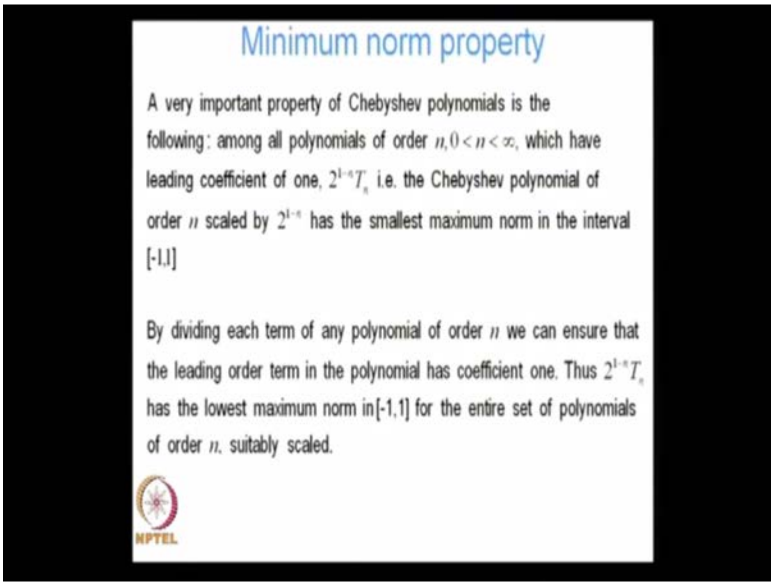
Numerical Methods in Civil Engineering
Prof. Arghya Deb
Department of Civil Engineering
Indian institute of Technology, Kharagpur

Lecture - 33
Orthogonal Polynomials-II

In lecture 33 of our series on numerical methods in civil engineering, we will continue with our discussion on orthogonal polynomials.

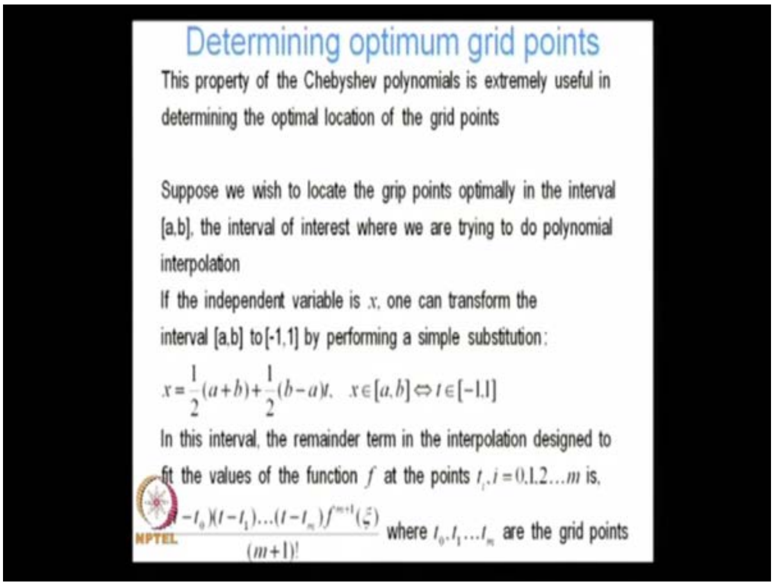
(Refer Slide Time: 00:29)



Recall that in the last lecture I talked about Chebyshev polynomials and I will at the end of the lecture while talking about the very important property of Chebyshev polynomials, namely the property that in the interval minus 1 to 1 the Chebyshev polynomials 2 to the power 1 minus n T n has the smallest maximum norm. Among all polynomials with coefficient equal to 1, among all polynomials with coefficient equal to1 the Chebyshev polynomial has the maximum has the lowest maximum norm that means that at any point in the interval, if I plot any other polynomial with coefficient 1 and I plot the Chebyshev polynomial 2 to the power 1 minus n time T n the magnitude of the Chebyshev polynomial will be smaller than any other polynomial.

Now, by dividing each term of any polynomial of order n we can ensure that the leading term of that polynomial is 1. So, for the entire set of polynomials of order n suitably scaled, we can say that the Chebyshev polynomial is the, has the lowest maximum norm.

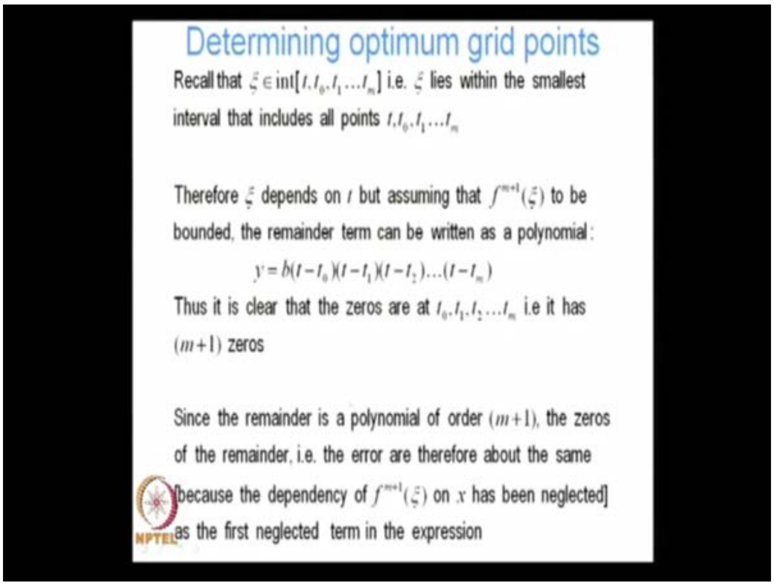
(Refer Slide Time: 01:45)



And we saw that this has particular implications. If we want to find out the optimum location of the grid points, what is the optimum location of the grid points? Well I know certain intervals say a b and I want to find out where in that interval, should I place my grid points, where should I know the function values in order to minimize the error so that, the error any polynomial approximation will have a truncation error and if I want to minimize the truncation error it turns out that it depends a lot on where I, sample my function and that is where the Chebyshev polynomial concern suppose, we wish to locate the grid points optimally in the interval of interest, where we are trying to do a polynomial interpolation and that interval may be a b but in by doing a suitable change of variables transformation of variables like this we can transform that into minus 1 and x becomes t.

And I know that in this interval the remainder term in the interpolation is like this we have seen earlier, t minus t 0 t minus t 1 t up to t minus t m f m plus 1 xi divided by factorial m plus 1 where, t 0 t 1 through t n are the grid points. Now, suppose there is a dependence on this, in this error term on t.

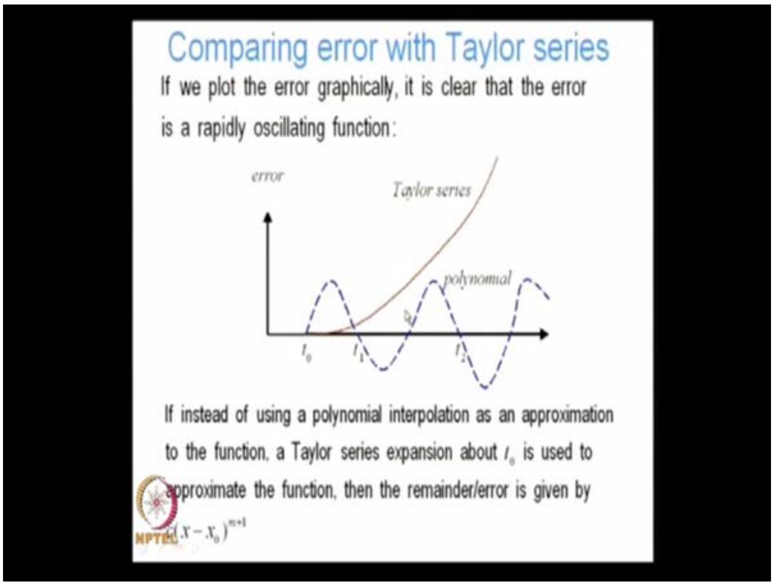
(Refer Slide Time: 03:27)



Because first of all xi must belong to an interval which is the smallest interval that includes all these points. So, xi depends on t but if we assume that f to the power m plus 1 xi is bounded that there is, it is bounded then the reminder term can be written like this, where I can represent, I can think of b as the bound.

