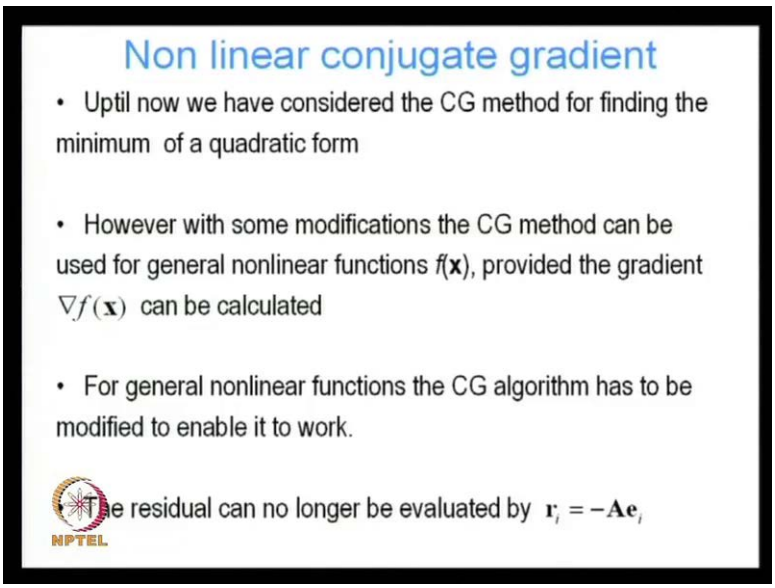
Numerical Methods in Civil Engineering
Prof. Dr. Arghya Deb
Department of Civil Engineering
Indian Institute of Technology, Kharagpur

Lecture - 20
Nonlinear Conjugate Gradient and Introduction to PDEs

Few of our series on numerical methods in civil engineering, we will complete our discussion of the conjugate gradient method by talking about non linear conjugate gradients and we will start our discussion on partial differential equations.

(Refer Slide Time: 00:36)

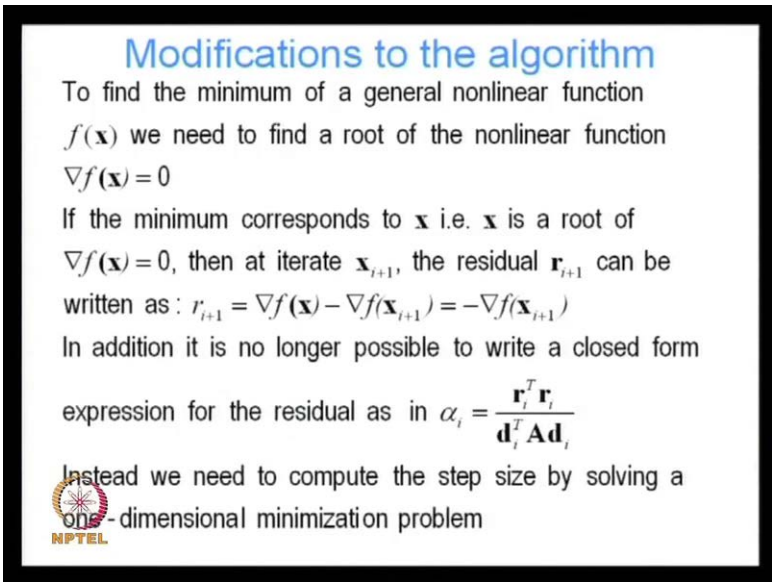


So, up till now we have considered the conjugate gradient method exclusively for finding the minimum of a quadratic form right. We have not looked at the general nonlinear problem; however, with some modifications the conjugate gradient method can be used for general non linear functions f of x basically to find the minimum of a general non linear function f of x provided the gradient can be calculated at every point in our domain of interest right.

So, provided we can calculate grad of f of x, we can still use the conjugate gradient method for general nonlinear problems with some changes. For general nonlinear functions, the conjugate gradient algorithm has to be modified to enable it to work what are the modifications necessary? We will first of all the residual can no longer be

evaluated by r i is equal to minus A i because there is no constant matrix A anymore right.

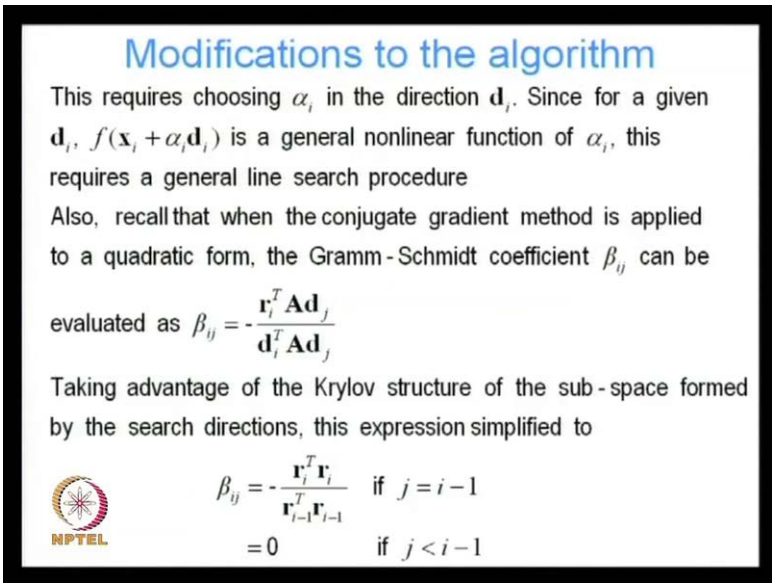
(Refer Slide Time: 01:40)



To find the minimum of a general nonlinear function f of x we need to find a root of the nonlinear function gradient of f of x is equal to 0. So, we want to find the value of x such that gradient of f of x is equal to 0. If the minimum corresponds to x that is x is a root of gradient of f of x is equal to 0. Then as we go towards that root at iterate x power x i plus 1 the residual r i plus 1 can be written as the gradient at the root at the exact root minus the gradient the current gradient at x i plus 1. And since we know that we add the exact root the gradient is got to be 0. So, gradient of f of x is going to be 0 and so, we get minus grad of f x i plus 1. So, that is going to be my residual.

So, that is one that is one thing. So, we cannot calculate the residual using our previous formula. So, we have to calculate the residual directly from the gradient of the function. In addition, it is no longer possible to write a closed form expression again I have a typo for the step size as in alpha i is equal to r i transpose r i divided by d i transpose A d i right. we cannot any more calculate a closed form expression for this step size.

(Refer Slide Time: 03:16)



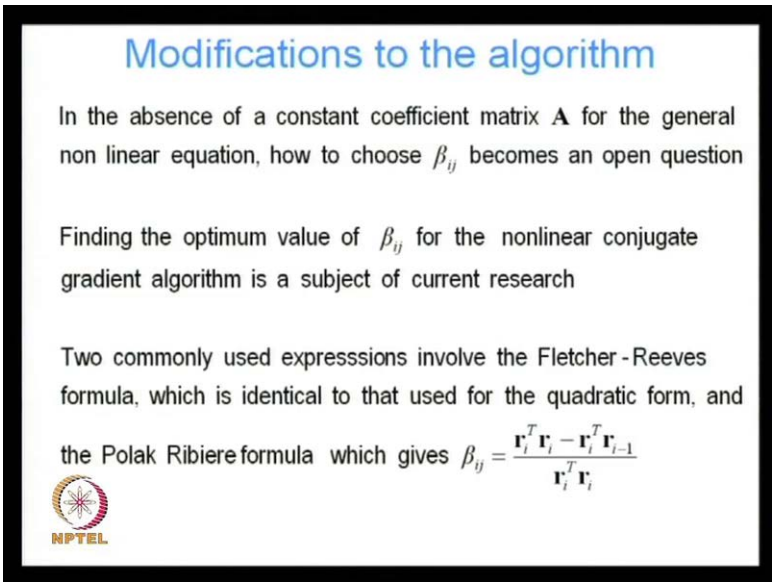
So, instead we compute the step size by solving a one dimensional minimization problem and what is that one dimensional minimization problem? Well it basically requires us to find the value of alpha in the search direction d, which is going to minimize my function f x i plus alpha i d i right, But this f x i alpha d i is fixed right, I am going along a particular search direction. So, d i is fixed for that iteration right.

So, I will that expression that function f of x i plus alpha i d i becomes a general polynomial in alpha right. It becomes a general polynomial in alpha hence we are required to find the value of alpha which minimizes that that function. Also, recall that when the generate when the conjugate gradient method is applied to a quadratic function that was the second change that we need to make right first of all the residual has to be calculated differently the step size has to be calculated differently; thirdly the gram-Schmidt coefficients are different why because when we use the conjugate gradient method for quadratic form the Gramm-Schmidt coefficient beta i j.

we evaluated it as beta i j is equal to minus r i transpose A d j divided by d i transpose A d j and we took advantage of the Krylov structure of the subspace formed by the search directions to simplify the expression for beta i j to get beta i j equal to minus r i transpose r i by r i minus 1 transpose r i minus 1 right and we found that only one beta i j is required right. Only one beta i j is required the others we do not need to enforce because

they are automatically orthogonal. The previous residuals are automatically, the previous search directions automatically orthogonal to the current residual.

