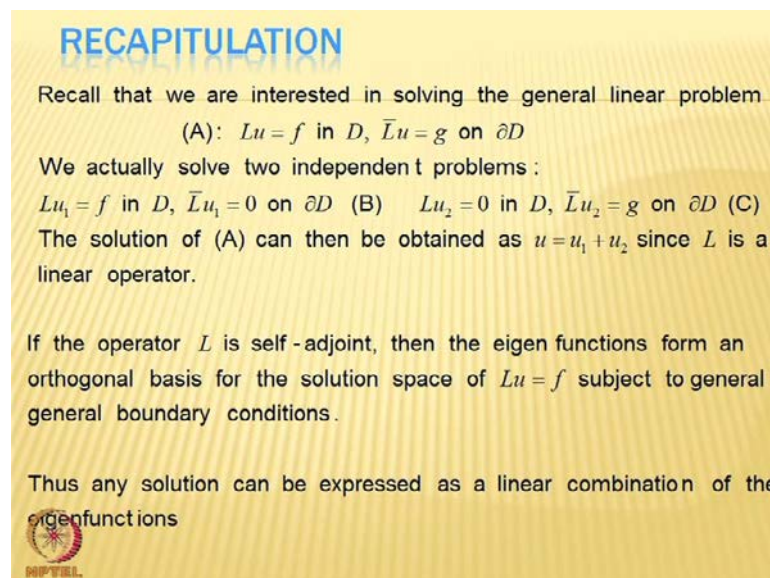


**Numerical Methods in Civil Engineering**  
**Prof. A. Deb**  
**Department of Civil Engineering**  
**Indian Institute of Technology, Kharagpur**

**Lecture - 37**  
**Orthogonal Basis Functions for Solving PDE's-II**

In this lecture 37 of our series on numerical methods in civil engineering, we will continue our discussion on using orthogonal basis functions for solving partial differential equations.

(Refer Slide Time: 00:30)



**RECAPITULATION**

Recall that we are interested in solving the general linear problem

(A):  $Lu = f$  in  $D$ ,  $\bar{L}u = g$  on  $\partial D$


We actually solve two independent problems :

$Lu_1 = f$  in  $D$ ,  $\bar{L}u_1 = 0$  on  $\partial D$  (B)     $Lu_2 = 0$  in  $D$ ,  $\bar{L}u_2 = g$  on  $\partial D$  (C)

The solution of (A) can then be obtained as  $u = u_1 + u_2$  since  $L$  is a linear operator.

If the operator  $L$  is self-adjoint, then the eigen functions form an orthogonal basis for the solution space of  $Lu = f$  subject to general boundary conditions.

Thus any solution can be expressed as a linear combination of the eigenfunctions



To recapitulate we recall that we are interested in solving the general linear problem where  $L$  is a linear operator,  $L$  operating on  $u$  is equal to  $f$  where  $f$  is a known function in a certain domain  $D$ . So, we are interested in finding that  $u$  which satisfies that equation subject to certain boundary conditions, boundary conditions defined on the domain main boundary which we denote as  $\partial D$ . So, on the boundary  $\bar{L}u = g$ . So, on the boundary  $\bar{L}u$  is the boundary condition. So,  $\bar{L}$  operating on  $u$  is an operator which defines the boundary condition.

So,  $\bar{L}$  operating on  $u$  is equal to  $g$  on the boundary. So, we are interested in solving this problem and we want to use orthogonal basis functions, for solving these problems.

And this  $L$  is typically a differential operator and since we are interested in solving real world problems in multi-dimensions, it is usually a partial differential operator, which involves partial derivatives, right?

And actually we in order to solve this problem we actually solve two problems, we divide this problem into two parts and we solve two parts the problem has part B. So, the problem A has part B and part C on part B it involves solving the problem  $L u_1$  is equal to  $f$  in  $D$  and  $L \bar{u}_1$  is equal to 0 on  $\partial D$ . And part B involves solving the problem  $L u_2$  is equal to 0 in  $D$  and  $L \bar{u}_2$  is equal to  $g$  on  $\partial D$ . So, you can see that is divided into two parts, in first part of the problem we satisfy we divide the solution into two parts  $u$  is equal to  $u_1$  plus  $u_2$ . And in the first part we solve the we solve the problem  $L u_1$  equal to  $f$  in  $D$  and  $L \bar{u}_1$  is equal to 0 on  $\partial D$ .

So, in the first part in within the domain it satisfies  $L u_1$  equal to  $f$ . So, it satisfies the actual equation  $L u$  equal to  $f$  within the domain, but on the boundary it satisfies the boundary condition in a homogeneous fashion. So,  $L \bar{u}_1$  is equal to 0 in  $\partial D$  in the second part of the problem  $L u_2$  is equal to 0 in  $D$  and  $L \bar{u}_2$  is equal to  $g$  on  $\partial D$ . So, within the domain it satisfies the differential operator in homogeneous manner and on the boundary, it satisfies the exact boundary condition. And we can divide this problem our original problem A into two parts B and C, solve them separately and then use those add those two solutions  $u_1$  and  $u_2$  together and claim that  $u_1$  plus  $u_2$  actually satisfies my original system A, because of the fact that my operator  $L$  is a linear operator, right?

Because it is a linear operator only because of that we can use this approach divide, the solution into two parts one part satisfies the actual differential equation and the boundary condition homogeneously. The other part satisfies the differential equation in the domain not the actual differential equation, but the homogeneous part of the differential equation in the domain, that has the actual boundary conditions which satisfies the actual boundary conditions.

So, get the two solutions we add them together and we get our final solution, and we saw earlier on that in our last lecture that if we can solve the problem B, it is very easy to find the solution C. So, we basically focus our efforts on solving the problem B basically instead of solving the problem  $L u$  equal to  $f$  in  $D$  and  $L \bar{u}$  is equal to  $g$  on  $\partial D$ , we

solve the same problem, but with homogeneous boundary conditions  $L u$  is equal to  $f$  in  $D$   $L \bar{u} = 0$  on  $\partial D$ , this part we can easily solve.

And we also saw that if the operator is self adjoint and we explain what is a self adjoint operator  $L u$  the inner product of  $L u$  with  $v$  is equal to  $u$  inner product of  $u$  with  $L v$  right. So, that is the self adjoint operator. So, if the operator is self adjoint then it has the very nice property that its Eigen functions form an orthogonal basis for the solution space of  $L u$  is equal to  $f$  subject to very general boundary conditions right. So, if  $L$  is a self adjoint operator, its Eigen functions form a orthonormal basis right and since it form they form an orthonormal basis, it is possible to express any solution  $u$  to the system  $B$  in terms of in terms of a linear combination of those basis functions, right? We have talked about these things in our earlier lectures this is just to recapitulate.

