refer Slide Time: 04:54)



These are the Hemachandra Fibonacci numbers. So these are the 1 by 1, 2 by 1, 3 by 2, 5 by 3, 8 by 5, 13 by 8, 21 by 13, and so on. After all, you just have a 1 at the end and you are simply going to add them up. So these, if you notice, this sequence of numbers, 1, 2, 3, 5, 8, 13, 21, the next one will be 34, and so on, these are what are earlier known as Fibonacci numbers.

(Refer Slide Time: 05:34)



But as we saw in the last lecture also that these were described by Hemachandra in 1150, more than 50 years back; more than 50 years before Fibonacci actually wrote about them. Fibonacci's account of these numbers is in 1202. Although it is 800 years back, but even then, Hemachandra knew more than 50 years before Fibonacci about this.

And the motivation for Hemachandra was as follows, that if in our music or in poetry, we have two types of syllables. There are the short syllables and then there are the long syllables. And the basic question is that if you had 8 beats, let us say, and you wanted to fill it with, either a short syllable, a sequence of short and long syllables, then how many ways are there to fill it with.

So one observes that if you put the last one to be a long syllable, then you have 6 beats before that, where you have freedom to fill then with anyway, and that gives you the way to fill a 6 beat sequence with long and short syllables. So this is when you have the last beat to be large syllable.

Similarly, if your last beat is the short syllable, then you have 7 beats left and you can fill them in any way as you want. Therefore, the number of ways to fill 8 beats sequence with short or long syllables is the sum of the corresponding numbers for 6 and 7. This is how Hemachandra had come to these numbers and then one learns that Fibonacci also came to these numbers by some biological motivation.

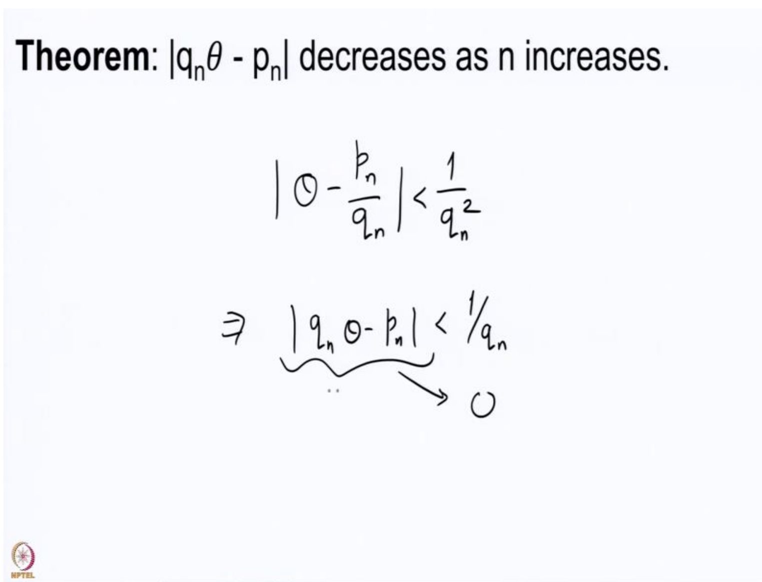
It is okay if these numbers are described as Fibonacci numbers but once you know that these were also described by Hemachandra much before Fibonacci then we should really call them Hemachandra Fibonacci numbers. In passing, in would like to mention this principle called The Arnold Principle

The principle has, the principle states that any mathematical term which is named after a mathematician is not discovered by the corresponding mathematician. That applies to the Fibonacci numbers because they are named after Fibonacci, but they are discovered by Hemachandra much before Fibonacci.

This happens quite regularly in mathematics. The names to certain terms are given by people who made them popular or people who write first about them. Although they may not discover about them, but they, their first written account of the concept might be due to them and so on.

So these are the things which routinely happen, what is important is that whenever we come to know about the original discoverer, we should keep in mind who is the discoverer and who is the person, who is not discoverer. A funny thing is that Arnold Principle does apply to Arnold Principle. It is discovered by somebody else before Arnold. I will let you search on Google for the Arnold Principle and we continue with our mathematics.