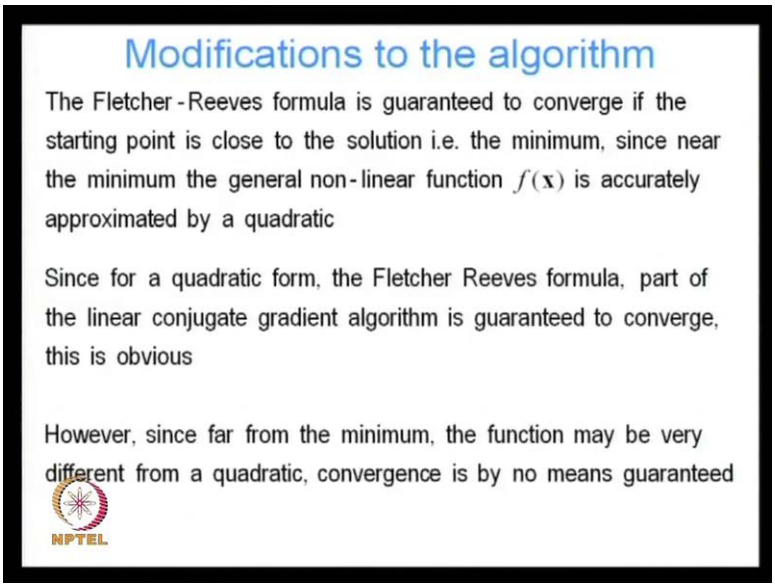
(Refer Slide Time: 05:18)



In the absence of a constant coefficient matrix A for the general nonlinear equation how to choose beta i j? Becomes an open question there is no definitive solution to that still people are doing I mean people are proposing different formulas for that right. one option is to just use the use the expression which we which we use for linear conjugate gradient right for conjugate gradient applied to the quadratic form that is we use this expression for beta i j right.

So, that is known as the Fletcher-Reeves algorithm the Fletcher-Reeves algorithm, which involve which is identical to that used for the quadratic form. And then there is the Polak Ribiere formula, which is a slight modification which gives beta i j is equal to r i transpose. If we compare these two you can see that it is slightly different this is mine this is this and that that right. So, it is slightly different right then there is other formula when since there is something called the Hestenes-Stiefel formulas Stiefel formula them there are other formulas also, but these are the main one.

(Refer Slide Time: 06:33)



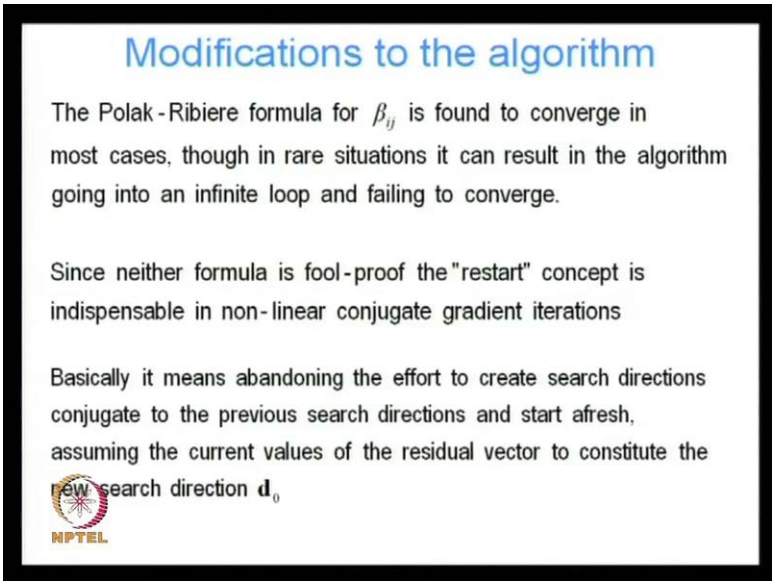
Now the Fletcher-Reeves algorithm when is the when is the Fletcher-Reeves formula going to work best well it is going to work best near the solution why is it going to work best near the solution? Because near the solution whatever be my nonlinear function, it is going to behave pretty much like a quadratic right. So, when it is when it becomes a quadratic then my linear conjugate gradient applies exactly right. So, near the solution that Fletcher-Reeves formula is going to work well because the function behaves quadratically right. Any function near it is root it is you can pretty much approximate it very well by quadratic.

So, the Fletcher-Reeves formula is guaranteed to converge if the starting point is close to the solution why because close to the solution it is going to be a quadratic the behavior is going to be more or less quadratic right. So, close to the solution the Fletcher-Reeves formula if you start with a if you start you are your start your iteration with a starting guess which is close to the minimum then the behavior of the function near the minimum is pretty much quadratic right. The Fletcher-Reeves formula is going to converge very well. But, since near the minimum the general nonlinear function f of x is accurately approximated by a quadratic since for a quadratic form the Fletcher-Reeves formula part of the linear conjugate gradient algorithm is guaranteed to converge this is obvious right.

However, far from the minimum the function may be very different from a quadratic. In that case convergence is by no means guaranteed, if we use the Fletcher-Reeves formula

for a general nonlinear function and we start far away from the true solution then there is no guarantee that we are going to converge because the behavior is no way near quadratic right.

(Refer Slide Time: 08:26)

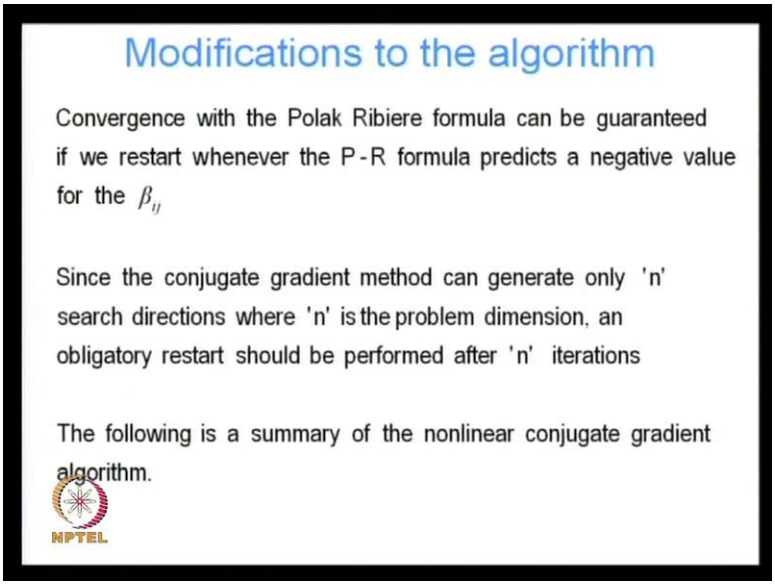


The Polak-Ribiere formula for beta i j is found to converge in most cases it is its converges in most cases irrespective of the starting point right. Though in rare situations it can result in the algorithm going into an infinite loop and failing to converge basically it starts cycling right it cannot converge right since neither formula is full proof the restart concept is indispensable in nonlinear conjugate gradient iterations. This is probably the most important concept at the restart concept when it comes to nonlinear conjugate gradient iterations.

Basically it means that the effort to create search directions conjugate to the previous search directions has to be abandoned at certain point. when we think that we are no longer converging right or we have exceeded n number of iterations right, we know recall that the conjugate gradient algorithm we it is it is an n dimensional space we construct n conjugate directions right that is there are only n conjugate directions right. So, suppose my iteration takes more than n it goes to n plus 1 right it still does not converge then; obviously, the new directions that I am using they are conjugate directions right those I have already exhausted the all the possible conjugate directions n conjugate directions have already been exhausted.

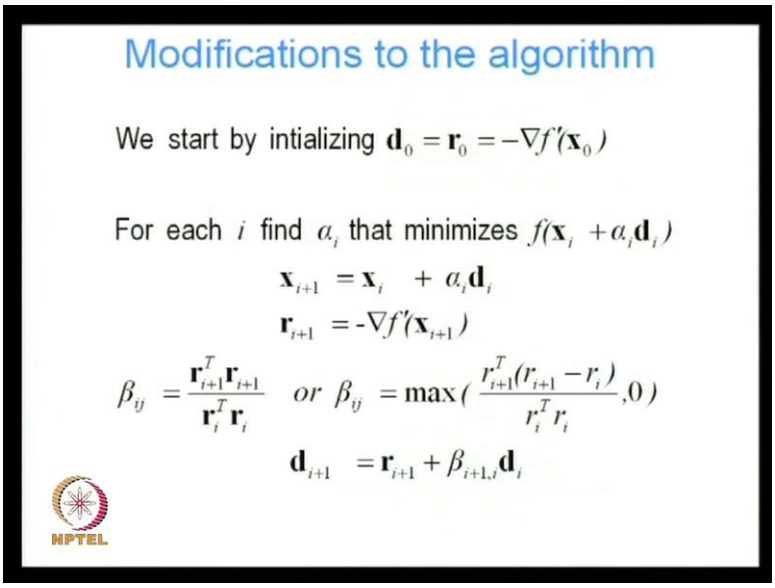
So, there is no point. So, I am no longer doing the conjugate gradient method. So, at that point I stop, I stop my iteration and restart my algorithm how do I restart? well I said the calculate the current value of the residual set it equal to the search direction like I did when I started my conjugate gradient and then I start the iteration again right. Basically, it means abandoning the effort to create search directions conjugate to the previous search directions and starting a fresh assuming the current values of the residual vector to constitute the new search direction d 0.

(Refer Slide Time: 10:28)



So, convergence with the Polak Ribiere formula can be guaranteed if we restart whenever the Polak Ribiere formula predicts a negative value for the beta i j. suppose, I got a Polak Ribiere beta i j which is negative. So, at that point I know that my Polak Ribiere formula is giving arbitrary result. So, I better abandon it and start again with the with the current residual as the first search direction. Since the conjugate gradient method can generate only n search directions where, n is the problem dimension an obligatory restart should be performed after n iteration right. The following is the summary of the nonlinear conjugate gradient algorithm.