(Refer Slide Time: 06:05)

**RECAPITULATION**

Given an appropriate set of linearly independent basis functions  $\{\psi\} = \{\psi_1, \psi_2, \dots, \psi_n\}$ ,  $\tilde{u} = c_1 \psi_1 + c_2 \psi_2 + \dots + c_n \psi_n$

To solve  $L \tilde{u} \approx f$  with  $\tilde{u} = c_1 \psi_1 + c_2 \psi_2 + \dots + c_n \psi_n \in H^n$ , we have to satisfy:  $c_1 L \psi_1 + c_2 L \psi_2 + \dots + c_n L \psi_n \approx f$

The criterion used to enforce the condition that  $f$  is indeed approximately equal to  $\sum_{i=1}^n c_i L \psi_i$  determines the solution method

If Least Squares Minimization is used, this condition is enforced by requiring that  $\left\| f - \sum_{i=1}^n c_i L \psi_i \right\|$  be a minimum

This leads to the requirement  $(f - \sum_{i=1}^n c_i L \psi_i, L \psi_j) = 0 \quad \forall \psi_j, j=1, 2, \dots, n$  (\*)

Any solution can therefore, be expressed as a linear combination of the Eigen functions; that means, that if we have if we know the Eigen functions of the operator  $L$ , if we have if we know the solution of the problem the Eigen value problem recall is given by  $L u$  is equal to  $\lambda u$  or  $L u$  equal  $\lambda u$ . So, if we can solve the if we know the solution of that of that problem, if we know those Eigen functions and since it is a self adjoint operator  $L$  is a self adjoint operator, we know that that those Eigen functions form an orthonormal basis. Then any solution  $u$  tilde can be expressed as a linear combination of those basis functions of those Eigen functions.

Those Eigen functions now form a basis right they form a basis for the function space right and. So,  $\psi_1 \psi_2 \dots \psi_n$  are my  $n$  independent basis functions and  $c_1, c_2$  through  $c_n$  are arbitrary constants. So, any solution to  $u$  to that equation  $L u = f$  in  $D$  and  $L u = 0$  on  $\partial D$  can be expressed as in this form. So, now our basic our basic problem is to solve  $L u = f$ , where we know that  $u$  is of this form and we have to and if this  $u$  is truly has solution then we know that this condition has to be satisfied right. Since,  $u$  has a solution which has substitute  $u$  here. So, we get  $c_1 L \psi_1 + c_2 L \psi_2 + \dots + c_n L \psi_n$  and that must be at least close enough, it must in order to satisfy the equation in a reasonable manner, it must be very close to the given function  $f$ .

How close determines the accuracy of the solution and the criteria we use to enforce the condition that  $f$  is indeed approximately equal to  $c_1 \psi_1 + \dots + c_n \psi_n$  determines the solution method. So, there are many, many ways in which we can enforce that condition that this minus  $f$  in some norm is very, very small if that if this thing this left hand side minus  $f$  norm of that in some norm is very, very small. Then we know that our solution  $u$  our solution approximate solution  $u$  is an acceptable solution to my system. And the way we enforce that criterion that condition determines the solution method.

So, one method which we have already talked about earlier is the least squares minimization method. So, we basically say that I want to calculate the norm of this  $f$  minus  $c_1 \psi_1 + \dots + c_n \psi_n$  I have calculate the norm of this, maybe I just calculate the norm using the  $L^2$  norm, right? And it take the inner product right and make sure that the square root of the inner product that is a minimum right. So, that condition can be used and that leads to this requirement we have seen that earlier also it leads to the requirement that this inner product  $\int L \psi_j u$  must be equal to 0 for all  $\psi_j, j$  is equal to one through  $n$  right.

For this to be a minimum for this to be a minimum for this norm to be a minimum, it must satisfy this condition. And we saw last time that leads this leads to a system of equations for the coefficients  $c$ , if we solve that system of equations we can find the  $c$ 's. And once we find the  $c$ 's we can construct our solution  $u = c_1 \psi_1 + \dots + c_n \psi_n$  like this. So, that is using least squares minimization.

(Refer Slide Time: 10:05)

**RECAPITULATION**  
Unlike the least squares minimization technique and the collocation method which follows therefrom, the Galerkin method determines the coefficients using a different criterion

In the Galerkin method the residual function  $f - \sum_{i=1}^n c_i L \psi_i$  is required to be orthogonal to each and every member of  $H^n$

This imposes the criteria  $(f - \sum_{i=1}^n c_i L \psi_i, \psi_j) = 0 \quad \forall \psi_j \in H^n,$   
 $j = 1, 2, \dots, n$  (\*\*)

Comparison of (\*\*) with (\*) reveals the difference between the Galerkin and the least squares method. In the least squares method the residual is required to be orthogonal to each and every linearly independent component of the best approximation to  $f$  i.e. to each and every  $L \psi_j$

So, unlike the least squares minimization technique and the collocation method which we also talked about briefly there is another method which is known the Galerkin method and that uses a different criteria recall, I said that there are different methods for finding the solution just depend on the criteria we use to enforce the condition right to enforce this condition that  $f$  is approximately equal to this right. So, the criteria is changing. So, for the least squares minimization method this was the criteria that this has to be a minimum.

In the Galerkin method we say that this, if we can think of this as the residual function the  $f$  minus sigma  $i$  equal to 1 to  $n$   $c_i L \psi_i$  that is the residual that is the part by which my approximate solution, which I use which I construct like this does not match my given function that is the residual. The residual is orthogonal to each and every member of this of this space  $H^n$ , recall the space  $H^n$  the size are the basis functions for the space  $H^n$ . So, this residual I have to make it orthogonal to each and every member of that space  $H^n$  right. So, this is basically this if we if I use this criteria instead of that criteria, I have the Galerkin method, while if I use this criteria I am using least squares minimization.

So, the difference in this ((Refer Time: 11:51)), but it is very, very important the only difference that you will notice between this and that is that here I am making it orthogonal to each and every  $L \psi_j$ . While in the Galerkin method I am I am enforcing I

am requiring that the residual be orthogonal to each and every basis function of my function space  $H^n$  that is the fundamental difference. So, the comparison of these two equations reveals the difference between the Galerkin and the least squares method in the least squares method. The residual is required to be orthogonal to each and every linearly dependent independent component of the best approximation to  $f$  that is it has to be orthogonal to each and every  $L\psi_j$ .