So, instead of writing it like this I can represent this whole thing as b. And then, I get this expression b into t minus t 0, t minus t 1 t minus t 2, t minus t m. So, the 0s of this order t 0 t 1 t 2 through t m now. Since the remainder is a polynomial of order m plus 1 the 0s of the remainder are about the same as the 0 of the first neglected term they are not exactly the same because I have ignored the dependency on t of f m plus 1 xi this should be t I am sorry, it is a same thing because there is the transformation of variables. So, I have ignored the dependency of f m plus 1 xi on t but it turns out that since I have ignored that I can assume that the 0s of this remainder are the same as the first neglected term in the series. Now, if we plot that error graphically, this error basically.

(Refer Slide Time: 05:04)



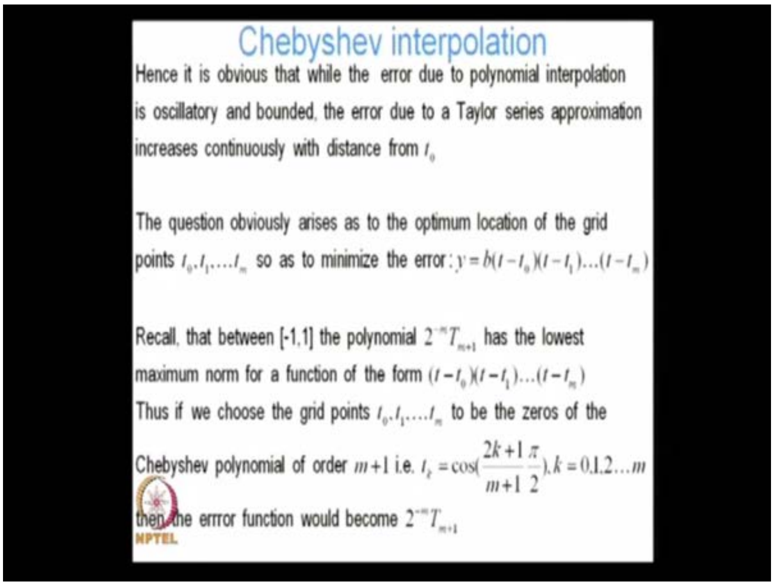
If I plot it graphically, then I will see that it will be something like this, why because it has to be 0 at t 0 it will be 0 at t 1 it will be 0 at t 2 t 2 t 3 and so on and so forth. So, basically it will be 0 at all these points at t equal to t 0 t equal to t 1 t 2 and t 3. However, look at this if, I instead of doing a polynomial approximation if I try to expand the function in a Taylor series about t 0 if I try to expand the function. So, why do we do an interpolation? Well, we know the function values at certain points and we want to know the function value at some other point, where the function value is not known that is why I do interpolation.

Now, suppose instead of doing that I say that, I know the function value at a certain point and I may be, I know higher order derivatives of that function also at that point I can expand that function in a Taylor series about that point I can expand it in a Taylor series and if I expand, I can find out the value at another point. But, the error in a Taylor series is of this form x minus x 0 to the power m plus 1 given that this is the remainder term that is of this order so if I expand the function in a Taylor series about t 0 the error term which is like this, will be like this as you go away from t 0 the point where I know the function value and it is derivatives the error is going to go as a power in again this should be t minus t 0 I apologize so t minus t 0 to the power n plus 1. So, it is going to increase like this.

However, if we do if our error term is of this form you can see that it will be oscillatory. But, more than that we have, we want something more than that we not only wanted to be oscillatory, we do not want to increase monotonically. But, we also want these values to be bounded I want that error term to be as small as possible. I do not wanted to grow monotonically that is good it is oscillatory that is 1 good thing but another requirement is that those values they must not be very big and that is where my Chebyshev.

So, any interpolation any polynomial interpolation is going to be better than Taylor series you can see that this is oscillatory and this is growing monotonically but if I locate my grid points at the location of the 0s of the Chebyshev polynomial in that case I am assure that this thing will also be the lowest possible this error term will also have the lowest possible magnitude. So, then it will not only be oscillatory but it will also have this smallest possible magnitude.

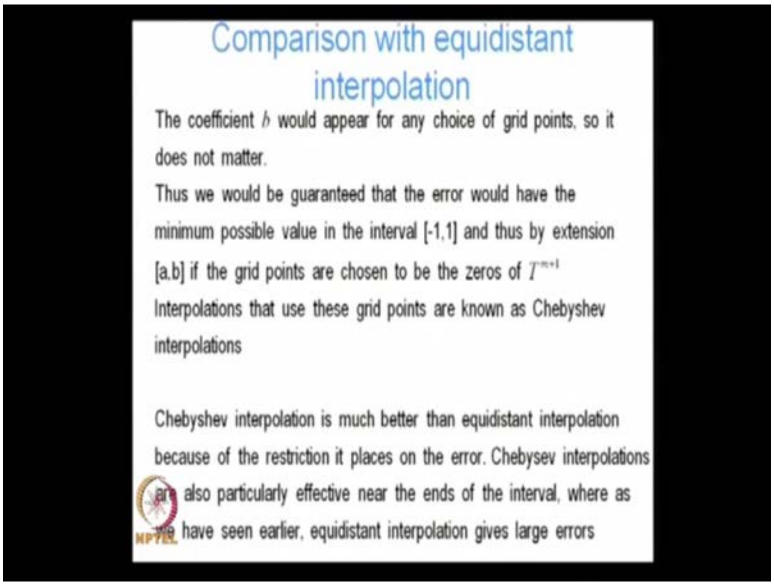
(Refer Slide Time: 08:14)



So, hence it is obvious that while the error due to polynomial interpolation is oscillatory on bounded the error due to our Taylor series approximation increases continuously with distance from t 0 the question obviously arises as to the optimum location of the grid points t 0 t 1 through t m. So, as to minimize that error y is equal to b t minus t 0 and so on and so forth we recall that between minus 1 and 1 the polynomial 2 to the power minus m T m plus 1 has the lowest maximum norm for a function of this form why of this form, because this is the form where the leading order term is going to have coefficient 1.

So, it has the lowest maximum norm thus, if we choose the grid points t 0 t 1 t m to the to be the 0s of the Chebyshev polynomial of order m plus 1 of order this is after all the Chebyshev polynomial of so this is a m plus I-th order polynomial. So, we choose this t 0 t 1 t m to be the 0s of the Chebyshev polynomial of order m plus 1 that is t k is of to be these which we saw earlier then the error function would become this function, 2 to the power minus m T m plus 1 this is the error function would basisally become this function.

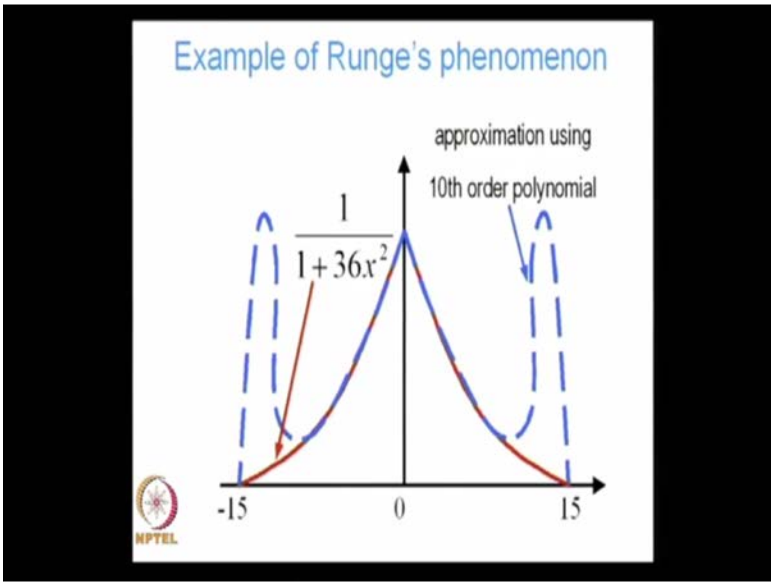
(Refer Slide Time: 09:50)



The coefficient b would appear for any choice of grid points so it does not matter b is going to be there wherever i choose t 0 so that, is not important. I am just concerned with this part I am ignoring the b because b is constant whatever t 0 t 1 I take b is always going to be there. So, I am looking at this part, and I am trying to minimize the norm of the magnitude of this part the magnitude of a function due to this part.