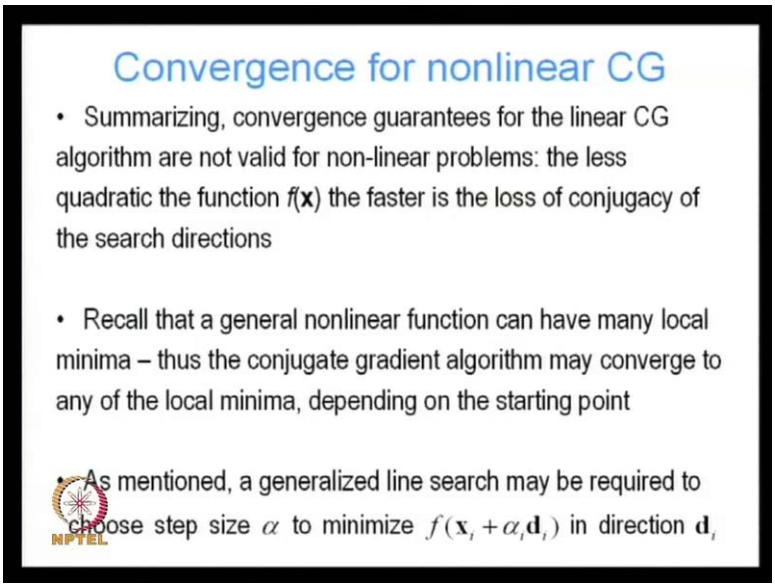
(Refer Slide Time: 11:16)



So, we start with initializing our initial search direction d 0 with the residual at r 0 right. Then for each I each iteration we find the alpha that minimizes f of x i plus alpha i d i at this stage I assume the already know d i right. So, I find the minimum of this function this function nonlinear function in alpha right, once I find the alpha i update the new iteration; the new iterate value x i plus 1 is equal to x i plus alpha i d i. So, once I know x i plus 1, I calculate the gradient now that is that is some strange notation right. So, I do not need that f prime here right I already have that gradient right. So, again I apologize for the typo.

So, I find gradient of f at x i plus 1. So, I calculate their residual at r i plus 1 once I calculate the residual I either use my Fletcher-Reeves formula or my Polak Ribiere formula to calculate the new beta i j. once, I get my new beta i j I found out my new search direction right. and, I make sure that any time I exceed n iterations I start again right from here right you can see that this Polak Ribiere beta i j, I have a max here this is to ensure that it is never negative right if it is negative then it is bad right. So, if it is if it is negative with 0 right that tells me that I better restart. Because, I am going to use I am not going to use any search direction to get a new search direction.

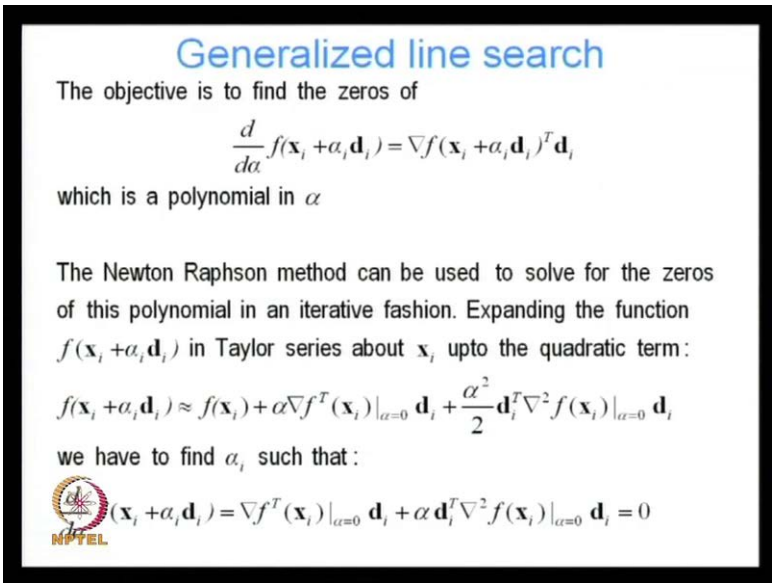
(Refer Slide Time: 13:02)



So, convergence for nonlinear conjugate gradient summarizing conjugate gradient nonlinear conjugate gradient we are guaranteed to there is no guarantee right then there is no all the all the convergence criteria we have obtained for the linear conjugate gradient that just goes out of the window right. There is no guarantee right the less quadratic the function f of x the faster is the loss of conjugacy of the search directions right. So, we are guaranteed to converge in n, n iterations if my search directions are conjugate if my search directions loose conjugacy no guarantees right and here there is no guarantee of conjugacy at all right.

Also, the recall this is a general problem for all general nonlinear equations, which is that all general nonlinear equations there is no unique minimum no global its very hard to reach a global minimum right. So, you get to the local minimum to which you start closest right. So, we can we will get we will converge if we converge we are going to converge to a local minima depending on our starting point and as mentioned the generalized line search may be required to choose the step size alpha to minimize f of x i plus alpha i d i in direction d i.

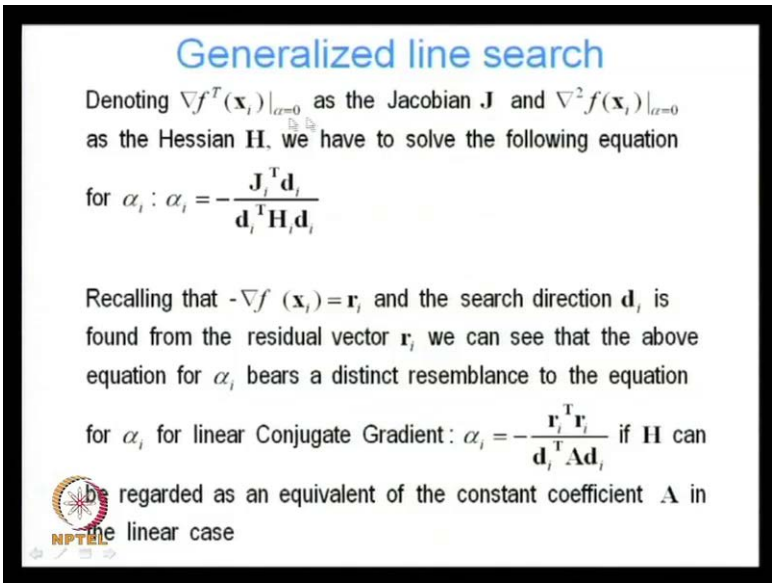
(Refer Slide Time: 14:25)



So, I talk little bit about the generalized line search problem the objective is to find the zeros of d d alpha f of x i alpha i d i. So, I want to minimize f of x x i plus alpha i d i. So, how do I minimize I said the derivative with respect to alpha to 0 right. So, that says the gradient you take the calculate the gradient and you take the projection of that in the d direction right. You take the projection of that in the d direction you get a polynomial n alpha right you get a polynomial n alpha and then you can use the Newton Raphson method. So, this is a polynomial a general polynomial in alpha you can solve it using the Newton Raphson method to solve for the zeros of this polynomial in iterative fashion.

So, how do you do that well you expand the function f of x i plus alpha i d i in Taylor series about x i up to the quadratic terms, you write f of x i plus alpha i d i approximately equal to f of x i plus alpha gradient of f transpose alpha d i gradient of f transpose x i plus alpha square by 2 d i transpose gradient of x i d i right. So, it is up to quadratic term and then we may have to find alpha i such that then why do we do that? I did that because this offers us some insight into the convergence of the conjugate gradient method let us follow through and we will see right. So, from this equation we set this to 0 we take the derivative with respect to alpha and we set this to 0 right. So, we get grad f transpose x i d i plus alpha d i transpose, this is the Hessian right the second derivative right Hessian d i equal to 0 right and. So, this is an equation which is which I can evaluate to obtain my alpha and what is that.

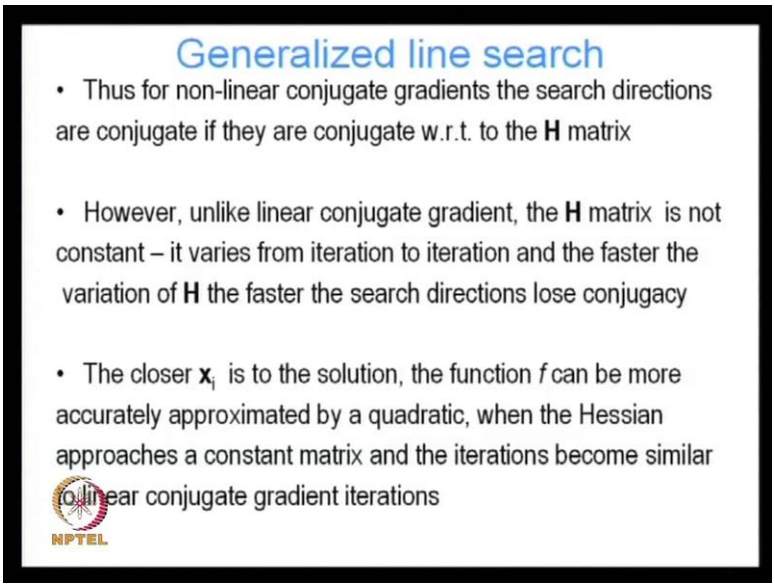
(Refer Slide Time: 16:25)



So, denoting grad f transpose x i at alpha equal to 0 as the Jacobian j and grad square f x i at alpha is equal to 0 as the Hessian H we have to solve for the following equation for alpha i which is this. Now, compare this equation compare this equation with the equation I had for the linear conjugate gradient and that is going to be very revealing or did I have that equation right. There, yes I had that equation right that this was my equation for the linear conjugate gradient. So, what is what do I have let us compare terms. So, this is j i transpose right, j i transpose I know is nothing, but this right it is the Jacobian, what is this is nothing but the residual right we saw that right gradient; gradient of f that is very similar to the residual right then I have d i.