(Refer Slide Time: 08:46)

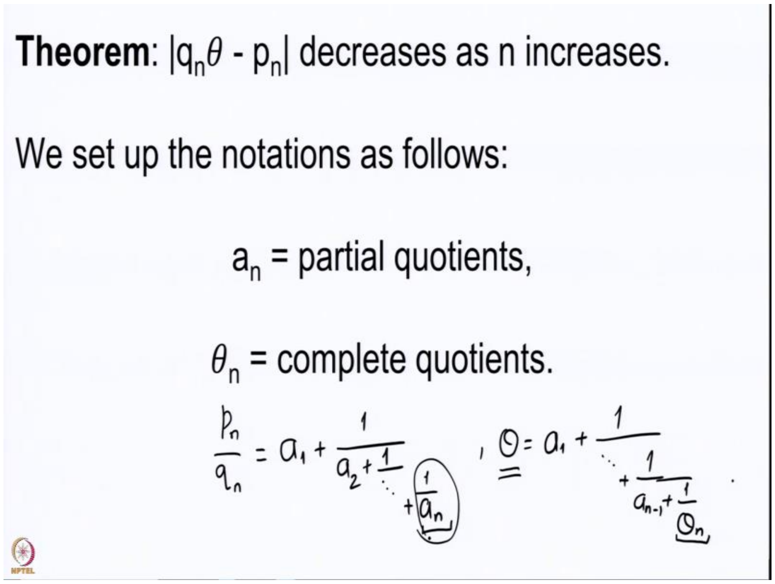


So this is a theorem that I would now like to prove for you. Let me just mention one small thing here. What we have noticed is that theta minus pn by qn is less than 1 upon qn square. These are the properties of the convergents that we have already noticed.

Therefore, if you have qn theta minus pn, then we know that this is less than 1 upon qn. So this difference of course goes to 0 as qn moves to infinity. This was lemma four in our proof of existence of a continued-fraction expansion for a given real number.

So, we know that this goes to 0 but what we want to prove here is that this sequence which goes to 0 is actually a decreasing sequence. That means that the property for the number for n is smaller than the number for n minus 1. The number for n plus 1 is smaller than the number for n. So it is actually a decreasing sequence. This is what we are going to prove.

(Refer Slide Time: 09:55)



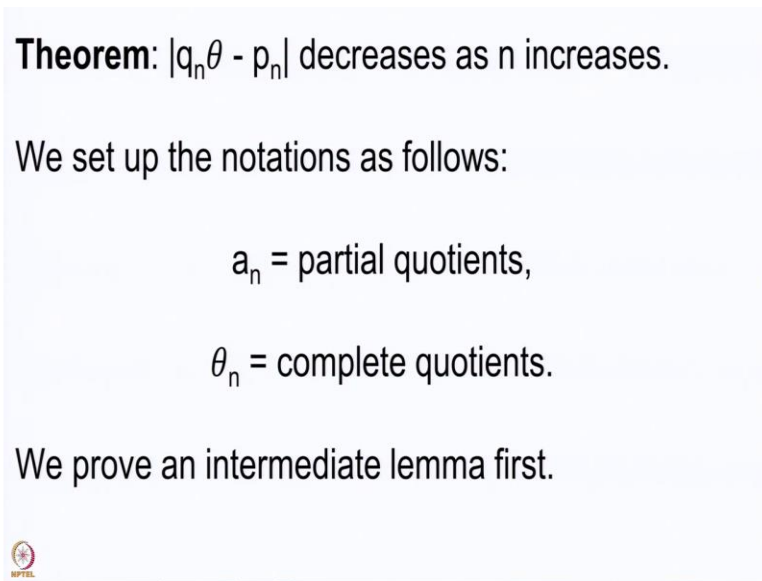
Let me also set up some notations because these are the things that will keep coming again and again. So whenever we have a theta, the corresponding integer is an, which appear in the continued fraction expansions will be called partial quotients to theta. And then we will have these theta n, these will be called the complete quotients.