(Refer Slide Time: 12:42)

**GALERKIN METHOD**

In the Galerkin method the residual function  $f - \sum_{i=1}^n c_i L\psi_i$  satisfies

$$\left(f - \sum_{i=1}^n c_i L\psi_i, \psi_j\right) = 0 \quad \forall \psi_j \in H^n, \quad j = 1, 2, \dots, n \quad (**)$$

Assuming  $L$  is self adjoint as well as positive definite, a new norm and associated inner product can be defined on the function space  $H^n$  as follows:  $\langle u, v \rangle = (u, Lv)$ ,  $\langle u, u \rangle = (u, Lu) = \|u\|^2 \quad \forall u, v \in H^n$

This inner product, as distinct from the inner product associated with the  $L_2$  norm, is defined with respect to the operator  $L$

If  $L$  is positive definite and  $u^*$  is the solution of (B) i.e.  $Lu^* = f$  on  $D$  and  $\bar{L}u^* = 0$  on  $\partial D$ , then the Galerkin method gives the best possible approximation to  $u^*$  in the function space  $H^n$  measured in the  $\langle \cdot \rangle$  norm

While in case of the Galerkin method the residual function has to be orthogonal to each and every basis function  $\psi_j$ . So, that is the fundamental difference, now if we assume that  $L$  is self adjoint as well as positive definite a new norm an associated inner product can be defined on the function space  $H^n$ . So, instead of using my  $L_2$  norm, I can define a new norm then that new norm, I can define a number of norms we have discussed norms earlier on we have our  $L_2$  norm we have the maximum norm we have so many different norms right.

So, this we defined another norm like this and this norm I denote by this sort of brackets right  $\langle u, v \rangle$ ,  $u$  is a is may be function  $u$  and  $v$  is another function  $v$ . So, I define the inner product of  $u$  and  $v$  as  $\langle u, v \rangle = (u, Lv)$  inner product of  $u$  with  $Lv$  this is my conventional inner product. And this new inner product I defined as  $\langle u, v \rangle = (u, Lv)$ , and this automatically gives rise to a norm because  $\langle u, u \rangle = (u, Lu)$  and that is norm of  $u$  square.

So, now this gives a norm which is somewhat different from my usual  $L^2$  norm, which I get from my inner product because here I am taking not just the inner product of the function with itself, I am taking the inner product of the function with  $Lu$  right. So, this inner product is distinct from the inner product associated with the  $L^2$  norm is defined with respect to the operator  $L$ . So, this inner product actually depends on the operator, it depends on the operator is defined with respect to the operator.

So, that is a norm and if  $L$  is positive definite and  $u^*$  is the solution is the exact solution to  $B$ . That is  $Lu^* = f$  on  $D$  and  $\bar{L}u^* = 0$  on  $D$  we can show that the Galerkin method has some very, very useful properties and what is that property? Where Galerkin method gives the best possible approximation to  $u^*$  in the function space  $H_n$  measured in this norm, that is that basically explains the popularity of the Galerkin method because we are guaranteed that given this norm given that we can define this norm. The Galerkin method gives the best possible the best possible approximation to the true solution this  $u^*$  is my true solution.

If I solve this problem using the Galerkin method that whatever  $u$  I get from solving the Galerkin method that is going to give me the best possible approximation to  $u^*$ , measured in this norm. Not measured in the usual norm usual norm, but measured in the norm we just defined with respect to the operator  $L$ . That basically contributes to the popularity of the Galerkin method and as we will discuss later, it also contributes the popularity of methods such as the finite element method which are basically Galerkin methods.

(Refer Slide Time: 16:08)


**BEST APPROXIMATION PROPERTY**

From the solution of the error minimization problem, the best possible solution  $u$  in  $H^n$  in the  $\langle \cdot, \cdot \rangle$  norm must satisfy the requirement  $\langle u - u^*, v \rangle = 0$  for each and every  $v \in H^n$

However  $\langle v, u - u^* \rangle = (v, L(u - u^*)) = (v, Lu) - (v, Lu^*)$   
 $= (v, Lu) - (v, f) = (v, Lu - f)$

Thus requiring  $\langle v, u - u^* \rangle = 0$  is equivalent to requiring  $(v, Lu - f) = 0$

But from (\*\*)  $(v, Lu - f) = 0$  is the Galerkin condition. Hence finding the best possible solution  $u \in H^n$  in the  $\langle \cdot, \cdot \rangle$  norm is equivalent to finding the solution using the Galerkin method.



So, from the solution of the error minimization problem the best possible solution  $u$  in  $H^n$  in the  $u$  norm, must satisfy the requirement that  $u - u^*$  with respect to this product must be 0 for each and every  $v$  belonging to  $H^n$ . This is the difference this is the Galerkin solution  $u^*$  is the true solution,  $u - u^*$  is the error right the true solution the Galerkin solution minus  $u^*$  that is the error. The error with the inner product which take the product with each and every  $v$  for because each and every  $v$  belonging to  $H^n$  I know that it is linearly independent and this must be equal to 0.

So, what does this mean this means that  $v(u - u^*) = 0$  then we can translate it in terms of our usual inner product that becomes  $vLu - vLu^*$  right because this inner this inner product is defined with respect to  $L$ . So,  $vLu - vLu^*$  and then taking advantage of the linearity of the operator, we can write it as  $vLu - vLu^*$ . We can write it as  $vLu - vLu^*$  that is  $vLu$  and  $vLu^*$  I know is the exact solution. So,  $vLu^*$  must be equal to  $f$  so  $vLu - f$ .

So, what does it mean it means that requiring this to be equal to 0 is equivalent to requiring  $vLu - f = 0$ . If I if I said if I require this to be equal to 0; that means, that this has got to be equal to 0; however, I know that this is actually my Galerkin condition, my Galerkin condition is exactly that right that the residual must be orthogonal to each and every basis function of the space  $H^n$ . So, finding the best



possible solution  $u$  belonging to  $H^n$  in this norm is equivalent to finding the solution using the Galerkin method right. So, finding the best possible solution in this norm is equivalent to finding the solution using the Galerkin method.

So, what does it mean it means that the Galerkin solution always gives me the best possible solution in this norm stating it just the converse of that. That means, that if I find the solution using the Galerkin method, I am automatically finding the best possible solution in this norm right because this is equivalent to that. So, finding the solution using the Galerkin method, whatever solution I find gives me the best possible solution in this norm. So, if I minimize this right if I minimize this and find the best possible solution, which for this then I am also solving the Galerkin problem.