Thus we would be guaranteed the error would have the minimum possible value in the interval minus 1 and thus by extension a, b. If the grid points are chosen to be the 0 of T m plus 1 and interpolations that use these grid points are known as Chebyshev interpolations. Is that clear, they might choose a different polynomial basis. But, the grid points if they choose like that then we are assure that the error term would behave like the T m plus 1 to the part 2 into m 2 to the power m minus 1 times T m plus 1, is that clear. So Chebyshev interpolation is much better than equidistant interpolation. We have seen that before but just recapitulate because of the restriction it places on the error. Chebyshev interpolations are also particularly effective near the ends of the interval were as, we have seen earlier equidistant interpolation gives larger errors.

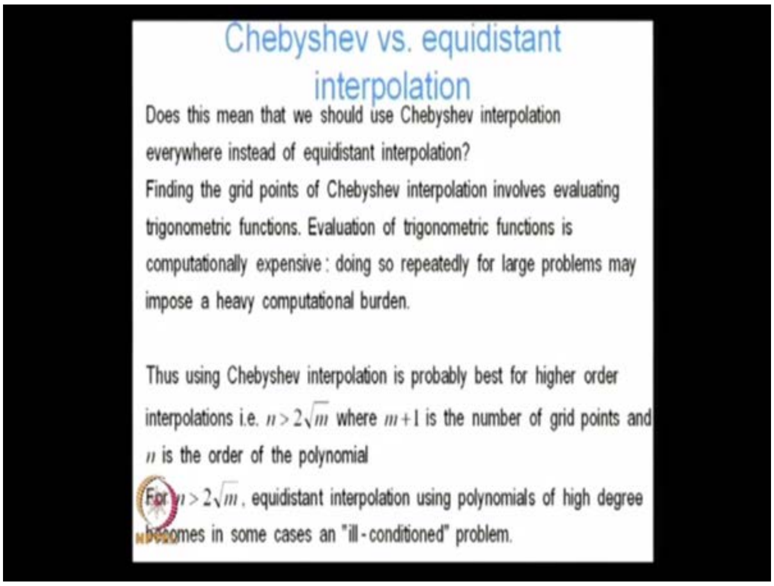
(Refer Slide Time: 11:28)



I looked at this example earlier when we are trying to fit this function. We have no function values for this 1 and I am trying to fit a polynomial to this and if I do equidistant interpolation I get something like this very large errors near the end points and this I do for these errors are particularly severe for higher order polynomials.

As you increase the order these errors becomes large and I think at that time we saw or maybe we did not in case if we use a Chebyshev instead of using equidistant interpolation it will use Chebyshev interpolation. basisally, you choose your grid points to be the 0s of the Chebyshev polynomial of order T m then, in that case the error is going to be maximum.

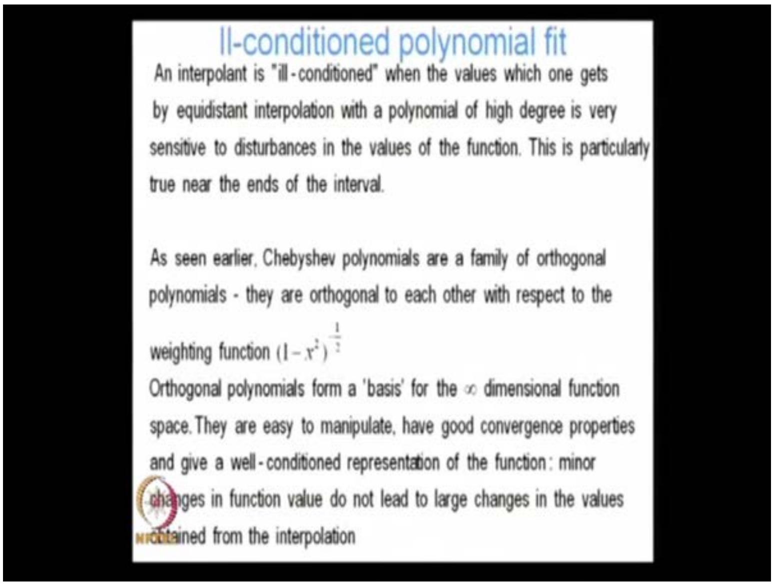
(Refer Slide Time: 12:25)



So, does this mean that we should use Chebyshev interpolation instead of equidistant interpolation all the time, well may be not because finding the grid points of Chebyshev interpolation involves evaluating trigonometric functions and evaluation of trigonometric functions is you can see it we need to evaluate this cos terms.

In order to find the 0’s we need to evaluate these terms. Evaluation of trigonometric functions is computationally expensive doing, so repeatedly for large problems may impose a heavy computational burden. Thus, using Chebyshev interpolation is probably best, for a higher order interpolation.

(Refer Slide Time: 13:14)



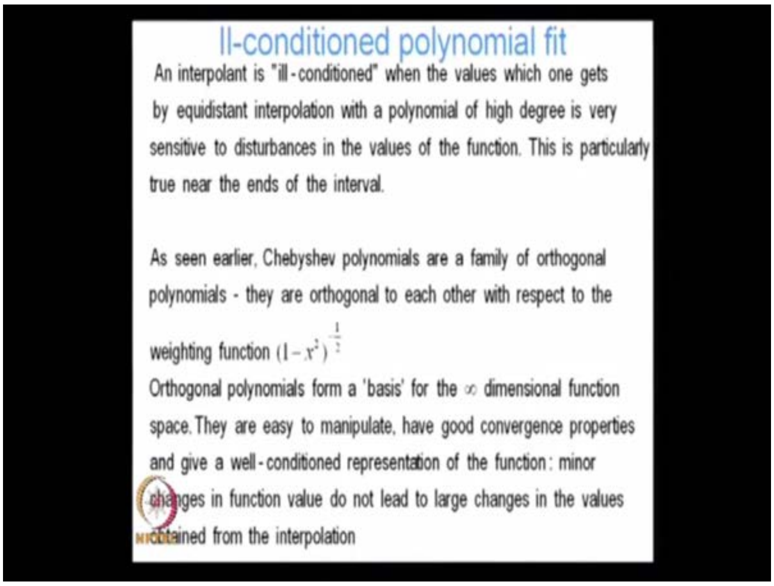
Particularly, because higher order interpolations has these problems, higher order interpolations have this problem so for these higher order interpolants it is better to use Chebyshev interpolation to use your grid points at the 0s of the Chebyshev function instead of the order of the remainder.

But, in case you using lower order interpolate then, maybe you can get away with using equidistant interpolation, maybe it is not that bad, if you use equidistant interpolation because well you are going to get larger errors that is true your truncation error is going to be larger but always in numerical analysis it is all a question of weighing the costs and the benefits. So, how much computational expense additional computational expense what is the gain in accuracy those things become important and people have done it people have done a lot of simulations they have reached the conclusion that well you should probably use Chebyshev interpolation for higher order polynomials because then the error just becomes too big and it is worth paying the price of calculating those grid points, calculating the grid points using for the Chebyshev polynomial.

But, for lower order interpolate may be you can get away with using equidistant interpolation, then using Chebyshev interpolation is probably best for higher order interpolations that is n greater than 2 to the power root m, I have m plus 1 grid points but I have n plus 1 grid points the highest order polynomial I can fit to that is m, I can fit in m’th order polynomial but instead of fit fitting an m’th order polynomial I am going to fit a polynomial of a much lower order and what is that much lower order n and if n less than 2 to the power root m, then I am fine doing equidistant interpolation.

But, if n is greater than 2 to the power root m, then I should go for Chebyshev interpolation for n greater than 2 to the power root m. Well, this is exactly why we have that behavior. It is because for n greater than 2 to the power root m equidistant interpolation using polynomials have a high degree have ill-condition what does ill-conditioning mean well, it means that if you have a slight change in the function value your interpolate is going to change a lot it is going to change a lot.