So to make sure that we are on the same page as far as the notations are concerned, this is a1 plus 1 upon a2 plus 1 upon dot dot dot, ultimately you have an 1 upon an. And theta is a1 plus 1 upon dot dot dot then we have an minus 1 plus theta n. So plus 1 upon theta n.

So instead of an, we would have theta n. Once you have theta n in place of an, you do get your theta back. Therefore, these are called complete quotients. They give you the complete theta. These are after all quotients because it is 1 upon theta n that you have here in the continued fraction expansion. So they are called complete quotients.

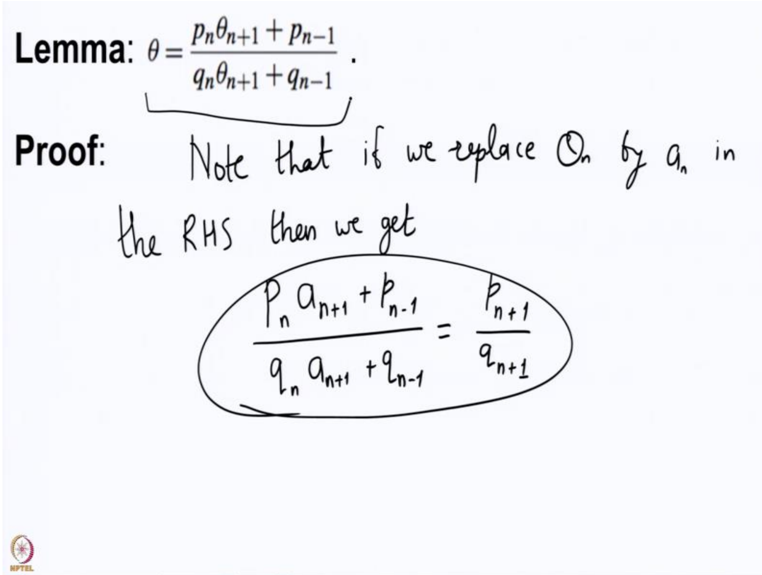
And the ans which come here, these are also quotients but they give you partial information. So they are called partial quotients. So an are the partial quotients and theta n are the complete quotients, that is our notation.

(Refer Slide Time: 11:41)



And before we prove this result, we need to prove an intermediate lemma first, which is similar to one of the results that we have proved earlier.

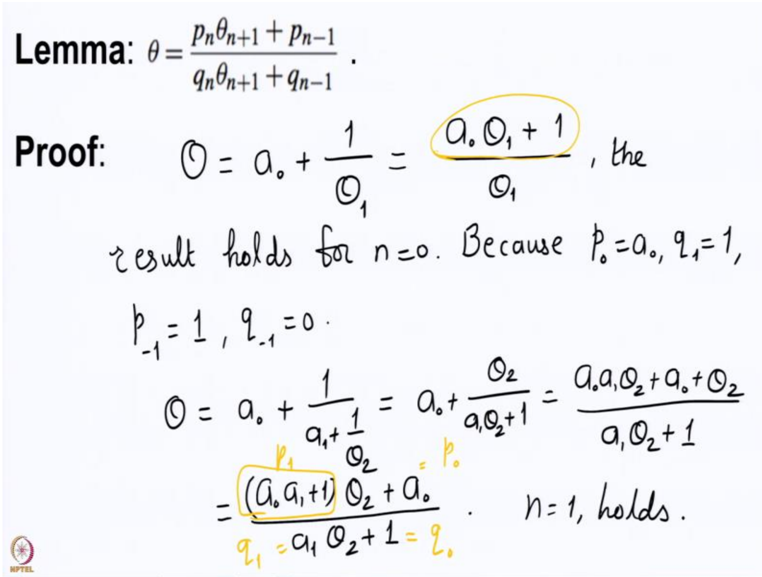
(Refer Slide Time: 11:56)



So the lemma says that your theta can be expressed in terms of pn, pn minus 1; qn, q minus 1, and the complete quotients theta n plus 1. So note, first of all, that if we replace theta n by an in the right-hand side, then we get pn an plus 1 plus pn minus 1 upon qn 1 n plus 1 plus qn minus 1. And if you remember the recursion formulae for the numerators and denominators of the convergents, then you will see that this is, these are nothing but pn plus 1 upon qn plus 1.