(Refer Slide Time: 19:36)

**MINIMIZATION OF A FUNCTIONAL**

It can be shown that the best approximation property of the Galerkin solution also ensures that it minimizes the functional  $(u, Au) - 2(f, u)$  if this functional can be written as  $\langle u - u^*, u - u^* \rangle - \langle u^*, u^* \rangle$ :

$$\begin{aligned} \langle u - u^*, u - u^* \rangle &= (u - u^*, Au - Au^*) \\ &= (u, Au) - (u, Au^*) - (u^*, Au) + (u^*, Au^*) \\ &= (u, Au) - (u, f) - (Au^*, u) + (u^*, Au^*) \quad (\text{due to symmetry}) \\ &= (u, Au) - (u, f) - (f, u) + (u^*, Au^*) = (u, Au) - 2(f, u) + \langle u^*, u^* \rangle \end{aligned}$$

This establishes equivalence. Thus if  $u = u^*$ ,  $(u, Au) - 2(f, u)$  is minimized among all functions  $u$  for which the functional  $(u, Au) - 2(f, u)$  exists.

So, it can be shown that the best approximation of the Galerkin method, this is another ramification of that the best approximation of the Galerkin method also ensures that it minimizes this functional  $u A u$  minus  $2 f u$ , where  $A$  is any operator any liner operator which is self adjoint right, if I can write a functional like this  $u$  comma  $A u$  minus  $2 f u$ .

Then I can it can be shown that the best approximation property of the Galerkin method ensures that this functional is also minimized, provided that this functional  $u A u$  minus  $2 f u$  can be written in this form. Provided this functional can be written in this form, and while I am going to just show that it is indeed possible for our assumptions, for the assumptions that we made that  $A$  is a liner operator that  $A$  is self adjoint that it is

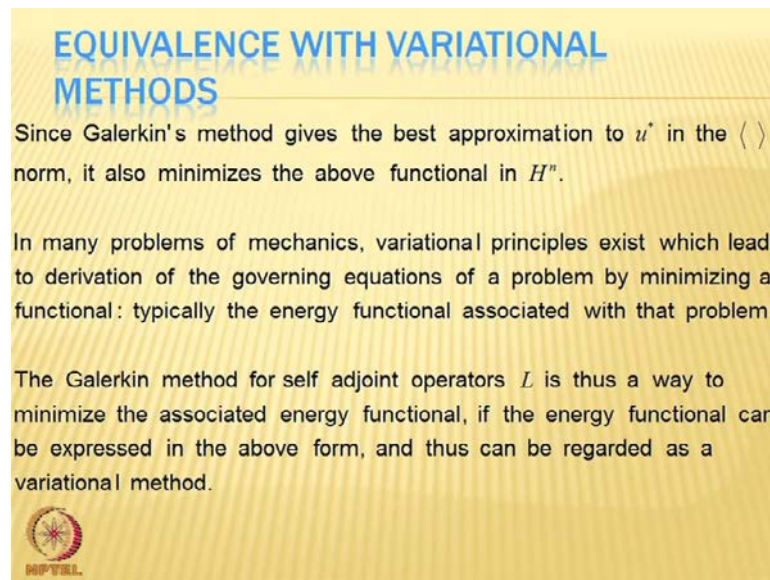
possible to write this in this form this functional in this form how are we going to show it let us look here.

So,  $\|u - u^*\|^2$  inner product of  $u - u^*$  with itself  $\|u - u^*\|^2$  defined in this norm right this is equal to again because of the definition of this norm that is the inner product usual inner product, but now with  $u - u^*$  and  $Au - Au^*$  that is the definition of this product. So, this can be written like this then again taking advantage of the linearity of the operators, we can write it like this  $\langle u - u^*, Au - Au^* \rangle = \langle u, Au \rangle - \langle u, Au^* \rangle - \langle u^*, Au \rangle + \langle u^*, Au^* \rangle$ , this remains the same  $\langle u, Au \rangle$  becomes  $\langle u, f \rangle$  because  $u^*$  is the exact solution. So,  $\langle u, Au^* \rangle = \langle u^*, Au \rangle$  here we are using the fact that the operator is self adjointed. So,  $\langle u^*, Au \rangle = \langle Au^*, u \rangle$  and here again we have  $\langle u^*, Au \rangle$ .

So, pulling all these terms together and again using the fact that  $Au^* = f$ . So, we have  $\langle u, f \rangle - \langle f, u \rangle$  and since the this since the order does not matter in this inner product. We can write it as  $\langle u, Au \rangle - 2\langle f, u \rangle + \langle u^*, Au^* \rangle$  in inner product of  $u - u^*$  using this new inner product.

So, we can write it like this is actually establishes equivalence because if  $u$  is equal to  $u^*$  what happens,  $u$  is equal to  $u^*$  then this thing is going to be minimized right because this thing is  $\langle u, Au \rangle - 2\langle f, u \rangle + \langle u^*, Au^* \rangle$  is equal to this minus that right. If this becomes 0 say right that is the minimum possible value for this part right, then it becomes if that is the minimum possible value for  $\langle u, Au \rangle - 2\langle f, u \rangle$ . So, thus if  $u$  is equal to  $u^*$   $\langle u, Au \rangle - 2\langle f, u \rangle$  is minimized among all functions  $u$  for which the functional  $\langle u, Au \rangle - 2\langle f, u \rangle + \langle u^*, Au^* \rangle$  exists.

(Refer Slide Time: 23:25)




**EQUIVALENCE WITH VARIATIONAL METHODS**

Since Galerkin's method gives the best approximation to  $u^*$  in the  $\langle \cdot \rangle$  norm, it also minimizes the above functional in  $H^n$ .

In many problems of mechanics, variational principles exist which lead to derivation of the governing equations of a problem by minimizing a functional: typically the energy functional associated with that problem.

The Galerkin method for self adjoint operators  $L$  is thus a way to minimize the associated energy functional, if the energy functional can be expressed in the above form, and thus can be regarded as a variational method.



Since the Galerkin method gives the best approximation to  $u^*$  in this norm in this norm; that means, it also minimizes this functional it also minimizes that functional this has got very important ramifications in mechanics why because in many mechanics problems there there exists what are known as variational principles. So, what these what do these variational principles do? They say that if you can set up a potential, which is some measure of the energy it is an it is a function, it is an energy functional right it is an energy functional. And if you minimize that potential the solution that you get is it you are going to recover the governing equations of that problem.