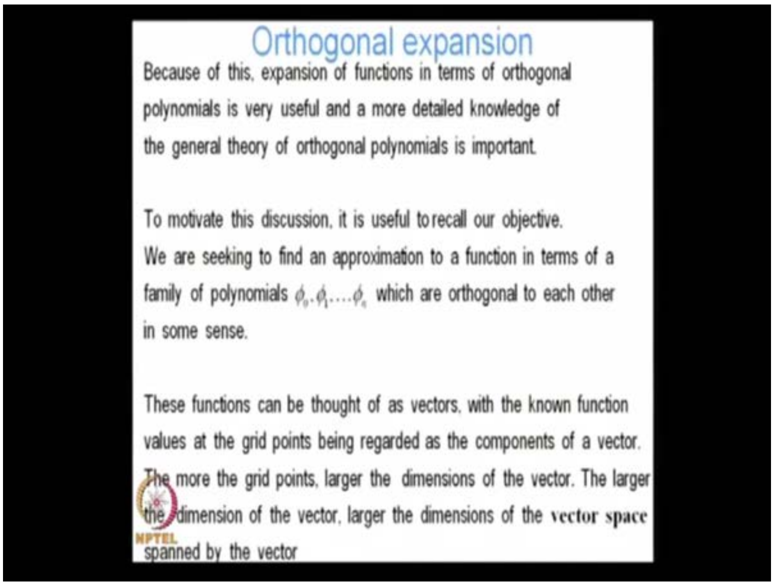
(Refer Slide Time: 16:00)



An interpolate is ill-conditioned when the values which 1 gets by equidistant interpolation with a polynomial of high degree is very sensitive to disturbances in the values of the function and this is particularly true near the ends of the interval. Chebyshev, well, that is enough about Chebyshev polynomials in particular. Now, I want to take a step back and talk about orthogonal polynomials in general, Chebyshev polynomials are just a particular instant or instants of orthogonal polynomials.

So, they are orthogonal to each other with respect to the weighting function 1 minus x square to the power minus half, why do we want to talk about orthogonal polynomials well, because orthogonal polynomials form a basis for the infinite dimensional function space. If I have an infinite number of orthogonal polynomials, the series of orthogonal polynomials infinite in number, they will form a basis for the infinite-dimensional function space. In addition, they are easy to manipulate, have good convergence properties and give a well condition representation of the function. Well, condition means that minor changes in function values will not lead to major changes in the values of the interpolates.

(Refer Slide Time: 17:25)

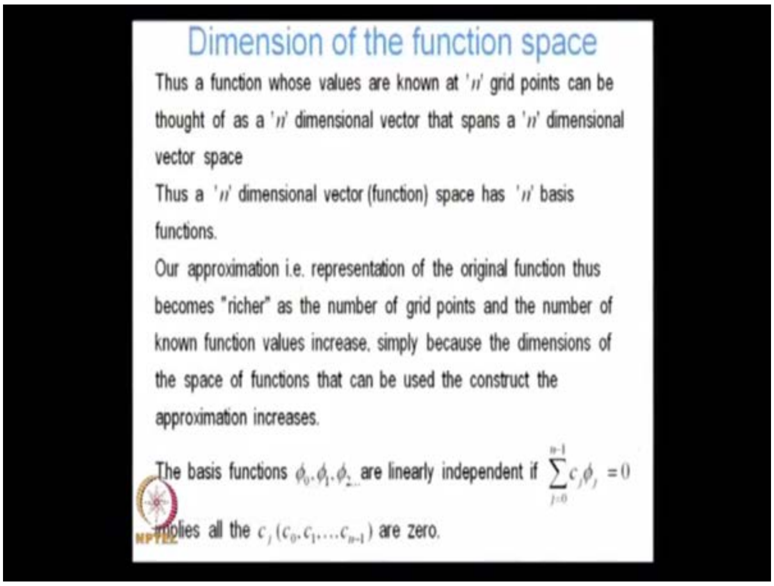


Because of this expansion of functions in terms of orthogonal polynomials is very useful and we want to talk about it in general, to motivate this discussion we are just to recall what we want to do.

We are trying to find an approximation to a function in terms of a family of polynomials of increase order which is orthogonal to each other in some sense they are orthogonal to each other. If we think of these functions as vectors just this is a mental exercise, if I think of those functions as vectors and if I know the, known function values at the grid points if I think of that as a vector the more the grid points larger the dimension of the vector that is clear .

So, I have grid points located and I evaluate the function values at those grid points, the larger, the more the number of grid points the larger the dimension of that vector. So if I have a function. I evaluate it at each of those grid points, I can write it as a vector and the more the grid points the longer that vector. So, the dimension of the space, the dimension of the vector larger the dimension of the vector space spanned by the vector. So larger the dimension of the vector space spanned, if I have 15 points, 15 grid points what is the dimension of the vector space it is going to be 15, so that, means that there must be 15 basis functions for that space there must be 15 independent vectors for that space any vector of size 15 can be written in as a linear combination of those 15 independent base vectors.

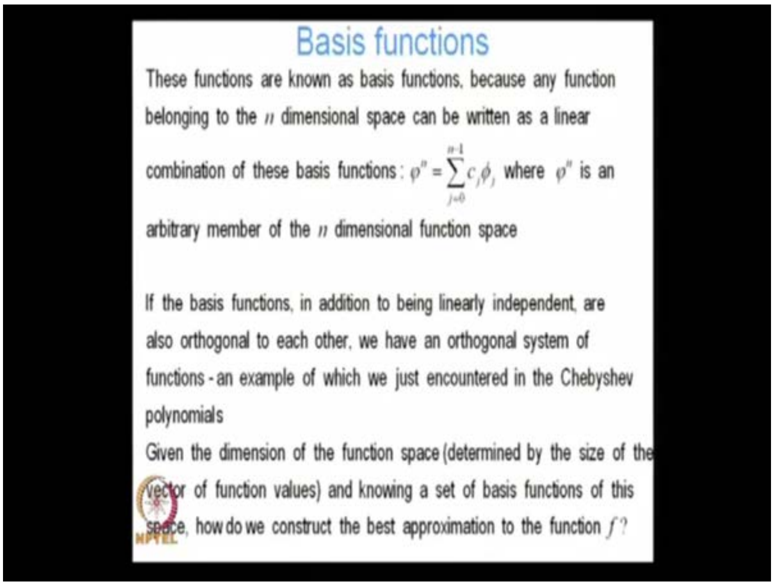
(Refer Slide Time: 19:24)



Thus, a function whose values are known at n grid points can be thought of as n dimensional vector that spans a n dimensional vector space. And then we go back to function. Thus, n dimensional function space has n basis functions. Here, instead of thinking of vector is just for making things clear because it is something that probably, most if we are familiar with but actually we are talking about functions and we are talking about the basis how many basis functions we need and it turns out that the number of basis functions, we need depends on the number of grid points number of grid points.

Our approximation that is representation of the original function thus becomes richer as the number of grid points and the number of known function values increases simply because the dimensions of the space of functions that can be used to construct the approximation increases. Now, as the dimension increases the number of basis functions also increase, it becomes in a sense richer my basis becomes more it can represent more complexity the basic functions phi 0 phi 1 phi 2 are linearly independent. If they satisfy this condition only if only, if all of the coefficients are 0, these functions phi j are linearly independent if I multiply all those phi j's with some arbitrary constants and if in case, they turn out to be 0 that means each of those constants have to be 0 what does it basically, it means that I cannot express any 1 of those phi j's as a linear combination of the remaining phi j's.