Here, but I have r i here but we know that we get the search directions d i from the residual r i right. So, that is also it is a very similar look at the bottom we have d i transpose here then we have the constant matrix A and we have d i. So, instead of A the major difference is that instead of A I have the Hessian I have the Hessian here and. So, that is the if it is the only difference is that in case of linear conjugate gradient the Hessian is a constant matrix A. Here the Hessian is a variable matrix right it is a function of it is a function of x i right. So, that is the difference. So, if So, this is exactly what I said right the. So, we can see there is a distinct resemblance of this equation to the linear conjugate gradient, if H can be regarded as an equivalent of the constant coefficient A in the linear case right.

(Refer Slide Time: 18:32)



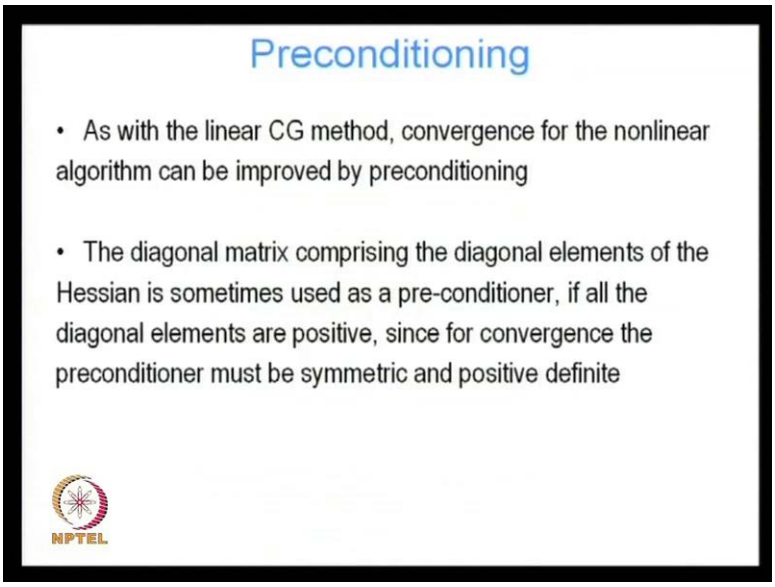
So, for nonlinear conjugate gradients the search directions are conjugate if they are conjugate with respect to the Hessian matrix because the Hessian here is plays exactly the same role as the A matrix. So, if the directions of conjugate with respect to the Hessian matrix right then the conjugate gradient method is going to do. Well, it is going to very close to its behavior in the linear and for the linear conjugate gradient situation right. Only problem is; however, unlike linear conjugate gradient the H matrix which in case of linear conjugate gradient is a constant matrix A is not constant it varies from iteration to iteration and the faster the variation of H the faster the search directions lose conjugacy.

The more constant H is the greater the chances of my retaining Conjugacy and this also tells me why closer to the solution I have I have better convergence. Because, closer to the solution my function is going to behave like a quadratic it is going to behave like a quadratic the Hessian got to be constant right. So, closer to the solution my Hessian is going to be constant.

So, there is a better chance of my conjugate gradient directions maintaining conjugacy, if they maintain conjugacy all the good things follow right all my convergence properties they follow and I can I am assure to converge right. So, the closer x i is to the solution the function f can be more accurately approximated by a quadratic, when the Hessian

approaches a constant matrix and the iterations become similar to linear conjugate gradient iterations right.

(Refer Slide Time: 20:27)



Let us talk little bit about preconditioning for the nonlinear conjugate gradient method, as with the linear conjugate gradient method convergence for the nonlinear algorithm can be improved by preconditioning right. We said that if we multiply, if you multiply by coefficient matrix with a certain operator with certain matrix or we operate on that coefficient matrix with a combination of matrices right and we get a coefficient matrix, which has a condition number; which is smaller than my original coefficient matrix a then I am going to get better convergence.

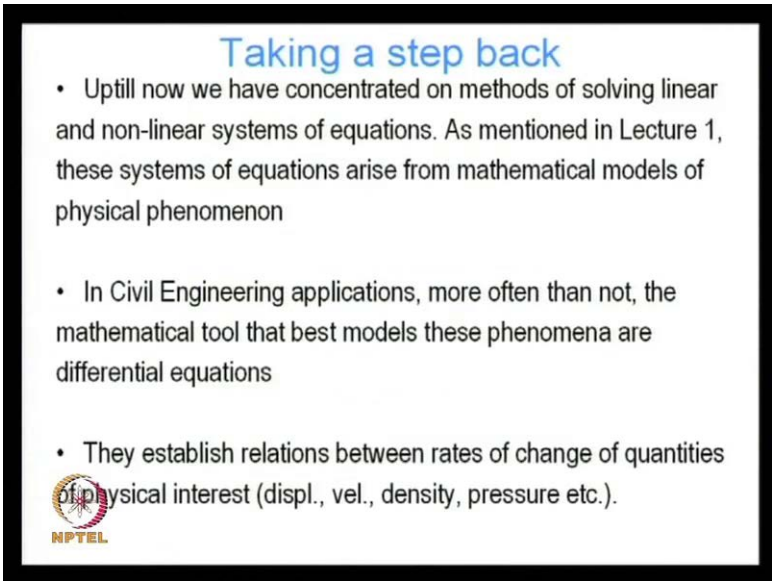
Because, it is the convergence of the conjugate gradient method depends on the how well conditioned my coefficient matrix is how close my Eigen values are to each other. So, best situation is when the Eigen values are like they are grouped together right when the Eigen values are grouped together that gives us the best convergence situations like that right.

So, this all an attempt to improve the condition number of my coefficient matrix. So, how do we use? How do we condition the coefficient matrix? in case of nonlinear conjugate gradient well sometimes what people do is to use the Hessian matrix right. To take the Hessian matrix throughout all the of diagonal terms and keep the diagonal terms if you keep the diagonal terms and if all the diagonal terms are positive then I must

assured that my precondition is symmetric and positive definite right. So, if the condition was the precondition has to be symmetric and positive definite. So, if I do that then I am assured that that is that is a valid preconditioning right. If we cannot do that well just do not do any preconditioning never try a precondition, where you have a negative element on the diagonal because that is the sure prescription for divergence right.

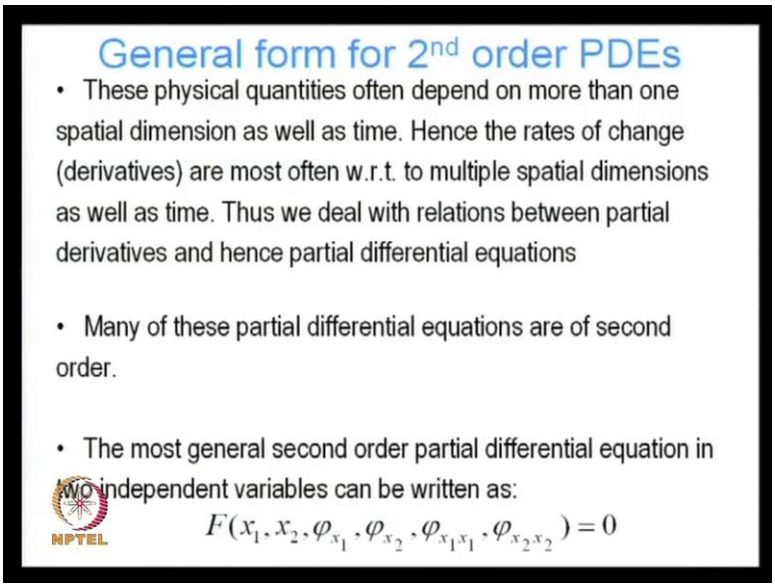
So, that is all I have to say about the conjugate gradient algorithm and that winds up our discussion on gradient based methods and it also winds up our discussion on methods for equations solving. So, that is all I am going to talk about equation solving for the rest of this course. So, the rest of this course we are going to focus on how do we get those equations, those equations that we get those matrix equations that we try to solve well how do we get them right.

(Refer Slide Time: 23:09)



So, up till now we have concentrated on methods of solving linear and nonlinear systems of equations. As mentioned in lecture, one these systems of equations arise from mathematical models of physical phenomena. In civil engineering applications more often than not the mathematical tool that best models these phenomenon are differential equations right. So, they establish relations between rates of change of quantities of physical interest for instance displacement velocity density pressure and things like that how these things vary, typically in response to some external excitation right external excitation or flux or whatever.

(Refer Slide Time: 24:03)



These physical quantities often depend on more than one spatial dimension as well as time right hence the rates of change that is the derivatives are most often with respect to multiple spatial dimensions as well as time right. So, these derivatives are derivatives of multiple independent variables right deal with and. So, we come with relations between partial derivatives and hence with partial differential equations. Now, partial differential equation is a vast subject analytical treatment of partial differential equations that there are entire courses which people teach it for those things right. I cannot even I cannot even attempt to try do try to cover any portion of it with any reasonable degree of coverage in during this course.