So this is really, if you put instead of theta, the partial quotients a then what we get is nothing but another proof for the recursion formulae and this proof is also quite similar to the proof that we had in proving the recursion formulae.

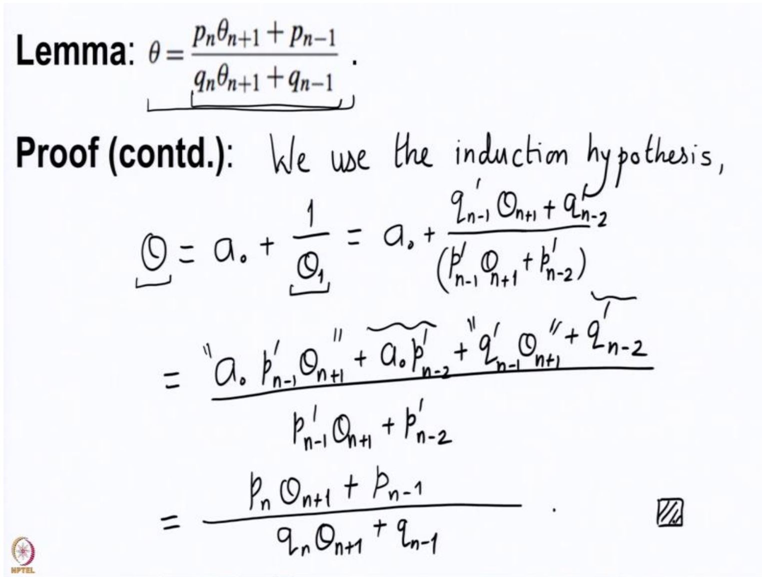
(Refer Slide Time: 13:18)



So here, we let, so here, of course, we have theta equal to a0 plus 1 by theta 1. And we see that this is nothing but a0 theta 1 plus 1 upon theta 1. So the result holds for n equal to 0 because p0 is a0, q1 is 1, p minus 1 is by convention you can take it to be 1, and pq minus 1 is taken as 0 by convention. If you are not happy with this, we can go to one more step and check that the result holds there also.

And here, of course, we have this to be a0 a1 plus 1 into theta 2 plus a0 upon a1 theta 2 plus 1. So here, a0 a1 plus 1 is our p1, as you will see that we can put instead of theta, if you put a, we get p and q. So this is our p1, this is our q1, this is our p0, this is our q1 and this is the q0. So the result holds for n equal to 1 and we are going to prove for the higher cases. So for n 2 onwards, we are going to use the induction hypotheses.

(Refer Slide Time: 16:13)



So we notice, first of all, that a, theta is a naught plus 1 upon theta 1, and assuming the induction hypotheses for n minus 1 for all real numbers, we can use it for theta 1. And then, we would have that the complete quotients for theta 1 are also the complete quotients for theta, there will be just the change of one number there. And the convergents for theta 1 are denoted by p prime and q prime as we have done in some of the last lectures.

So, theta 1 will have the property, we are going to assume this for n minus 1 for theta 1. So we are going to have p prime n minus 1 theta n plus 1, the nth complete quotient for theta 1, which is what we should take here is the n plus 1th complete quotient for theta.