(Refer Slide Time: 21:25)

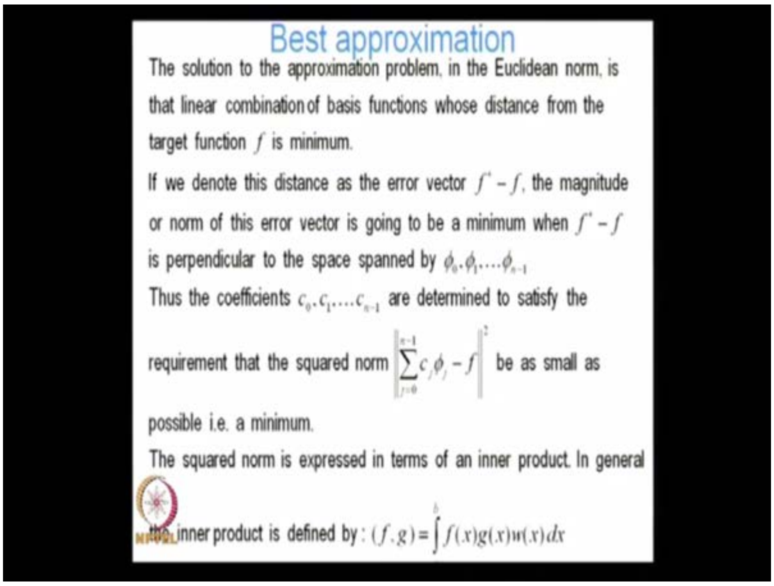


These functions are known as basis functions. Because, any function belonging to the n dimensional space can be written as a linear combination of these basis functions, if phi n is any arbitrary member of that n dimensional function space. I can write phi n as a linear combination of my basis functions for that space if the basis functions .In addition, to being linearly independent is also orthogonal to each other. We have an orthogonal system of functions like the Chebyshev polynomials they represented orthogonal series of functions orthogonal system of functions.

So, given the dimension of the function space determined by the size of the vector of function values. So, my vector of function values has a certain size evaluate the function at all the grid points. I get a vector, that vector has a certain size that gives me the dimension of the function space, how do we construct and knowing a set of basis functions, how do we construct the best approximation to the function f I know, I have a certain function f I know the size of my function space. Because, that function is given only at certain points at a finite number of points I have a finite dimensional function space and given and I know the basis functions for that function space.

So, how I am going to get the best approximation to f from by taking a linear combination of those basis functions.

(Refer Slide Time: 23:14)



The solution to the approximation problem in the Euclidean norm is that linear combination of basis functions whose distance from the target function f is minimum, So, you can imagine, I have these basis functions I construct a linear combination, I construct a function by taking a linear combination of those basis functions and that function that I construct is going to be the best possible approximation. If it is distance from the known function f is minimum and what do I mean distance, well I mean distance in the sense of Euclidean distance and we are going to talk about that in more detail.

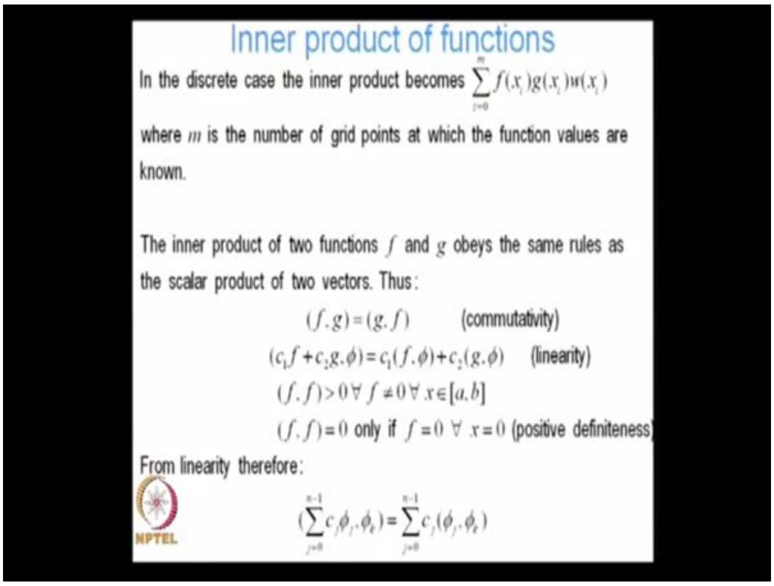
So, if we denote this distance as the error vector f star minus f the magnitude or norm of this vector is going to be a minimum, when f star minus f is perpendicular to the space spanned by phi 0 phi 1 phi n minus 1, think of very simple geometry. I have a straight line, I want to find a point not lying on the straight line, whose distance is minimum and I, it has on a certain path. So that, will be, it has to be perpendicular to the straight line. It has to be perpendicular, if I draw a line from that point to the straight line, if that line is perpendicular then the distance is minimum.

So, similar exactly we have something like this so f star minus f star is what I construct by taking a linear combination of my basis functions. If the distance, if the difference between f star and f star minus f is orthogonal is perpendicular to the space spanned by these functions, then I know that the magnitude of this error vector is going to be a minimum. Thus, the coefficients c 0 c 1 through c n minus 1 are determined to satisfy the requirement that, this Euclidean norm sigma c j phi j which, I construct like f star sigma c j phi c j phi j minus f.

If I take the norm of that square that is got to be the minimum that has got to be the small have to the smallest possible value and whatever value or whatever coefficient c j make this norm the smallest possible make this norm. Assume, this smallest possible value, those are the coefficients of my best approximation those coefficients multiplied by my basis functions are going to give me the best approximation.

Now, the squared norm, this norm is also expressed in terms of an inner product and we define inner an product like this is the notation first bracket 2 functions f g that means, I am taking the inner product of f and g and the inner product is always with respect to a certain weighting function and that weighting function is w of x.

(Refer Slide Time: 26:40)

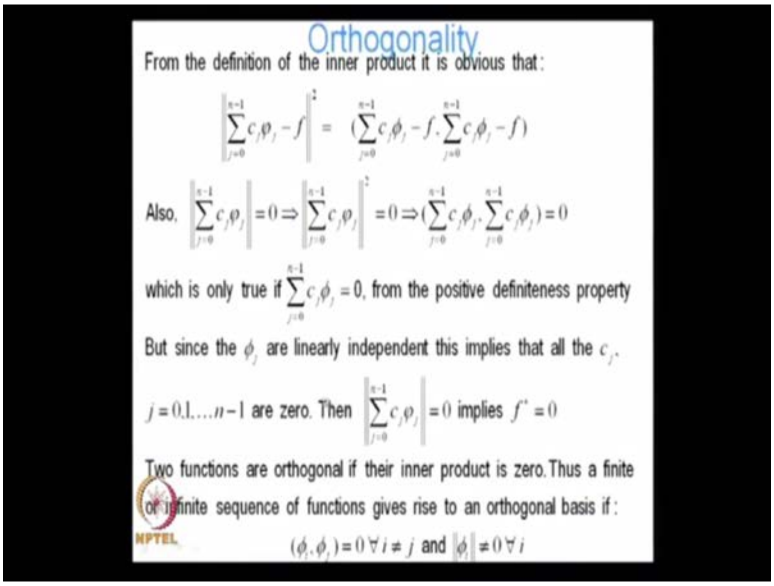


What that is in the continuous case? In the continuous case you can write the inner product in terms of an integral, what about the discrete case, where I just know the function values at a certain number of grid points. Well in that case you construct the inner product like this sigma i is equal to 0 f of x i g of x i w x i, I know the function values at points x i evaluate the function f at x, I evaluate the function g at x i multiplied by the value of the weight function at x i sum them together that gives my inner product of the discrete case.

And the inner product of 2 functions obeys the same rules as your vector inner product vector, dot product, what are those rules, well, f g is equal to g f that is commutative linearity c 1 f plus c2 g operating on phi inner product with phi, sorry not operating on phi inner product with phi that is going to give me c 1 f phi inner product of f and phi plus c 2 g inner product g of phi. That is linearity inner product of f and f is always greater than 0 for any f not equal to 0, when I say a function is not equal to 0 I mean the function is never 0 at any point in it is domain and it is domain being the interval on which it is defined. So, for all x belonging to a b f is not equal to 0 only then do I say that f is not 0.

So, f dot f comma f is greater than 0 for all f not equal to 0 for all x and f comma f is equal to 0 only if f is equal to 0 at every point x this is very important we are going to use that by document so f comma f is equal to 0 only if f is equal to 0 at all points x in the interval from linearity therefore, if I take this inner product the inner product of c j phi j is equal to 0 to n minus 1 with phi k, I can write it like this, I can bring out my constants outside and calculate the inner product of each of those functions with phi k each of those functions with phi k first calculate the inner product scale it by my coefficient and add it together so direct consequence of linearity.