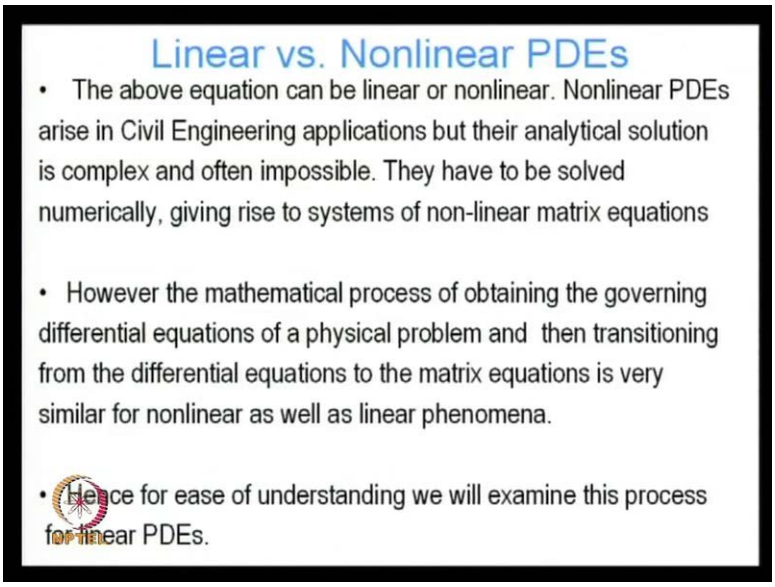
So, what a but I but I do think that when someone is trying to solve partial differential equations numerically, it is very important that at least we understand the basics of partial differential equations right and at and at least we understand some basic analytical tools that people can use to solve partial differential equations right. So, in the in the next few lectures we will talk about some simple analytical tools for solving partial differential equations, which are valid for a wide range of partial differential equations.

Before actually embarking, on how to do the numerical formulation right how to get that get the get from the partial differential equation to the matrix equation ? Right that step we are going to postpone it for some time and talk about some more brief introduction to analytical methods for partial differential equations. So, many of this partial differential

equations that arise in civil engineering applications are of second order and the most general second order partial differential equation in two independent variables can be written as the following f of x 1 x 2 phi x 1 phi x 2. So, you can see that this is a second order partial differential equation why? Because the highest derivative is of order two.

So, the subscript denotes derivative with respect to x 1 and x 2 right the highest derivative is of order two. So, but this function is a perfectly general function there is no guarantee that is that is going to be a linear function. it can be a nonlinear function of the derivatives right, but when it comes to nonlinear on linear partial differential equations we are not even going to try to try to solve them; try to talk to talk about analytical solutions we are going to do numerical methods that is the whole purpose of this course right. But, at least for linear partial differential equations we want to look at some common analytical methods that can be used for solving them.

(Refer Slide Time: 27:24)

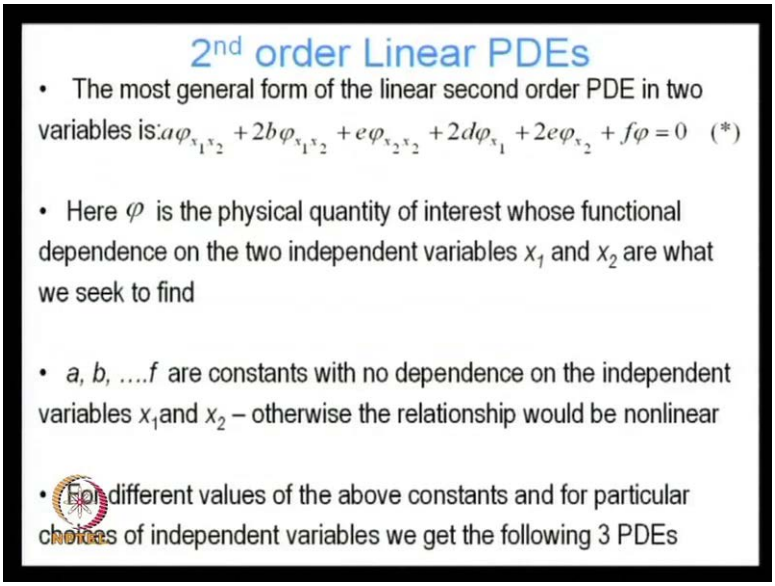


The above equation can be linear or nonlinear, non linear partial differential equations arise in civil engineering applications but their analytical solution is complex and often impossible they have to be solved numerically giving rise to systems of nonlinear matrix equations which we have already seen how to solve right; however, the mathematical process of obtaining the governing differential equations of a physical problem and then transitioning from the differential equations to the matrix equation is very common for both linear as well as nonlinear phenomena.

So, if we understand that basic mechanism and it is common to both linear and nonlinear equations right the equations themselves are different the solution methods; the matrix solution methods are different right we have seen that right Newton Raphson conjugate gradient steepest descent those are all for nonlinear equations while straightaway Gaussian elimination or link can be used for linear equations right.

So, the solution methods are very different the equations are very different but to go from the equations to the solution method, the process is not that different right some differences but not that different right. So, the process we are going to look at for linear equations right hence for ease of understanding we will examine this process linear partial differential equations right.

(Refer Slide Time: 29:05)

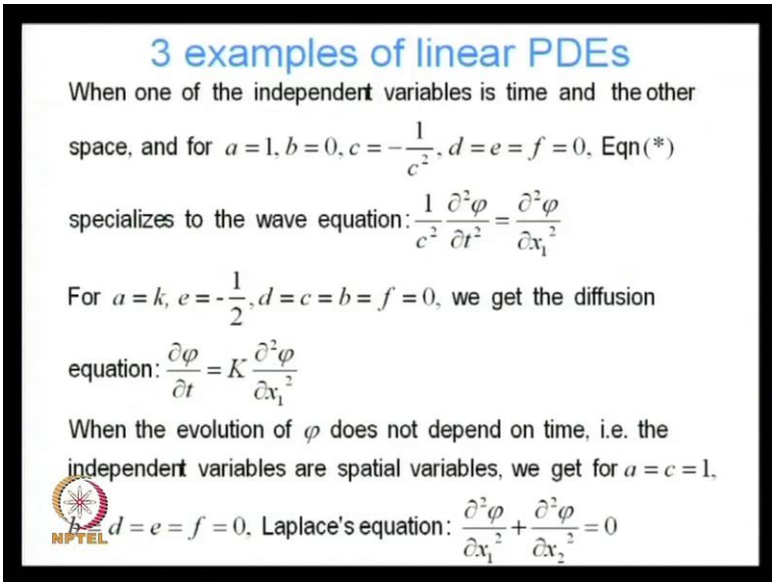


So, so the previous equation that we show that was for the general non linear equation the second order non linear equation if it is a linear equation then in two variables then it has this form. this is the general form of the linear equation where now a b e d e and f are all constants right, if a is a function of either of the independent variables or if for that matter any of these constants or functions of either of the independent variable we are no longer going to have a linear equation.

So, this is what makes it phi is the of course, physical quantity of interest whose functional dependence on the two independent variables x 1 and x 2 are what we seek to find right. So, we are given this equation and we want if we want to solve it analytically

we want people to tell us what is the function what is the functional form of phi in terms of x 1 and x 2 right a b f are constants with no dependence on the independent variables x 1 and x 2 otherwise the relationship would be nonlinear for different values of the above constants and for particular choices of the independent variables we get the following three PDEs right which are very very sort of generic PDEs right.

(Refer Slide Time: 30:34)



So, what are they well when one of the independent variables is time and the other is space remember that we are now considering only two independent variables there is no restriction that on that we can consider second order partial differential equations in more than two independent variables. But, for the for the time being let us consider that we are considering two independent variables right and one of them is space and the other is time.

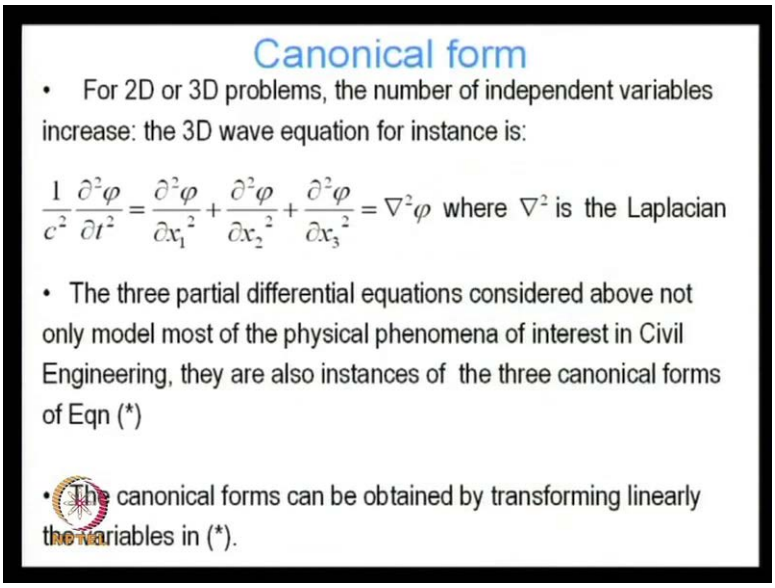
So, if that is true and then this these constants these constants a b e d e and f they sorry this should be c again this is a typo right this is a c right. So, if that is. So, these functions have the following values a is equal to 1 b is equal to 0, c is equal to minus 1 by c square which is related to I mean little complex notation. But, this c is very different from this c right this c is the wave speed right since we are taking about the wave equation this is the wave speed I did not want to change the symbol because this is sought to be universal for wave speed and the d, e and f are equal to 0. So, all the cross terms are there is no cross term and also the single derivatives also vanished right.

So, in that case this equation this equation becomes my wave equation right this equation becomes the wave equation. For a equal to K, K being the diffusivity right k being the diffusivity e equal to minus half d c b and f equal to 0 we get the diffusion equation So, this equation if you are if for civil engineers this is basically your equation of motion right. This is exactly the equation of motion which you encounter in mechanics right solid and structural mechanics this is the equation of motion this for you for civil engineers. This is probably you do not encounter it too much in structural mechanics, but you encounter it in heat transfer right where you consider heat flow right diffusion of heat right the phi becomes the right. So, in that case that becomes the heat transfer equation.