So, if you are solving a mechanical system, and you are interested in finding the equations of motion, the governing equations, which define that system, the one way to set up to get those governing equations is to write an expression for the energy for the energy functional for the energy in that system, for the energy potential of that system. And then if you set the first variation of that to 0, it is possible to come up with to arrive at the equilibrium equations of that system.

So, this is known as the variational method and what we has to show in here is that for self adjoint problems, the solution to the variational the variational method, which basically involves minimizing this functional  $u^T A u - 2 f^T u$ . If that functional represents the energy of the system represents a potential energy potential of that system,

then minimizing that functional that functional is equivalent to solving the Galerkin problem.

It is equivalent to solving the Galerkin problem right. So, what does this mean that for self adjoint linear operators. The Galerkin method is actually going to give you, it can be regarded as a variational method because it is going to give exactly the same solution, as the variational method which involves minimizing this potential right which involves minimizing this potential.

However, it is important to emphasize that the Galerkin method is much more it is much broader, it is much more powerful than the variational method because there are many classes of problems for which it is not possible to write down an energy potential like this. And then minimize that and come up with the governing equations, for instance for non conservative systems it is not that simple right it is not possible to write that, but for conservative systems it is possible to do that.

But for conservative systems it means that the Galerkin method and the variational method are going to give identical solutions, the Galerkin method can be thought of as a variational method in such a case. But for other problems for which it is not possible to write down a potential energy potential like this, it is not the Galerkin method can still be used. While the variational method cannot be used the simple variational method cannot be used straight away right. So, that is why the Galerkin method is much more in a sense it is much more general it is a much more general method. But for the simple case, that is when we have linear self adjoint, when we have self adjoint operators  $L$  then it is possible to do, so right to establish equivalents between these two methods.

(Refer Slide Time: 27:28)


**WEIGHTED RESIDUAL METHODS**

The Galerkin method is also an instance of a broader class of methods known as Weighted Residual Methods.

In the Weighted Residual Method, the residual  $f - \sum_{i=1}^n c_i L \psi_i$  is required to be orthogonal to function space  $H^m$ , not necessarily identical to the space  $H^n$ , spanned by  $\{\psi_j\}, j = 1 \dots n$

If the space  $H^m$  is spanned by orthogonal basis functions  $\psi'_j, j = 1, 2, \dots, n$ , this imposes the criteria  $(f - \sum_{i=1}^n c_i L \psi_i, \psi'_j) = 0$

$\psi'_j \in H^m, j = 1, 2, \dots, n$



The Galerkin method is also an instance of a broader class of methods known as weighted residual methods, recall in the Galerkin method what did we say? We say that  $f$  minus this residual has to be orthogonal to each and every  $\psi_j$  right, it has to be orthogonal to each and every, every basis function of that function space  $H^n$ . But if there is a broader the Galerkin methods, or a particular instance of a broader class of methods known as weighted residual methods. Then this weighted residual methods we do not require this to be orthogonal to the same function space to the basis functions of the same function space. But they can be they can be it requires orthogonality with some set of basis functions, but those basis functions can belong to a totally different space right when the for the class of weighted residual methods.

The residual  $f$  minus  $\sum_{i=1}^n c_i L \psi_i$  is required to be orthogonal to the function space  $H^m$ , which need not be necessarily the same as the function space  $H^n$  spanned by the  $\psi_j$ . So, I am constructing my approximate solution using basis functions  $\psi_j$  which belong to the function space  $H^n$  right, but when I am requiring orthogonality, I am saying that it is this residual is orthogonal to the basis functions of may be another function space  $H^m$  right. So, that is why Galerkin weighted residual methods are more general, but the Galerkin method requires that it would be orthogonal to the basis functions belonging to the same function space  $H^n$ .

If the function space  $H^n$  is spanned by orthogonal basis functions  $\psi_j$ ,  $j$  is equal to 1 to  $n$  this imposes the criteria that the residual be orthogonal to the basis functions of that function space. Where  $\psi_j$  belongs to  $H^n$ ,  $\psi_j$  does not anymore span  $H^n$  right it belongs to it spans  $H^n$ . So, this is the weighted residual method the Galerkin method, you can is a specialization of the weighted residual method it only says that it says that no a  $\psi_j$  cannot belong to any arbitrary space. It has to belong to the same space from which you construct, your trial solution from which you construct your trial solution  $c_i \psi_j$ .

(Refer Slide Time: 30:22)

**FINITE ELEMENT METHOD**

The Finite element method is a weighted residual method. There are two main variations of the finite element method - the Galerkin finite element method and the Petrov-Galerkin finite element method

In the Galerkin Finite element method, the spaces  $H^n$  and  $H'^n$  are taken to be identical - thus the Galerkin finite element method is identical to the Galerkin method, with the basis functions  $\{\psi_j\}, j = 1 \dots n$  specialized to satisfy certain special requirements that make them particularly useful for the solution of partial differential equations in solid and fluid mechanics

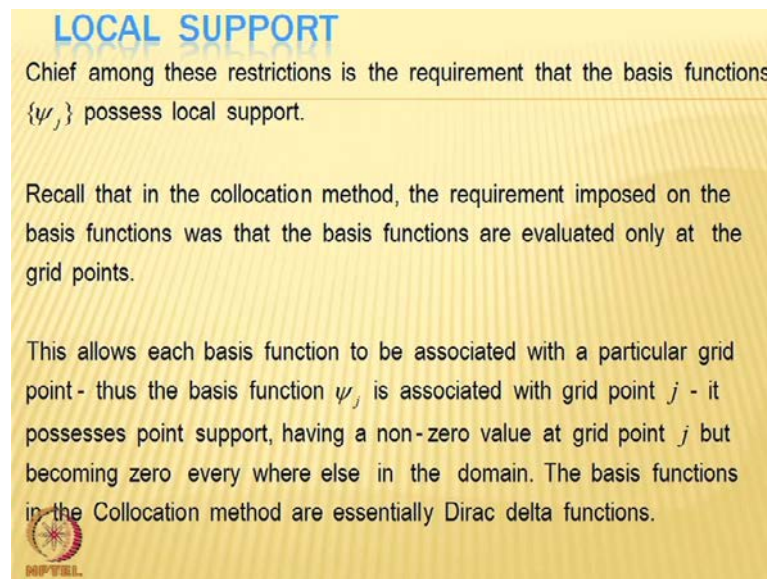
In the Petrov Galerkin finite element method, the function spaces  $H^n$  and  $H'^n$  are not the same, but the basis functions  $\{\psi_j\}$  and  $\{\psi'_j\}$  satisfy very similar requirements as in the Galerkin finite element method

The finite element method is actually a weighted residual method and there are two main variations of the finite element method. The Galerkin finite element method and the Petrov-Galerkin finite element method, both are weighted residual methods. In the Galerkin finite element method the spaces  $H^n$  and  $H'^n$  are one and the same because we just saw that the Galerkin finite element method requires that those two basis must be the same. But on the other hand the Petrov-Galerkin finite element method, says those function space is  $H^n$  and  $H'^n$  need not be the same. That is the basis function  $\psi_j$  and  $\psi'_j$  are not the same.