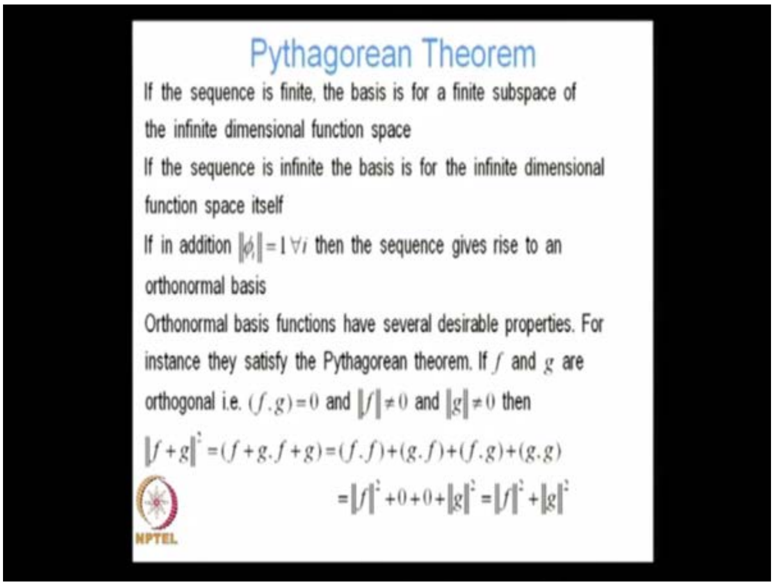
(Refer Slide Time: 29:30)



From the definition of the inner product it is obvious that this thing sigma c j phi j minus f norm square I can write it like that it is just the inner product is just the norm square sigma c j phi j minus f sigma c j phi j minus f also sigma c j phi j norm of norm of sigma c j phi j from j is equal to 0 to n minus 1 if this norm is equal to 0 this mean that the square of the norm must be is equal to 0 if the square of the norm is equal to 0 then i can write the inner product is equal to 0. But, we just saw that if the inner product is equal to 0.

If the inner product is equal to 0 then, the function itself must be equal to 0. So, that means if the inner product is equal to 0 sigma c j phi j j equal to 0 to n minus 1 must be equal to 0. But, since the phi j are linearly independent this implies that all the c j equal to 0 through 1 through n minus 1 must be equal to 0 then what does this mean, this means that if this is equal to 0 if norm of sigma j equal to 0 n minus 1 c j phi j equal to 0 then all the c j's must be equal to 0 and phi star f star is also going to be 0 because f star is constructed by from by this sigma c j phi j, 2 functions are orthogonal if the inner product is 0 that is a finite or infinite sequence of functions give rise to an orthogonal basis if phi i phi j is equal to 0 for all i not equal to j and norm of phi i not equal to 0.

(Refer Slide Time: 31:36)

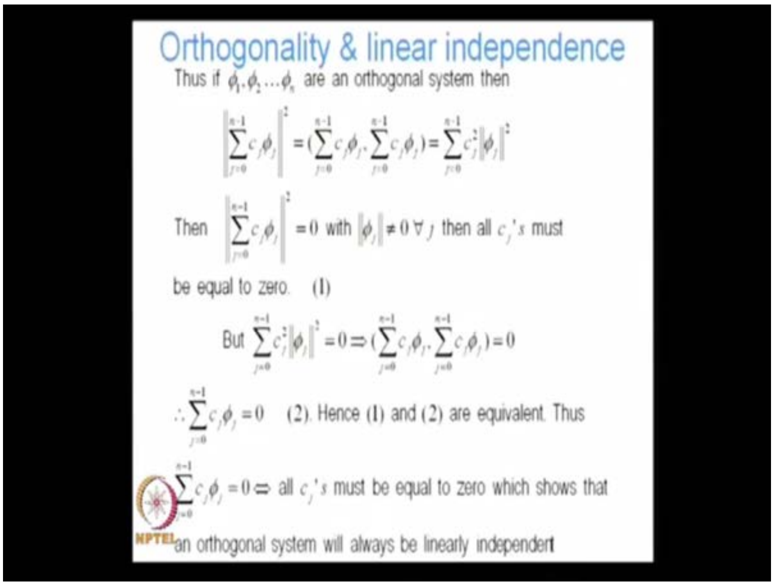


For all i we have something call the Pythagorean theorem if the sequence is finite the basis is for so will the basis is for a finite subspace of the infinite dimensional function space what I am trying to say is that if my function space is infinite dimensional my function basis is going to be infinite dimensional if my functions basis its sort of a trivial but it is worth saying if so incase I have a finite dimensional function space that is really a subspace of my infinite dimensional function space and the basis of that of that finite dimensional function space or a subspace of the basis functions of my infinite dimensional function space.

So, in addition to all these , we talked about Orthogonality that in addition to these if my norm of these of my functions is equal to 1 if the norm of my functions is equal to 1 then I have a orthonormal basis, I have an orthonormal basis ,orthonormal functions have actually it is not really orthonormal, its orthogonal they have very several desirable properties for instance they satisfy Pythagorean theorem, what does it say, it says that f and g or orthogonal that is f comma g inner product of f and g or 0 and norm of f is not equal to 0 and norm of g is not equal to 0 then norm of f plus g square, I can it in terms of an inner product f plus g f plus g which I can write as f comma f again using the linearity I can write it as f comma f plus g comma f plus f comma g plus g comma g.

And since, they are orthogonal f comma g and g comma f are going to be 0 and I am going to be left with norm of f square plus norm of g square, if I have orthogonal functions they satisfy this relationship norm of f plus g square is equal to norm of f square plus norm of g square. It is very similar to your trigonometry.

(Refer Slide Time: 34:13)



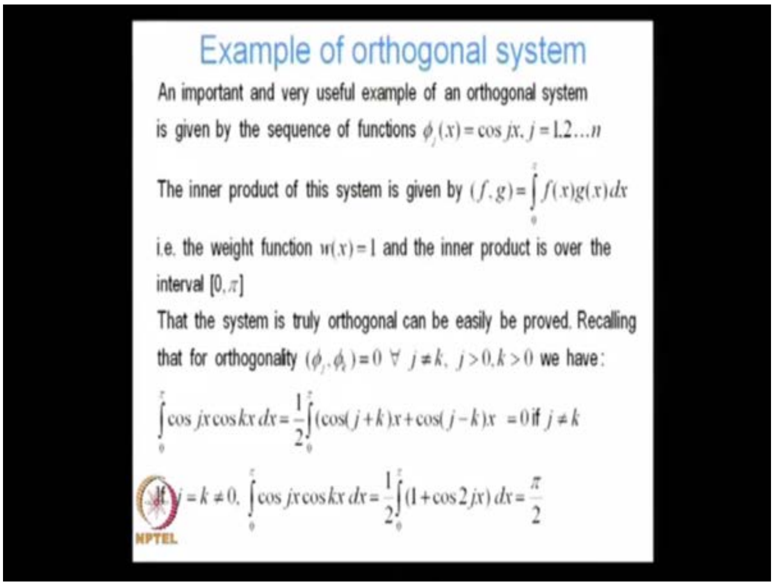
Thus, if phi 1 phi 2 through phi n are an orthogonal system then norm of sigma j equal to 0 to n minus 1 c j phi j square I can always like this in terms of an inner product and if I perform this inner product I am going to get sigma j equal to 0 to n minus 1 c j square norm of phi j square why because phi if the inner product will involve i again use the linearity and the inner product will involve phi c j c k instead of writing instead of writing it as j may be we would have been better if, I wrote it as k so c j c k phi j phi k some of those angle terms where j is not equal to k those inner products are going to vanish because of orthogonality.

I will be only be left with those inner products where the 2 functions are identical and that is going to give me norm of phi j square, I am going to get something like this then if this thing is equal to 0 if the norm of sigma j equal to 0 to n minus 1 c j phi j square is equal to 0 with norm of phi j not equal to 0 for all j then all the c j's must be equal to 0 you can see that why because i have these phi c 1 square, phi 1 square plus c 2 square phi 2 square plus c 3 square phi 3 square up to c n square phi n square all the phi n squares I know or not 0 with norm of phi j not equal to 0 if norm of phi j is not equal to 0 norm of phi j square cannot be 0 so I have this some of squares some of squares which are equal to 0 so that means each of those terms have to be equal to 0.