So, transient heat transfer that is the equation, when the evolution of phi does not depend on time. So, you can see here we have independent variables t and x time and spatial dimension when the evolution of phi does not depend on time that is the independent variables are spatial variables we get for a equal to c equal to one and b equal to d equal to e equal to f equal to 0 Laplace's equation right; Laplace's equation, which Laplace's equation is nothing. But, the wave equation if I ignore the transient term right. So, if you solve the equations of motion instead of solving the dynamical equations of motion if you solve the make the quasi static assumption we say that my solution my phi my variable phi is varying. So, slowly thus del two phi del t square is very small right. So, this part becomes zero. So, I get this right. So, that becomes that. So, I get Laplace's equation.

So, these are three main types of equation and we will see later on slightly later on that they have very they map onto certain things, which are known as the canonical forms. The canonical forms of second order partial differential equations each of these equations is a representative of a canonical form of a partial of the second order partial differential equation.

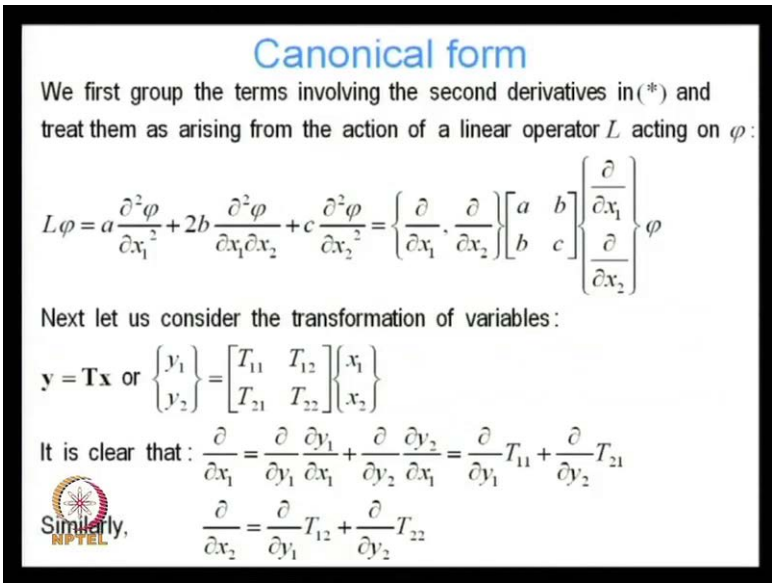
(Refer Slide Time: 34:30)



So, for 2D or 3D problems the number of independent variables increase the 3D wave equation for instance is this right. So, instead of x 1 I have x 2 x 2 x 3 right second order partial derivatives with respect to x 1 x 2 x 3, which I can write as 1 by c square l del 2 phi del t square is equal to Laplacian of phi well that is the Laplacian operator right.

So, there is three partial differential equations we considered like the wave equation the diffusion equation and Laplacian equation are modeled most of the physical phenomena of interest in civil engineering they are also instances of the three canonical forms of equation star which is my this equation right. So, they are examples of the three canonical forms of my general second order linear partial differential equation how can we get the canonical forms well you can get the canonical forms by transforming linearly the variables in star right by doing a linear transformation.

(Refer Slide Time: 35:45)



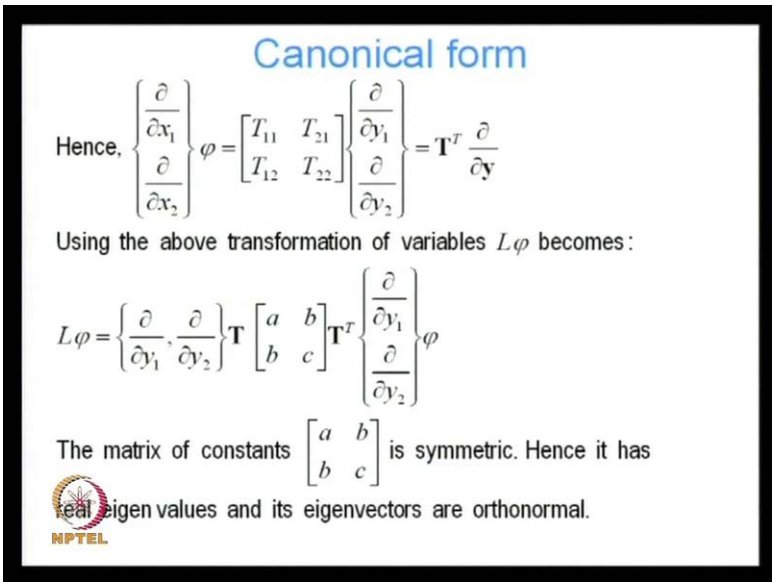
So, how do we do that well we first group the terms involving the second derivatives in star and treat them as arising from the action of a linear operator L acting on phi.

So, these are the terms which involve the second derivatives right second derivative with respect to x 1 with respect to x 2 and this is the cross term and we say that to get this by the action of some linear operator acting on phi right why is it a linear operator well because it satisfies linearity l of phi 1 plus phi 2 is equal to L of phi 1 plus l of phi 2 l of alpha phi is equal to alpha l of phi right. So, it satisfies definition of linearity. So, this is like a linear operator right and I rewrite this operator in this way. So, I have this row matrix I have this little row matrix del del x 1 del del x 2 right and then I have this little two by two matrix a b b c and then I have this color matrix del del x 1 del del del x 2. So, you can see that if I carry out this matrix vector the sorry the vector matrix vector a multiplication I am going to get exactly that equation.

So, next we consider a transformation of variables right we say that we are going to go from x 1 x 2 to another set of variables y 1 y 2 right. And this t is my transformation matrix. So, y is equal to t of x right or y 1 y 2 is equal to. So, I wrote out the transformation matrix in full form. So, just let us think of just as a transformation from x 1 x 2 to y 1 y 2. So, it is clear that using the chain rule using the chain rule del del x 1 equal to del del y 1 del y del y 1 del x 1 plus del del y 2 del y 2 del x 1 right and what is del y 1 del x 1 that is nothing, but t 1 1 from this equation right. So, does this equation

gives me y 1 is equal to T 1 1 x 1 plus T 1 2 x 2. So, del y 1 del x 1 equal to T 1 1 del y 2 del x 1 is equal to T 2 1 right. So, that gives me del del x 1 in terms of del del y 1 del del y 2 similarly I can get del del x 2 in terms of del del y 1 del del y 2 right. So, what do I get?

(Refer Slide Time: 38:24)

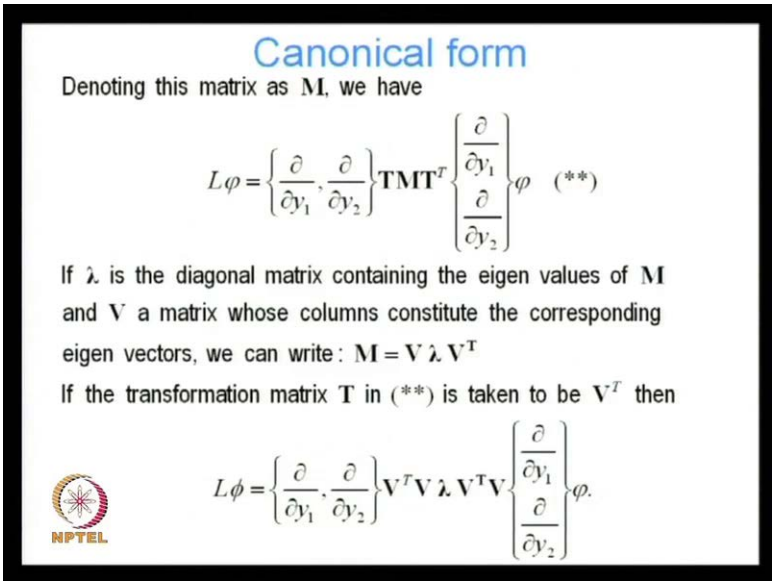


So, I can write del del x 1 del del x 2 phi is equal to this T 1 T 2 1 T 1 2 T 2 2 just grouping them together into a matrix again right. So, basically I am rewriting these last two equations as matrix equation right T 1 1 T 2 1 T 1 2 T 2 2 as del del y 1 del del y 2 which is nothing but T transpose del del y if del del y. I denote this column vector del del y 1 del del y 2 this is equal to T transpose del del y because my t had T 1 1 T 1 2 T 2 1 T 2 2 and this matrix has T y 1 T 2 one T 1 2 T 2 2.

So, this I get a relationship between the partial derivatives right using the above transformation of variables what does l phi become well l phi becomes, let us look at l phi again. So, it is del del x 1 del del x 2 right. So, that I can write as del del y 1. So, this is just the transpose of this equation right. So, this becomes del del y 1 del del y transpose t right. So, this becomes del del y transpose T and then I have my little sorry I have my little matrix a b b c. So, that remains the same that remains the same and then I have del del x 1 del del x 2 which I know is nothing, but t transpose del del y right. So, I get it in this form right.

Now, look at this matrix a b b c this matrix is symmetric right this matrix is symmetric and because it is symmetric I know that it is Eigen values cannot be imaginary numbers it is Eigen values got to be real numbers and it is Eigen vectors form an orthonormal basis right this form an orthonormal basis. So, I can do a spectral decomposition of that matrix I can find it is Eigen values I can find it is Eigen vectors and I can write that matrix that matrix a b b c denoting it as M right.