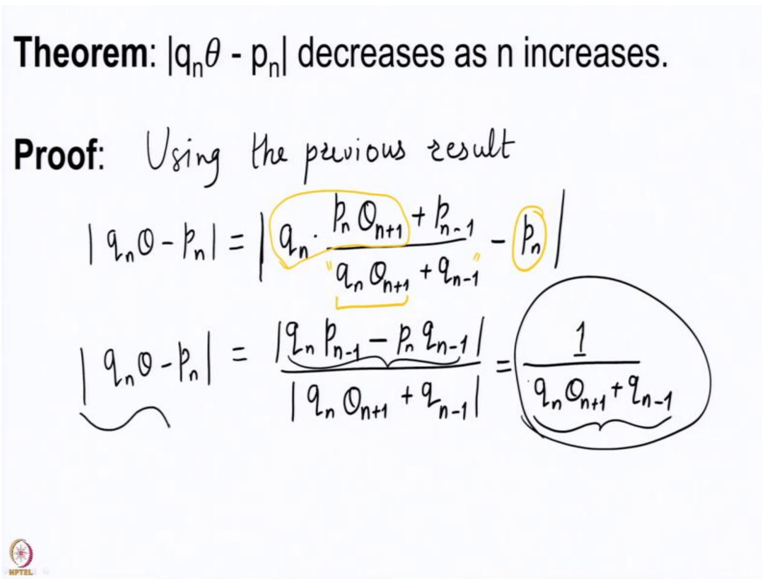
So we will have theta n plus 1 there, plus p prime n minus 2 divided by the corresponding expression in terms of q, qn minus 1 prime theta n plus 1, plus qn minus 2 prime. And now, if we simply expand this out, we get a naught p prime n minus 1 theta n plus 1 plus a naught p prime n minus 2 plus q prime n minus 1 theta n plus 1 plus q prime n minus 2 divided by the same constant, p prime n minus 1 theta n plus 1 plus p prime n minus 2.

And now, the relation between the pj prime and pj or qj prime and qj is that qj is p prime of j minus 1. So by that, we already obtained the expression for our denominator, which is qn theta n plus 1 plus qn minus 1, the one that we wanted. And the numerator will have this terms, so you have a0 p prime n minus 1 into theta n plus 1, plus q prime n minus 1 into theta n plus 1, which is nothing but pn into theta n plus 1. And similarly, a0 p prime n minus 2 plus q prime n minus 2 is nothing but pn minus 1, which completes the proof.

So this proof is very similar to what we had done in obtaining the recursion formula for pn and qn. And there, we had obtained the expressions for pj prime, pj qj in terms of pj prime and qj prime, which is what we have used here using the a0 and so on.

So the real number theta can be expressed in terms of the complete quotient theta n plus 1 using pn pn minus 1, qn and qn minus 1. This is something that we will have to remember when we go our, our next result which is that this mod qn theta minus pn decreases as n increases.

(Refer Slide Time: 20:12)



So this is the result that we now want to prove. This is a very important result. It is one of the stepping stones towards proving that the convergents give the best possible approximations for any real number theta.

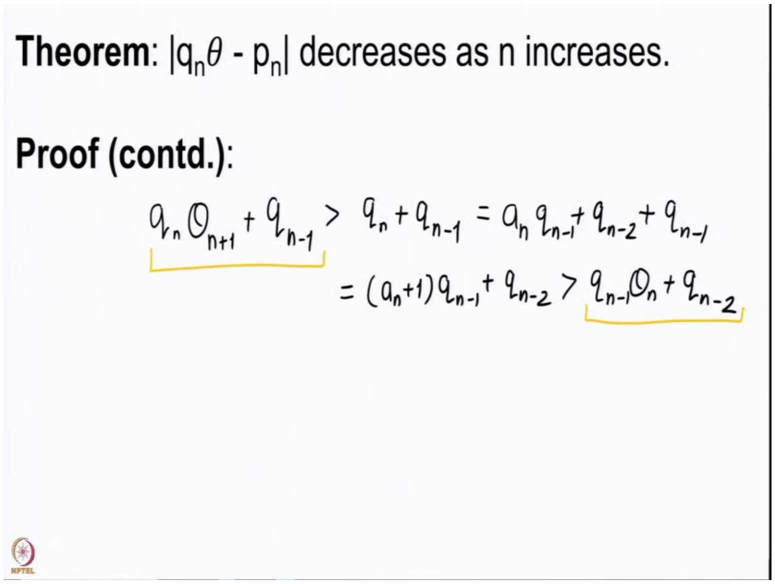
So we notice, by previous, using the previous result, mod qn theta minus pn, we have qn into pn theta n plus 1 plus pn minus 1 upon qn theta n plus 1 plus qn minus 1 minus pn. And now, we observe that this pn will be multiplied to the denominator, we are going to clear the denominator which would mean that we will take this term multiply it to pn and keep the common denominator.