So that is the finite element method is basically a Galerkin method or a Petrov-Galerkin method, but it is not just any the why the reason why finite element methods are. So, popular is because the basis functions that that are used in this finite element methods are

really very, very simple functions. And they are they are specialized to solve particular problems and they can be specialized to solve different types of problems, but they are basically very simple functions, very simple polynomials right. And they have very nice properties, which makes the finite element method were popular and simple to use and can be solved used to solve a wide variety of problems.

(Refer Slide Time: 32:09)



**LOCAL SUPPORT**

Chief among these restrictions is the requirement that the basis functions  $\{\psi_j\}$  possess local support.

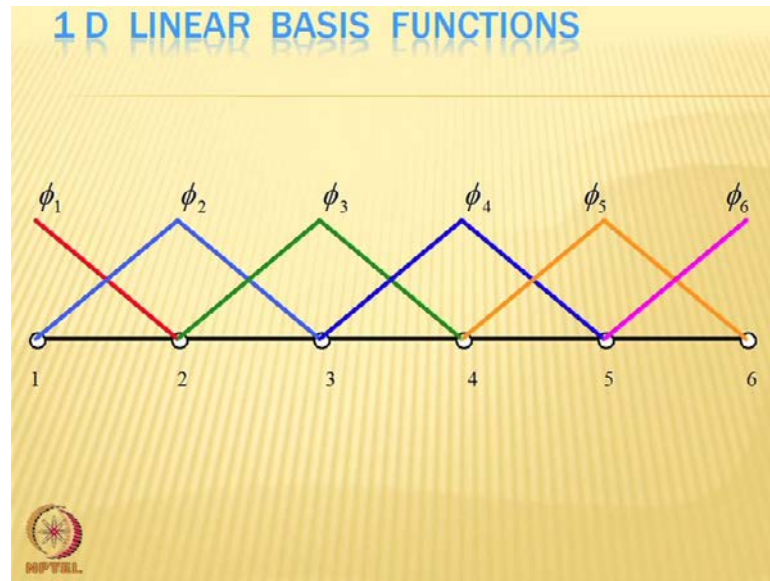
Recall that in the collocation method, the requirement imposed on the basis functions was that the basis functions are evaluated only at the grid points.

This allows each basis function to be associated with a particular grid point - thus the basis function  $\psi_j$  is associated with grid point  $j$  - it possesses point support, having a non-zero value at grid point  $j$  but becoming zero every where else in the domain. The basis functions in the Collocation method are essentially Dirac delta functions.

So, chief so the finite element method requires that these basis functions possess something known as local support, what do we mean by local support? Recall that in the collocation method which we discussed in the previous class, we require there impose the requirement on the basis functions that the basis functions are evaluated only at the grid points, this allows each basis function to be associated with a particular grid point. Thus a basis function basically each basis function is like a Dirac delta function it is one at a particular grid, basis function associated with a particular grid point  $j$  is one at that grid point  $j$ , but it is 0 at every other grid point. So, that those are those are the Dirac delta functions, which are used which are the basis functions of the collocation method.

So, in the finite element method the basis function they are slightly more general they say that no the basis functions are not just Dirac delta functions, they are not 0 everywhere else. And only one at a particular grid point with which it is associated, but they have local support what do we mean by local support it basically means something like this.

(Refer Slide Time: 33:28)



It basically means that my basis functions are they have a maximum at a grid point, the basis function  $\phi_2$  is associated with the grid point two it has a maximum value at that grid point. But then it linearly decreases as we move away from the grid point and by the time we reach another grid point a neighboring grid point to grid point two the basis function has become 0 right. So, it has got local support it is locally non-zero, but everywhere else the basis function associated with two is 0 at 3, 4, 5, 6, 1 everywhere else it is 0 at every other grid point it is non-zero at grid point two. But it is not just non-zero at grid point two it is non-zero at in a neighborhood of the grid point two, that is why it is said to have local support.



(Refer Slide Time: 34:26)

**FINITE ELEMENTS**

The finite element basis functions are similar - in that they too are associated with a grid point, however instead of having point support they have local support.

Thus they are non-zero in a neighbourhood of the grid point with which they are associated but are zero at every other grid point. The neighbourhood in which  $\psi_j$  associated with node  $j$  is non-zero naturally gives rise to the notion of "finite" elements - since this neighborhood is finite.

In the simplest possible representation, assuming linear basis functions in one dimension, the requirement that the basis function associated with a grid point be zero at every other grid point, gives rise to hat shaped basis functions as shown below.

So, basis functions the simplest possible the finite element basis functions, they are associated with the grid point. Instead of having a point support they have local support thus they are non-zero in a neighborhood of the grid point with which they are associated. But are 0 at every other grid point the neighborhood in which  $\psi_j$  is associated with node  $j$  is non-zero, the  $\psi_j$  is non-zero is it is a finite neighborhood and this gives rise to the idea of finite elements because the it is 0 in a finite region, t is 0 in a finite region. So, 0 in a well-defined region and that is why this method is called the finite element method. First of all finite number of basis functions and these are associated with local regions.

If a simplest possible representations as I have already given this is the this we can assume that these basis functions are linear basis functions that they are non-zero at a particular grid point they are 0 at every other grid point and they reduce linearly to 0 at the neighboring grid point. So, it is non-zero at this grid point and reduces linearly to 0 at the neighboring grid points and 0 everywhere else right at grid points, which do not share any which are not neighboring grid points to grid point two. It is identically 0 or in parts of the domain which are not which do not neighbor the grid point two they are identically zero.


(Refer Slide Time: 36:13)

**LOWER ORDER BASIS FUNCTIONS**

The linear basis functions shown in the figure are not really idealizations: such basis functions and their extension to two and three dimensions are commonly used in the finite element method

In fact one of the attractive features about this method is that it allows approximation using relatively lower order polynomial basis functions (eg. linear) while at the same time possessing the "best approximation property", common with all Galerkin methods.