(Refer Slide Time: 40:49)



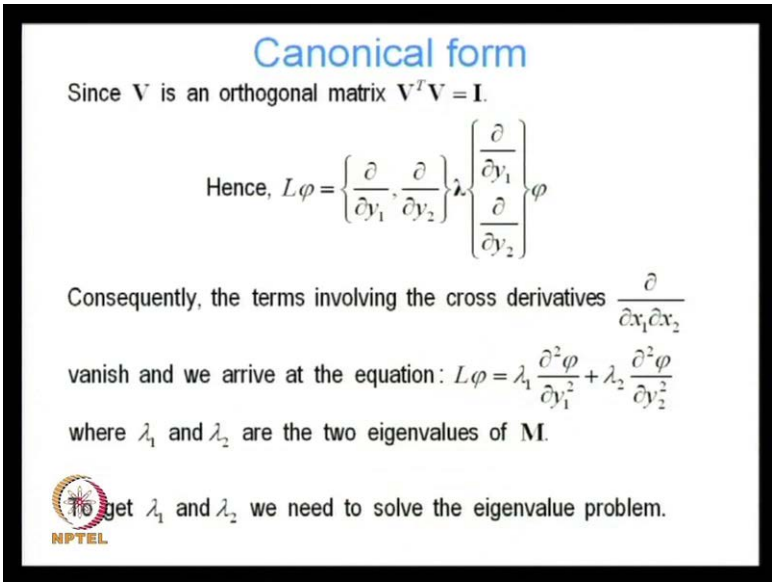
Denoting it as M I can write its write it as TMT transpose right. So, that is I am just rewriting that equation right am just rewriting that equation after replacing this a b b c by M right, but then the interesting part is that follows right because I because M is a symmetric matrix I can do a spectral decomposition and I can write it as V lambda V transpose right where V is a matrix its columns are the Eigen vectors of M right and lambda is a diagonal matrix whose diagonal elements are the corresponding Eigen values right. So, this is this is this is similarity transformation applied to symmetric matrices because it is symmetric. So, V transpose instead of V inverse right. So, is.

So, any way I can write as M is equal to V lambda V transpose right and then So now M become So, now, l phi is equal to this vector times t times V lambda V transpose T transpose times this vector right now what can I can I am free to choose my transformation matrix right I can choose any transformation I like right I am going to choose whichever transformation simplifies my equation most or best right. So, if I

choose the transformation matrix to be the matrix V transpose then what happens then if I replace T by V. So, I get a T by V transpose I get V transpose V lambda V transpose V right and we know that the matrix V since these Eigen vectors form an orthonormal basis the matrix V is a orthogonal matrix right.

Since it is an orthogonal matrix V transpose V are to be equal to the identity matrix right.

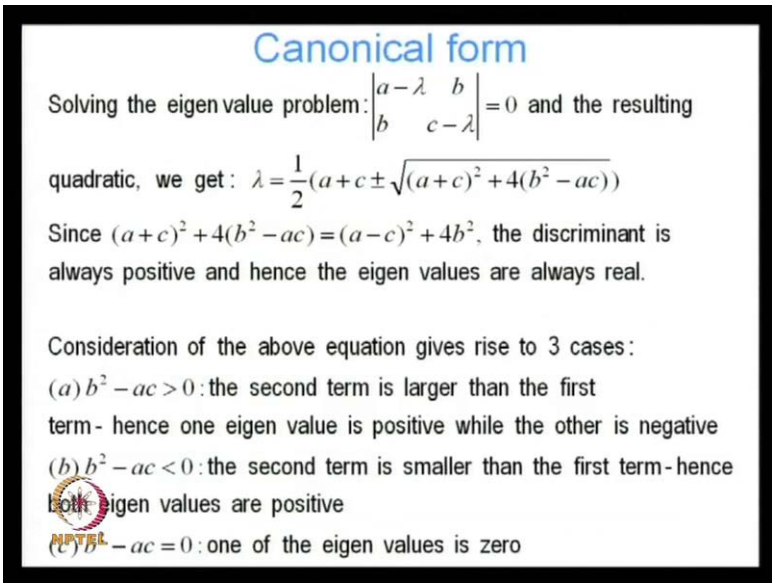
(Refer Slide Time: 43:02)



So, eventually we get l phi is equal to del del y 1 del del y 2 lambda del del y 1 del y 2 phi right and notice that the matrix lambda is a diagonal matrix. So, that is the biggest advantage we are getting. Because now, lambda is a because lambda is a diagonal matrix and we could bring it in this form it is found that there are no cross terms right. The cross derivatives do not there no more cross derivatives right and we arrive at the equation we only contain there are no cross. So, we only have derivatives with respect to y 1 and y 2. So, there are no cross terms rights because we did this transformation and because we choose our transformation matrix to be the matrix of Eigen vectors of my little coefficient matrix transpose that is why we could get this simplification right.

So, if we can do that we get we get this uncoupling right we get rid of the cross derivatives right; however, we still need to find lambda 1 lambda 2 find lambda 1 lambda 2 well that is just easier. Now, because we have to find we saw if the Eigen value problem two by two Eigen value problem find out lambda.

(Refer Slide Time: 44:27)



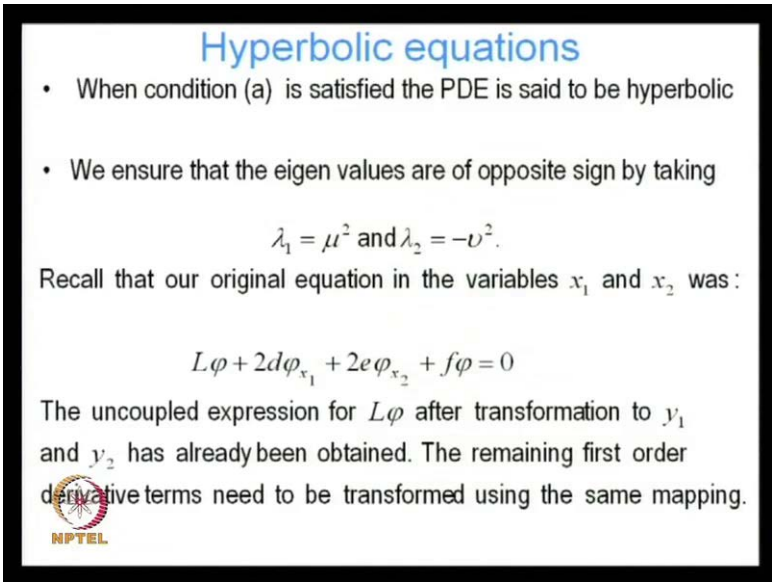
So, we saw if the Eigen value problem this is my Eigen value problem determinant of a minus lambda b b c minus lambda is equal to 0 resulting we get a polynomial and this case in situ two by two it is a polynomial is a quadratic we solve for lambda we get that equation in terms of a and c right.

So, you look at that equation if you look at the term under the square root, it actually it is you can write it as a sum of two squares right; you can write it as a sum of two squares and since you do that it is always going got to be positive right. It is also always got to be positive and hence the Eigen values are always real which we already knew right because this is a symmetric matrix. So, it is a Eigen value is got to be real, but this is just we are seeing that.

Now, consideration of the above equation gives rise to three cases what are those three cases? well one case is when b square minus a c is positive right when b square minus a c is positive you can see that the second term is larger than the first term right because this contribution b square minus a c is going to increase the term within the square root right. So one of, so at least one of the terms is going to be negative right while the other term for the positive sign is going to be positive right. So, b square minus a c is positive the second term is larger than the first term hence one of the Eigen values is positive while the other is negative. If b square minus a c is negative the second term is smaller than the first term in that case we are guaranteed that both are Eigen values are going to be

positive right; value of b square minus a c is equal to 0 we are sure that one of the Eigen values is going to be 0 because a plus c minus a plus c that is going to give me 0 and a plus c plus a plus c that is going to be the positive Eigen value right. So, each of these cases correspond to a canonical form for the second order linear partial differential equation.

(Refer Slide Time: 46:46)



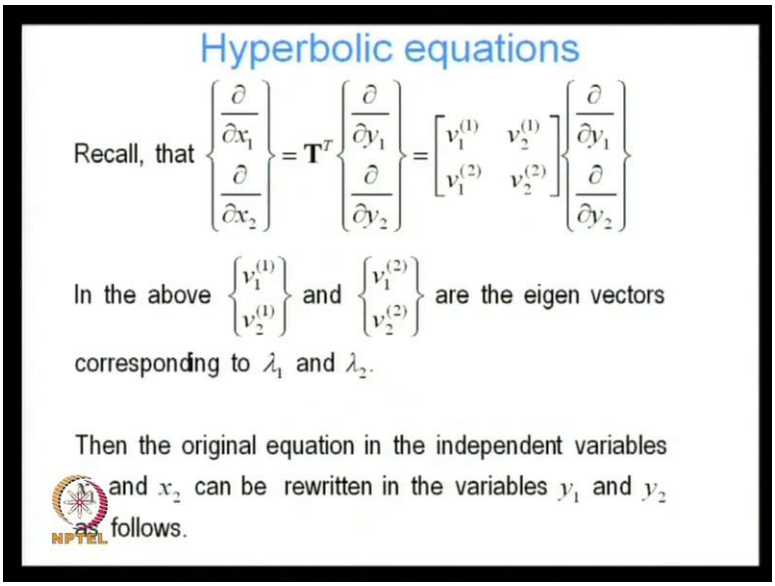
When is a condition a is satisfied; when condition a is satisfied the partial differential equation is said to be hyperbolic right we. So, that that is just a definition right when that condition is satisfied the partial differential equations what is that condition? Well it is Eigen values are of opposite signs one of them is positive one of them got to be negative right how do we ensure that we said lambda one the first Eigen value is some square and the second Eigen value is the negative of its square say it got to be a negative number right recall that our original equation the in the variables x 1 and x 2 was like this it was L phi plus 2 d phi x 1 plus 2 e phi x 2 plus f phi equal to 0 right .

The uncoupled expression for l phi after transformation to y 1 and y 2 has already been obtained. So, after we did the transformation that we did we know that this l phi is going to be uncoupled right the term involving the second derivatives the L phi term the L phi contribution it is more than one term that is going to be uncoupled that was a cross derivatives are going to vanish right but so but this so we have to what we have to do we

have to transform this single derivative the first order first order derivatives into the y 1 y 2 space right we have to do that right we have to be consistent right.