So, each c j c 1 c 1 c 1 square phi 1 square must be equal to 0 c 2 square phi 2 square must be equal to 0 and so on and so forth. And I know that phi 1 square norm of phi 1 square is not 0 so that means c 1 square must be equal to 0 c 1 square equal to 0 means c 1 must be equal to 0. So, if this is this satisfy if this is satisfied along with this condition then all the c j's must be equal to 0 that is 1 condition.

But, sigma c j square norm of phi j square equal to 0 implies again I am writing this in terms of an inner product sigma c j phi j sigma c j phi j equal to 0 so inner product equal to 0 that that means that this must be equal to 0 so that means sigma j equal 0 to n minus 1 c j phi j is equal to 0 hence what does it mean that means this is equivalent to this because this condition i got from this assumption and this condition I, I have also got from this assumption so hence 1 and 2 are equivalent thus sigma j equal to 0 n minus 1 c j phi j equal to 0 implies that all the c j's must be equal to 0 which shows that an orthogonal system will always be linearly independent an orthogonal system will always be linearly independent is the is the converse true well as certainly not every linearly independent system will not be orthogonal there is nothing requiring that but every orthogonal system has to be linearly independent that is should be sort of obvious but this is a nice little proof.

(Refer Slide Time: 38:14)



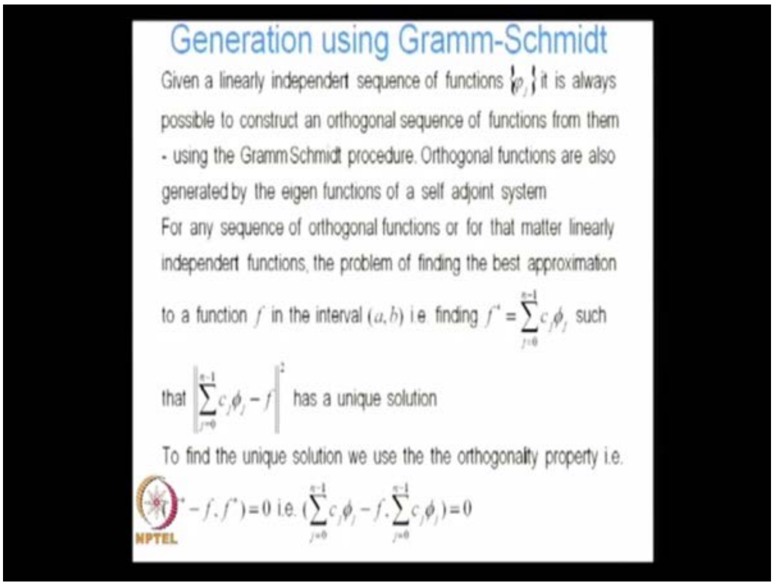
So, an important and very useful example of an orthogonal system is given by the sequence of functions cos j x or for that matter sin j x is of your Fourier series. So, cos j x you know that all of the all these cos j cos cos x cos 2 x cos 2 x they all orthogonal to each other well why is that well the inner product of the system is given by defined like this f of x g of x integral between 0 to pi that is the weight function is equal to 1 and the inner product is over this interval that the system is truly orthogonal can be easily proved with sort of trivial but let us do it quickly. So, integrals I am saying that each of these cos j x cos j x cos k x j not equal to k if I integrate that that if I take the product integrated between 0 to pi I am going to be get 0.

So, integral 0 to pi cos j x cos k x if I write it in terms of a sum of trigonometric functions I can write it like that half of this and this is going to give me sin j plus k this is going to give me sin j minus k if I integrate sin between 0 to pi I am going to get 0. So, this they are indeed orthogonal but if j is equal to k and both are not equal to 0 then cos j x cos k x nothing but cos square j x. So, I can as 1 plus cos t 2 j x this term is going to give me 0 on integration because again it is going to give me sin but this term is going to give me a non-zero contribution and that term is going to come out as phi by 2 so again so this is orthogonal any j not equal to any k not equal to j I am going to get 0 any 2 and j if i take the inner product of cos j with itself and I am not going to get 0.

So, suppose I have a linear earlier on in this lecture may be in the first may be sometime in the first quarter of, we talked about gram-Schmidt orthogonalization in the context of gauss elimination. So but particularly in the context of the q i method we talked about gram-Schmidt orthogonalization but actually the gram-Schmidt orthogonalization it is important for vectors the same thing it is very important for functions also why so given a set of functions which I know are linearly independent but not orthogonal they are linearly independent but not orthogonal, I can always construct an orthogonal basis from those linearly independent functions using gramm-schmidt orthogonalization using exactly the way I constructed given n linearly independent vectors I constructed a basis for the n an orthogonal basis for the n dimensional vector space using gramm-schmidt orthogonalization I can do exactly the same thing for functions the idea is exactly the same so that so given linearly independent functions I constraint construct an orthogonal series of functions when else do we get orthogonal functions we looked at that earlier too when we solve the eigenvalue problem .

When we call when we solve this when we had a self -ad joint linear system and when we tried to solve the Eigen value problem for that system we found that the Eigen functions the Eigen functions were orthogonal so again the eigenfunctions arising from the eigenvalue problem of a self – ad joint linear system, self – ad joint system they give me orthogonal functions they give me orthogonal functions.

(Refer Slide Time: 42:26)

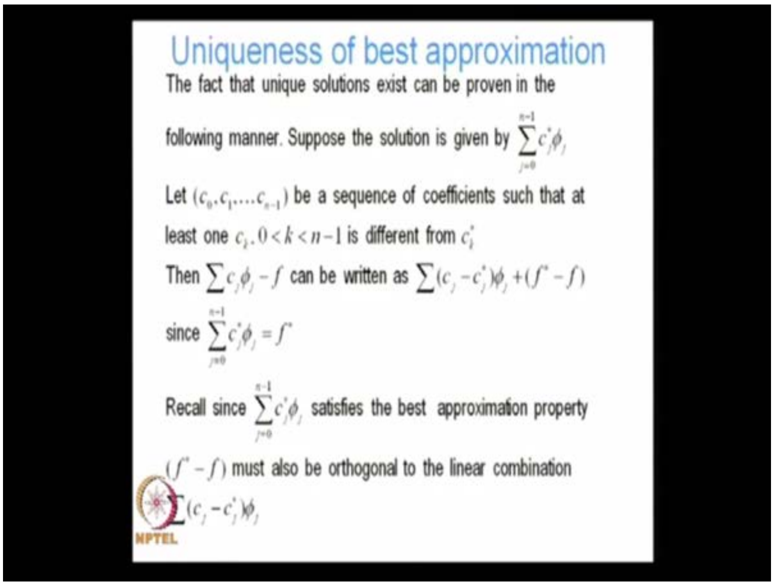


So, these are 2 ways we can generate orthogonal functions. So, for any sequence of orthogonal functions are for that Matrilineal independent functions the problem of finding the best approximation to the function f in the interval a b that is finding f star is equal to sigma j equal to 0 to n minus 1 c j phi j such that norm of sigma j equal to 0 to n minus 1 c j phi j minus f square has a unique it has a unique solution. So, this f star this f star which gives me the best approximation this f star is unique. So, given an orthogonal basis and given a certain function the best approximation the best guess that i can construct 2 f using those basis functions is unique there is only 1 combination of coefficients that is c 0, c 1, c 2, c 3 up to c n then, assume only certain unique value which is going to give me the best approximation there is no more than 1 c 0 c 1 c 2 have unique values they cannot have more than 1 value for that best approximation.

To find the unique solution we use the orthogonality property what is the orthogonality we just talked about that that is f star minus f comma f star must be equal to 0 f star why do I say that well because f star minus f is my error f star minus f is it was my error and I know that f star minus f must be orthogonal to the basis spanned by the basis functions by phi j but f star f star is a linear combination of those phi j's f star is a linear combination of those phi j's therefore, f star minus f if I take the inner product with f star itself that is got to be equal to 0 because I can write that f star as c 1 phi 1 plus c 2 phi 2 plus c 3 phi 3 up to c n phi n then i break it up so f star minus f comma c 1 phi 1 plus f star minus f comma c 2 phi 2 plus f star minus f comma c n phi n so f star minus f comma c 1 phi 1 has to be equal to 0 plus because f star minus f is orthogonal to the space spanned by all the phi's so it must be orthogonal to each phi as well .