Recall that the basis functions  $\psi_j \in H^n$  are being used to construct approximate solutions  $u = \sum_{j=1}^n c_j \psi_j$  which satisfy a partial differential equation with linear operator  $L$ , of the form  $Lu = f$



The linear basis functions are not really idealizations, such basis functions and their extensions to 2 and 3 dimensions are commonly used in the finite element method, why is this possible? And we will see that it is not immediately obvious that this very simple linear basis functions can be used to solve any useful problems. But we will see that it is indeed possible and it is fact one of the attractive features of this method is that we can use really simple basis functions to solve that problem.

Why can we use is really simple basis functions to solve this problem? Well we will see about that. So, basis the two things which make the finite element pattern popular is that first of all it posses the best approximation property. Since, it is a Galerkin method right it possesses the best approximation property, we know that in that norm defined with respect to that operator right the finite, the Galerkin solution, the finite element solution gives the best possible approximation to the true solution, that we are assured right the best approximation property of the finite element method, and the second thing which makes the finite element method. So, very popular and.

So, widely used is the these basis functions possess these very, very useful properties first of all they are relatively lower order polynomials number one, number two they have local support. So, basis function associated with the grid point is 0 at every other grid point, but it is not 0 it is not as restricted as the collocation method meaning, that it is not a Dirac delta function, but it is it has got local support.

Recall that the basis function  $\psi_j$  belonging to  $H^n$  are being used to construct basis. Being used to construct approximate solution  $u$  is equal to  $\sum_{j=1}^n \psi_j$ , which satisfies it is just restating the problem, which satisfies the partial differential equation with linear operator  $L$  of the form  $Lu = f$  subject to boundary conditions right.


(Refer Slide Time: 38:36)

**WEAK FORM**

Supposing that  $L$  is the Laplacian operator in 1D:  $\frac{\partial^2}{\partial x^2}$  then it is clear that the solutions  $u$  to  $\frac{\partial^2 u}{\partial x^2} = f$  must have non zero and non singular second derivatives.

If linear polynomials are used as basis functions for the above equation, it is clear that there is a problem. However finite element solutions regularly use linear basis functions to solve Laplace's equation.

This is possible because instead of directly solving  $Lu = f$  (referred to as the strong form of the equation), in the Finite element method, an equivalent weak form of the equation ( $\hat{L}u = f$ ) which involves less stringent continuity requirements is solved.



And suppose let us look at a particular example now. So, suppose we are interested in solving the Laplace equation right. So, the Laplace equation and we are interested in solving the simplest possible Laplace equation the Laplacian operator in 1 D. So, interested in solving  $\frac{\partial^2 u}{\partial x^2} = f$  in a certain domain and it is clear if we look at this equation that the solutions  $u$  to that equation must have non-zero and non-singular second derivatives, why?

Otherwise how are we going to solve this equation the  $u$  if the second derivative singular, how can if the function  $u$  has the second derivative which is singular how on earth can solve it, cannot solve this equation if provided that  $f$  is a reasonably well behaved function it is not possible for  $u$  to solve that equation. But we just claim that we can use linear polynomial basis functions to solve this problem, we can use linear polynomial basis functions to solve this problem why is that?

We know that linear polynomial basis functions then there is a problem right, since it is linear we cannot assure that the second derivatives is going to be non-zero and non-singular right? Then how is it possible to use linear basis functions of the finite element

method how can we use linear basis functions in the finite element method to solve this problem. Well the reason why we can do that is because instead of directly solving the problem  $Lu = f$ , which we refer to as the strong form of the equation in the finite element method, we solve an equivalent weak form of the equation.

How do we get the weak form, we will talk about that later. So, instead of solving this problem we solve this problem  $L^*u = f$ , where you note that the operator  $L^*$  is not the same as the operator  $L$ , it is not the same as the operator  $L$ . And if it turns out that the requirements  $L^*$  involves say first derivatives only, then it is clear that even if we use linear basis functions, it is possible to come up with a solution to that problem. Because it involves a less stringent continuity requirement, less stringent continuity requirement because it involves the linear operator  $L^*$  only involves first order partial derivatives right. How is that possible, well let us take a quick look.

(Refer Slide Time: 41:19)

**INTEGRATION BY PARTS**

This can be easily seen by considering the solution of  $Lu = \frac{\partial^2 u}{\partial x^2} = f$  in the interval  $x \in [0, l]$  subject to the boundary conditions  $u = 0$  at  $x = 0$  and  $\frac{\partial u}{\partial x} = 0$  at  $x = l$ .

Using Galerkin finite elements and considering weight functions  $w$  belonging to the same function space as  $u$ , from the method of weighted residuals :

$$\int_0^l w \left( \frac{\partial^2 u}{\partial x^2} - f \right) dx = 0$$

On integrating by parts:  $w \frac{\partial u}{\partial x} \Big|_0^l - \int_0^l \frac{\partial w}{\partial x} \frac{\partial u}{\partial x} dx = \int_0^l wf dx.$

So, if we consider this we are considering solutions to this in the interval  $x$  belonging to 0 to  $L$  subject to boundary conditions  $u$  is equal to 0 at  $x$  is equal to 0 and  $\frac{\partial u}{\partial x}$  is equal to 0 at  $x$  is equal to  $L$ . So, here the boundary condition is directly on  $u$  at  $x$  is equal to 0 the boundary condition is on the partial derivative of  $u$  at  $x$  is equal to  $L$  this type of boundary condition, as we have seen earlier is a Dirichlet type of boundary condition well that is a Neumann type of boundary condition.

So, using Galerkin finite elements and considering weight functions  $w$  belonging to the same function space as  $u$ , from the method of weighted residuals we have that equation, why we know that my  $L u$  minus  $f$ . This is basically the term within first brackets here is nothing but  $L u$  minus  $f$  that must be orthogonal to each of the basis functions  $w$ , right? If it is a pure Galerkin method  $w$  belongs to the same space as  $u$  if it is not  $w$  can be basis functions, which are totally different from the basis functions of the space to which the space, which is used to construct  $u$ .

So, this is my this is the requirement right this is the Galerkin this is the Galerkin requirement right. And this is basically the weighted residual requirement which becomes equal to the Galerkin Galerkin requirement, if the basis functions for  $w$  are the same as  $u$ . But for the time being let us think of it is a Galerkin requirement assuming that the basis functions for  $w$  are the same as the basis functions for  $u$ . So, then in 1 D if we integrate by parts, we get this right.