So pn with qn theta n plus 1 will come with a negative sign, whereas this qn pn theta n plus 1 is going to come with a positive sign. So these two are going to get canceled anyway. Therefore, let us write the remaining parts, which are qn pn minus 1, qn pn minus 1 minus pn qn minus 1, and the denominator is qn theta n plus 1 plus qn minus 1. We put a modulus sign to both the numerator and denominator.

Now, we observe here that this is a term in the numerator that we have seen quite often. It is plus 1 or minus 1. Therefore, when you put it under the modulus sign, you are going to just get 1. And here, we have a positive quantity. That is because qs are all positive and theta from 1 onwards are also positive. In fact, the complete quotients theta are bigger than 1, if they are not 0. Even is theta is 0, you have the corresponding qn minus 1 sitting there. So this quantity in the denominator is always a positive quantity.

So, what we have proved is that mod qn theta minus pn is equal to this quantity, 1 upon qn theta n plus 1 plus qn minus 1. And now, we simply need to observe that the denominator here increases as n increases. That would tell us that the mod qn theta minus pn decreases as n increases.

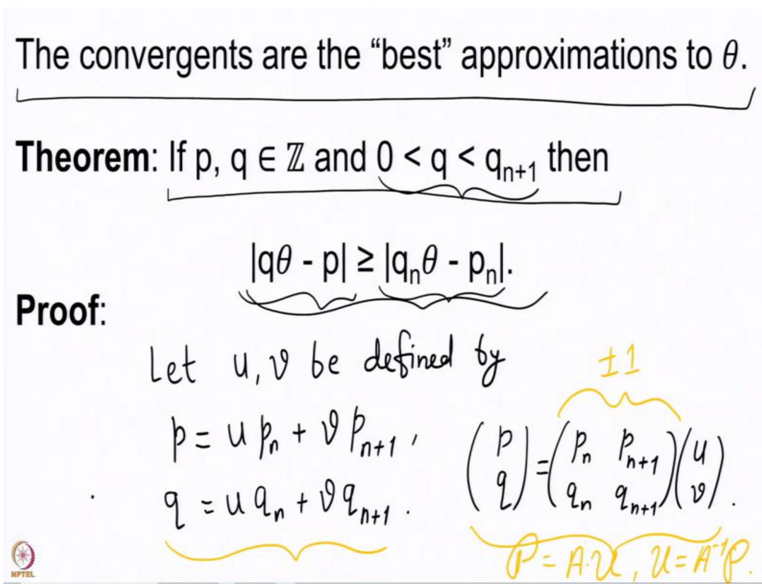
(Refer Slide Time: 23:38)



So we observe that qn theta n plus 1 plus qn minus 1 is, in fact, bigger than qn plus qn minus 1 because theta n plus 1 is bigger than 1. But the recursion formula for qn will tell you that this is nothing but an qn minus 1, plus qn minus 2, and then we have an extra qn minus 1. Therefore, this is an plus 1 qn minus 1 plus qn minus 2. an, remember was the integral part of theta n. So we have that this is strictly bigger than qn minus 1 theta n plus qn minus 2.

Therefore, this expression for n is strictly bigger than this corresponding expression for n minus 1. As n increases, the denominator is increasing; and therefore, mod qn theta minus pn which was the reciprocal of these expressions will decrease as n increases. So this is one very important result that we have and let me just prove one small result for you before we complete our lecture for today.

(Refer Slide Time: 25:32)



We note here that this convergents, we are going to prove this very important result. That the convergents are the best approximations to theta. That means, first of all, if I take any p and a natural q with the property that q lands between 0 and qn plus 1, then for any p, mod q theta minus p is bigger than or equal to mod qn theta minus pn.

So once you have a small bound on qn, then the q theta minus p will never be smaller than qn theta minus pn. Remember that we are, we know that this is going to 0. But it is going to 0 better than any such q theta minus p provided your q is smaller than qn plus 1.