So, f star minus f comma c 1 phi 1 is equal to 0 f star minus f comma c 2 phi 2 is equal and so on and so forth, f star minus f comma f star has also got to be equal to 0 and then I write, I have just written that f star I have written as sigma j equal to 0 to n minus 1 c j phi j minus f comma sigma j equal to 0 n minus 1 c j phi j that is equal to 0.

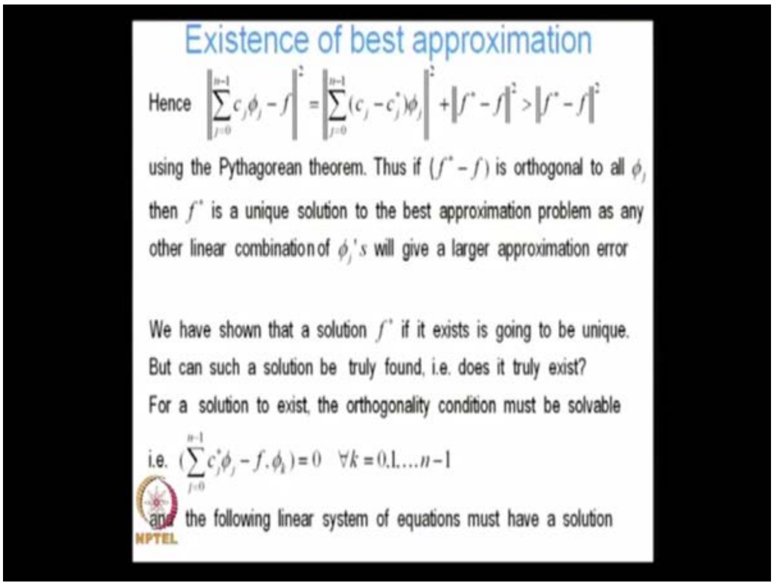
(Refer Slide Time: 45:51)



And the fact that unique solutions exist can be proven in the following manner suppose the solution the unique solution is given by this that is suppose that unique solution is given by this c j star comma phi j and let us suppose c 0 c 1 c n minus 1 where sequence of coefficients such that at least 1 of those coefficients is not equal to c j the star values.

So, suppose c k is not equal to c k star then sigma c j phi j minus f I can write it as sigma c j minus c j star phi j plus f star minus f c j phi j minus f c j phi j minus c j star phi j what is c j star phi j c j star is my best approximations that is f star. So, I have minus f star plus f star so that cancels out and then, I have minus f. So, I can write this as sigma c j minus c j star phi j plus f star minus f since this is equal to f star. Recall since sigma j equal to 0 to n minus 1 c j star phi j satisfies the best approximation property since it satisfies the best approximation property f star minus f must also be orthogonal to the linear combination c j minus c j star ph j. I just talked about that f star minus f must also be orthogonal to this because this is after all a linear combination of the phi j's. So, if take f star minus f and take the inner product with this thing c j minus c j star phi j I am basically taking the inner product with each of the phi j's.

(Refer Slide Time: 47:54)

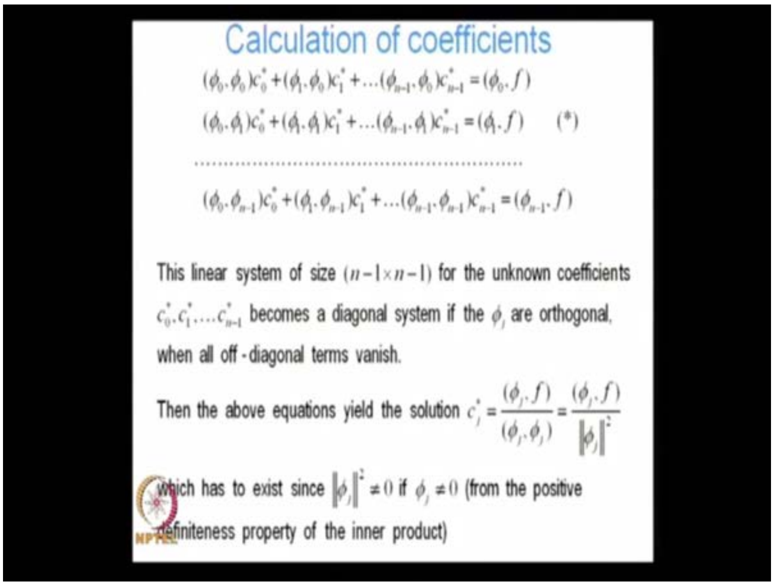


So, what do we have? We if take norm of sigma c j f i j minus f and square it I can write it as sigma j equal to 1 j equal to 0 to n minus 1 c j minus c j star phi j square plus f star minus norm of f square why can I do that because these are orthogonal. Because I just said that this must be orthogonal to that and I express this as the sum of this and the sum of that and each of this is orthogonal then I use my Pythagorean theorem and say that this norm of this must be equal to the norm of this square norm of this square must be equal to the norm of this square plus the norm of that square simply because these are orthogonal.

So, what does that mean? So, this square plus that square which must be greater than f star minus norm of f square because what is this is norm of f star minus f square plus something that something is either has got to be 0 or greater than 0. So, it is a square. So, this thing has to greater than f star minus f square. So, what does that mean? That means that this thing, this c j phi j minus f were remember this c j are not all the optimal coefficients they are not the best approximation coefficients. So, this the difference of this from f in the Euclidean norm is greater than the difference of f star from f which obviously means that f star is my best approximation thus if f star minus f is orthogonal to phi j which was the condition we started with then f star is a unique solution to the best approximation problem as any other linear combination of phi j's will give a larger approximation error any other linear combination is going to give me a larger approximation.

So, we have shown that a solution f star if it exists is going to be unique the coefficients which define f star are unique but can such a solution be truly found that that is does it truly exist. So, we always concern ourselves with in mathematics at least they concern on themselves with existence and uniqueness no matter does not matter, if something does not exist I mean it can proved to be unique but just because it is unique does not mean that it does not exist that, it exists. So, you have to prove uniqueness as well as existence but we are not but we just do it very simply. So, for a solution to exist the orthogonal orthogonality condition must be solved that the what so there must be some c j star phi j such that this condition is satisfied such that c j star phi j minus f if i take the inner product with any of the phi k's that is got to be equal to 0. So, this condition has to be satisfied so this gives rise to a linear system of equations.

(Refer Slide Time: 51:25)



What is that linear system of equations it is going to be something like this. Just by expanding this out c j star phi j minus f comma phi k equal to 0. Expanding that out, phi 0 comma phi 0 c c 0 star phi 1 comma phi 0 c 1 star equal to phi 0 comma f i expand this out i get a system of equations and you can see this is a linear system of size n minus 1 times n minus 1 and this linear system is for the for the coefficients c j star c 1 star through c n minus star 1 star and if I have an orthogonal basis you will see that this linear this coefficient matrix so basically i want to find c 0 star c 1 star c n star minus n. So, these are my unknowns and the coefficient matrix is phi 0 phi 0 term each term in the coefficient matrix is phi 0 phi 0 phi 1 phi 0 and so on and so forth. May if i have an orthogonal basis you will see it is obvious that only the diagonal terms are going to be non-zero each of the off-diagonal terms is going to be 0.

So that, becomes a diagonal system of the phi j are orthogonal when all the off-diagonal terms vanish then the above system yields the solution c j star is equal to phi j comma f divided by norm of phi j square which has to exist why because this thing cannot be equal to 0 because this thing because we have assumed that phi j is not equal to 0 so this thing has got to be different from 0 is got to be positive actually and if this thing is positive then each of these c j stars exist.

We, I also want you to so the I proved the existence for when my basis functions are orthogonal it can also be proved which is likely more involved but it can also be proved that when my basis functions are not really orthogonal but are linearly independent in that also I can in that case also I can prove the existence of my solutions. So, I start with that in the next lecture.

Thank you.