So, the remaining first order derivative terms need to be transformed using the same mapping right which was T T is equal to V transpose using that mapping right.

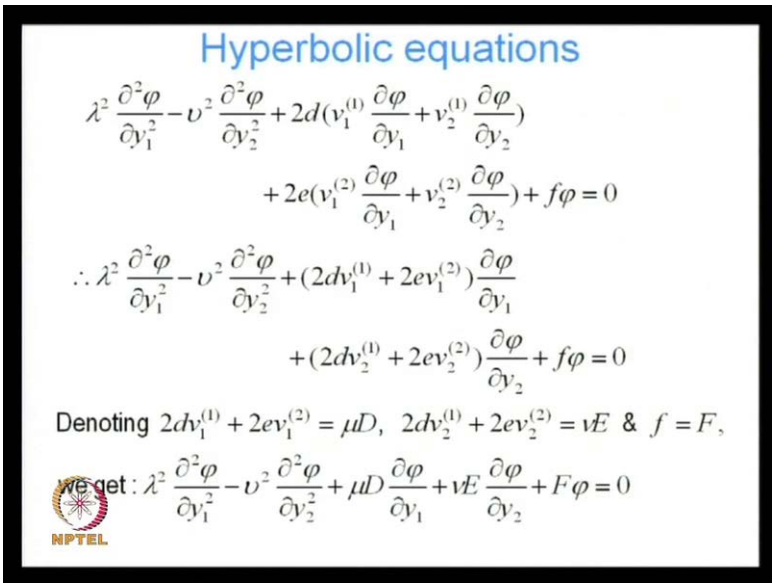
(Refer Slide Time: 48:31)



So, recall that del del x 1 del del x 2 is equal to t transpose del del y 1 del del y 2 and t is we know is the matrix of Eigen vectors transpose. So, if my Eigen vectors v 1 1 v 2 1 correspond to my Eigen value lambda 1 and my Eigen vector v 1 2 and v 2 2 corresponds to my lambda to my Eigen value lambda 2 then I can write it like this right.

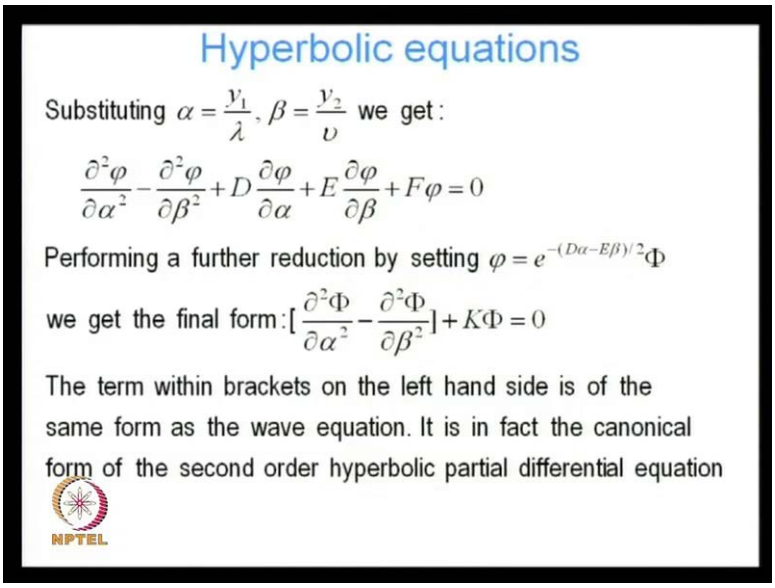
So, once I do that my entire original equation is now transformed in terms of my new variables y 1 and y 2 right.

(Refer Slide Time: 49:13)



And, what form does it take well it takes this form lambda square del 2 phi del y square minus nu square del 2 phi del y y 2 square. So, this is the l phi term that we saw earlier right that is del phi term and this is what happens to my first order partial derivatives when I write them in terms of y 1 and y 2 in rather than x 1 and x 2 right and then I pull out the coefficients of del phi del y 1, I get this term involving 2 d v 1 1 plus 2 e v 1 2 right and similarly I pullout the coefficients of 2 del phi del y 2 I get 2 d v 2 1 1 plus 2 e v 2 2 right plus I have the term f phi right I have the term f phi right and then I do just some again, I want to replace this 2 d v 1 1 plus 2 e v 1 2 by mu D 2 d v 2 1 2 e v 2 2 by nu e right and I replace small f by capital f I get this equation right I get this equation at the bottom of the slide right.

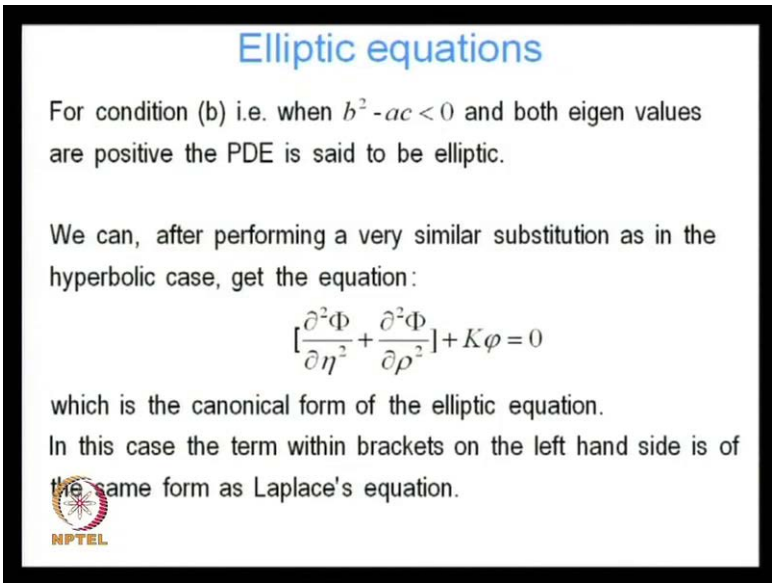
(Refer Slide Time: 50:32)



Then I do a further transformation of variables I say that I am going to replace y 1 y 2 by new variables alpha and beta. And this is how I define my alpha and beta alpha equal to y 1 by lambda beta equal to y 2 by nu right and then I get this form del 2 phi del alpha square minus del 2 phi del beta square plus d times first derivative plus e with respect to alpha e times first derivative with respect to beta plus that term then we perform a third transformation we replace phi by this function in terms of capital phi right and if we do that we get a final form of this equation right the final form of this equation.

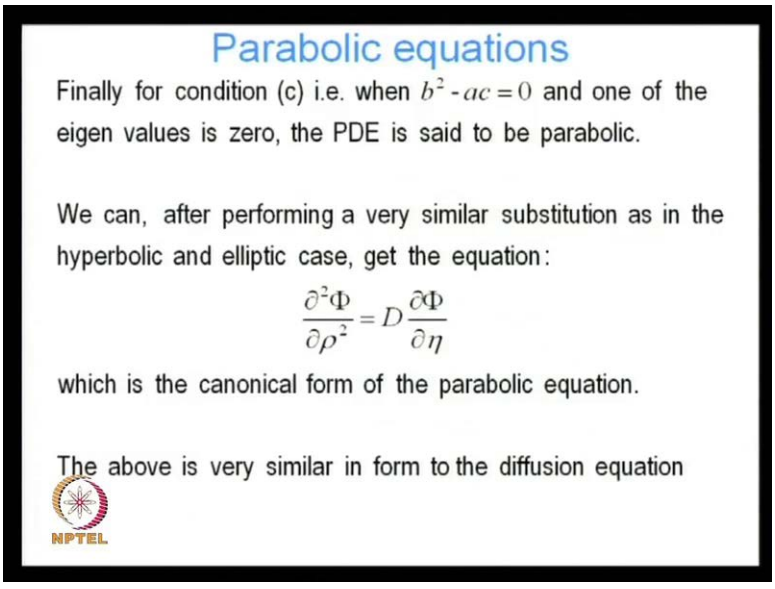
And if you look at the term within brackets on the left hand side which is of the same form as the wave equation right It is In fact, the canonical form of the second order hyperbolic partial differential equation right it is the. So, all hyperbolic partial differential equations can be obtained from this form right. So, whether it is the Helmholtz equation the wave equation any equation any hyperbolic second order partial differential equation you can get from this form. So, this is the canonical form of the hyperbolic partial differential second order linear partial differential equation right.

(Refer Slide Time: 52:04)



Similarly, we can do similar things right when b square minus a c is less than 0 and both the Eigen values are positive then the partial differential equation is said to be elliptic, it is an elliptic partial differential equation if we perform very similar substitutions as in the hyperbolic case we get this equation. del 2 phi del eta square plus del 2 phi del rho square plus k phi equal to 0, which is again the canonical form of the elliptic equation. So, all elliptic equation whether be Laplace's equation or Poisson's equation any elliptic equation you are going to get from this canonical form of the of the elliptic equation right.

(Refer Slide Time: 52:50)



And finally, for the condition when b square minus a c is equal to 0 that is when one of the Eigen values is 0 the partial differential equation is said to be parabolic partial differential equation. Again, performing substitutions very similar to what we perform for the hyperbolic case we can get it is canonical form, which is del 2 phi del rho square is equal to d del phi del eta, which is the canonical form of the parabolic equation and this equation if you look at that that is very similar to the diffusion equation.

So, we looked at it is a linear second order partial differential equations; we looked at the canonical forms. Next lecture, we are going to talk about some how do we get some of those physical equation may be we look at the wave equation right how do we get that equation? And then we will try to find look at some simple analytical techniques for solving that those equations right before actually moving on to numerical techniques.

Thank you.