So this small quantity is, this very small result is very important. There is something that we should note here and which is as follows. So let me just note that and then we will see the proof.

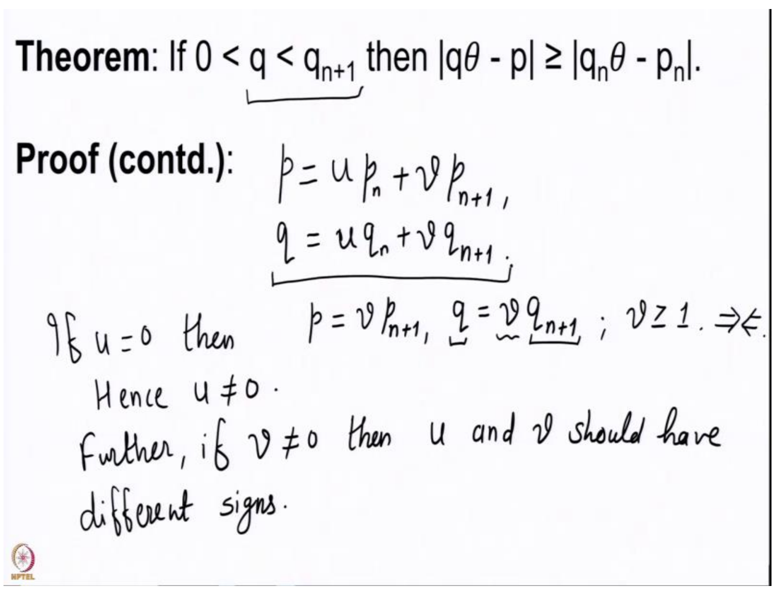
So let u comma v be defined by p equal to u pn plus v pn plus 1, and q equal to u qn plus v qn plus 1.

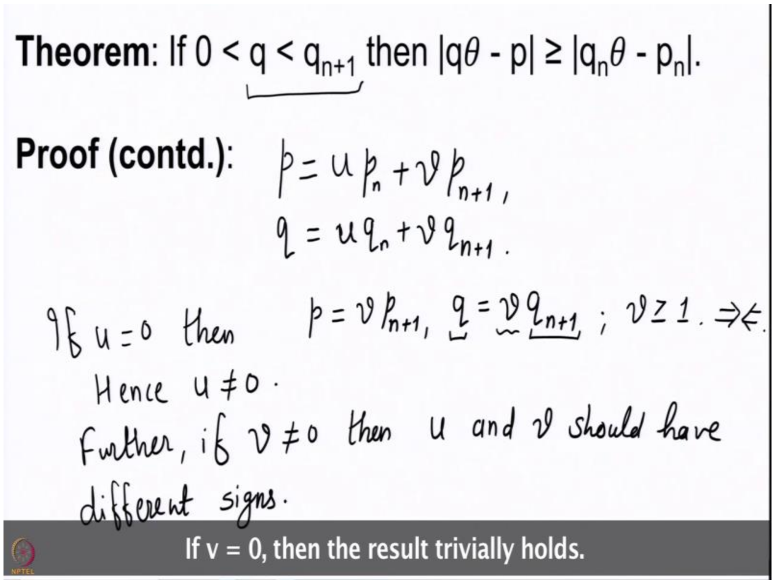
So note here that we are defining the two integers u and v. But we are defining them in a convoluted way. We are not defining them straight away. The reason for defining this is that these two expressions are going to be important for us. But the question still remains, whether we can indeed define u and v in this way.

So we can do that because this can be written as the column matrix p q equal to pn, pn plus 1, qn, qn plus 1; a two by two matrix into the column matrix u v. This is what we have in this equation. So the equation, this equation, these pair of equations is the same as one equation in matrices. But we notice that this matrix has determinant plus or minus 1.

Therefore, it is an invertible matrix in integers. The inverse also has integer entries. So you can simply, if this matrix, so let us call this script p equal to the matrix a into script u, where script p and script u are column vectors, then a is invertible. And we can actually solve for u in terms of a inverse and script p. So the u and v as we have defined do indeed exist and can easily be computed using pn, pn plus 1, qn, and qn plus 1.

(Refer Slide Time: 29:06)





So what is important for us is the pair of equations. Let me simply write them down because the expression is very important. So p is u pn plus v pn plus 1, and q is u qn plus v qn plus 1. Now, we observe first of all, that u cannot be 0 because if u is 0, then p is v times pn plus 1 and q becomes v times qn plus 1.

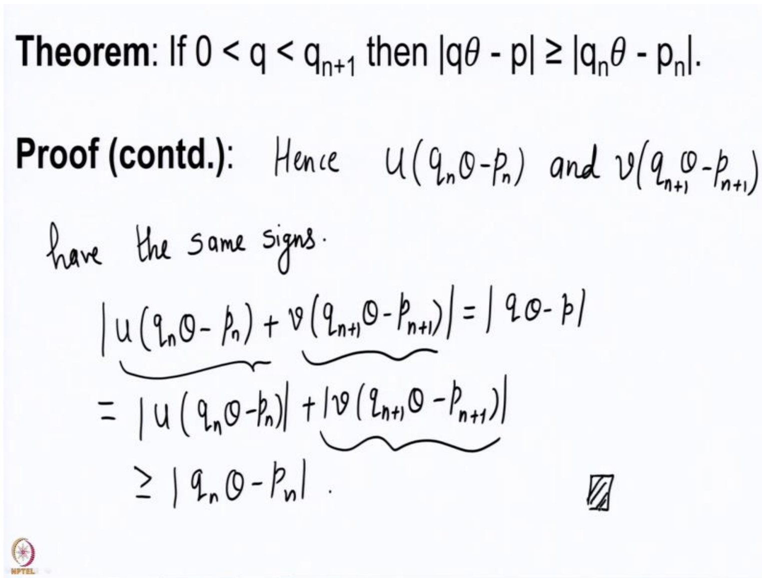
But q is positive, remember, our convention is that the denominator is always a positive integer. q is positive, qn plus 1 is positive, and then v, if you take v also to be 0, that would give you a contradiction because q then becomes 0. So q positive, qn plus 1 positive means that v is at least 1. And this is a contradiction because you had started with q to be less than qn plus 1. So u equal to 0 cannot happen. So hence, u is non-zero.

Then, further if, v is non-zero, then u and v should have different signs. This is again observed from this same equation. If u is positive, clearly v cannot be positive because then, u into qn, which is a positive number, if v is also positive then we are going beyond qn plus 1. So if u is positive, v cannot be positive.

If v is positive, if u is negative then v cannot be negative because u into qn is going to be negative number. Remember, all qns are positive. So u qn will be negative, v qn plus 1 will be negative, which forces q to be negative.

So whenever u is positive, v cannot be positive. If u is negative, v cannot be negative. So u and v have different signs. We also know one more such instance when the signs are different, which is that theta minus pn upon qn and theta minus pn plus 1 upon qn plus 1, these also have different signs because we know that theta is sandwiched between any two successive quotients; any two successive convergents.

(Refer Slide Time: 32:35)



So hence, u into qn theta minus pn and v into qn plus 1 theta minus pn plus 1 have the same signs. Both are either negative or both are positive. Because u and v, the signs are different, and similarly, the things in bracket that we have, their signs are different.

So whenever you have the things of the same sign, their modulus will have the corresponding property that u into qn theta minus pn plus v into qn plus 1 theta minus pn plus 1; this we know already because u and v were defined in a particular way. This is q theta minus p but because these have the same signs, both of these, we have that this is equal to mod of u into qn theta minus pn plus mod of v into qn plus 1 theta minus pn plus 1. This is because they have the same sign.

This we notice is at least 0 and u is at least 1 because u can never be 0. So we get that this is bigger than or equal to qn theta minus pn. So if you have just the denominator q to be less than qn plus 1, then q theta minus p will have the property that qn theta minus pn is smaller than or equal to that.

We are going to use this, and going to prove in the next lecture that any rational number which tries to give us slightly better approximation to theta has to be actually a convergent to theta. So we will see that in the next lecture. See you then. Thank you very much.