So, we integrate this first term  $w \frac{d}{dx} u \frac{d}{dx} x$  evaluated at  $l$  and zeros may take the difference, and then we have this part and this must be equal to integral of  $w f$  over  $0$  to  $l$ . If we expand this out what do we see? So, we have something like this  $w \frac{d}{dx} u \frac{d}{dx} x \Big|_0^l$  I can write it like that  $w \frac{d}{dx} u \frac{d}{dx} x \Big|_0^l$  minus  $w \frac{d}{dx} u \frac{d}{dx} x \Big|_l$ , I know this term is going to be automatically equal to  $0$ . Why is this term going to be automatically equal to  $0$ ? Because I know that this is my boundary condition, right  $\hat{x}$  is equal to  $L$   $\frac{d}{dx} u \frac{d}{dx} x$  is equal to  $0$ .

(Refer Slide Time: 44:14)

**REQUIREMENT ON WEIGHT FUNCTIONS**

If we require the weight functions  $w$  to not only belong to the same function space as  $u$  but in addition satisfy the essential boundary conditions on  $u$  homogeneously, i.e.  $w(x=0) = 0$  then from (\*):

$$w \frac{\partial u}{\partial x} \Big|_0^l = w(0) \frac{\partial u}{\partial x}(0) - w(l) \frac{\partial u}{\partial x}(l) = 0$$

and we are left with  $\int_0^l \frac{\partial w}{\partial x} \frac{\partial u}{\partial x} dx = \int_0^l -w f dx$

The integrand on the left involves only first derivatives. If we solve this equation (referred to as the weak form) rather than the strong form of the equation, then it is clear that the continuity requirements on  $u$ , and by extension  $w$ , are less stringent: now the requirement is merely that the first derivative of  $u$  exists

That means, that this term is going to be automatically equal to 0, but on top of that if I impose the restriction, that my weight functions. Which I am calling weight functions basically these are my these are the basis functions, with which with respect to which my residual has to be orthogonal. These are the basis functions with respect to which my with respect to which my residual has to be orthogonal. If I impose the condition that at the Dirichlet boundary the weight functions also satisfy the Dirichlet boundary condition homogeneously they become 0 at x is equal to 0.

So, if I choose my weight functions to satisfy that condition if I am choosing I am I am perfectly, but it is my weight function since this is a weighted residual method, right? I can do I can choose any set of functions from for my weight functions right. And if I make them I if I choose them such that they become 0 on the Dirichlet boundary then this term automatically become 0 as well because w 0 is equal to 0. So, this terms becomes this term becomes 0, this term is already 0 because del u del x at l is equal to 0 from my Neumann boundary condition. So, the first term here entirely vanishes. So, this term here becomes identically equal to 0 and I am left with this equation.

Now, let us look at that equation again the integrand on the left involves only first derivatives. So, if we solve this equation, rather than that equation, what is our requirement? Our requirement is that the first derivatives exists in this domain right that

these things  $\frac{\partial w}{\partial x}$  and  $\frac{\partial u}{\partial x}$  does not become singular, if we have to make sure that within the domain 0 to 1 these derivatives do not become singular.

So, if we solve this equation referred to as the weak form rather than the strong form of the equation, then it is clear that the continuity requirements have become less stringent right. So, it is no longer necessary for the second order derivatives the partial derivatives to exist, it is quite enough if the functions  $u$  and  $w$  possesses first order partial derivatives, which are non-singular in the interval 0 to 1. Because this continuity requirements have been relaxed that is why it is possible to solve this problem,  $L u$  is equal to  $\frac{\partial^2 u}{\partial x^2}$  is equal to  $f$  using piecewise linear basis functions. Rather than having to solve using having to solve using very quadratic functions right.

So, you can see this is the advantage of the finite element method, it first of all it is a Galerkin method which by virtue of it being a Galerkin method, ensures that my finite element solutions give me the best possible approximation with respect to the norm, with respect to the norm defined in terms of the operator  $L$  number one. Number two it allows these very, very useful very, very simple basis functions with less continuity requirements to be used right, we does not require very high order it could be one can solve complicated problems higher order partial differential equations with comparatively lower order polynomial basis functions.

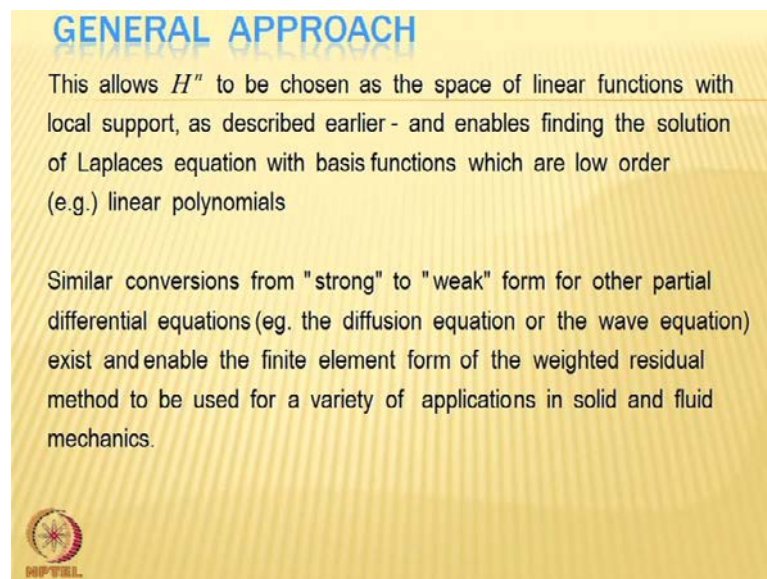
And these basis functions possess local support which also makes them attractive right we can associate each basis function, if we are solving a real problem and then the physical problem, we can associate each basis function with a small part of the domain right. It is not possible to integrate each integrate these carryout these integrals over huge, huge areas right. Because I know that these basis functions are 0 everywhere in non-zero in only locally right. And they have local support these basis functions have local support. So, they are non-zero locally.

So, my finite elements can be I can I can carry out my integrations over small parts of my domain right. So, my elements become truly finite right, they can I can make them as small as possible depending on how close I divide my how closely I place my grid points. I can make my domains as small as possible and then my domain of integration becomes small.

If necessary if I know that in a second part of the problem, I am not interested at there are there are or I am interested in a certain part of the problem, I am truly there are sharp gradients the solution where is a lot. I can make those regions with, I can define very small grid, I can define very closely spaced grid points in those regions. And then I can get more accurate solutions in those regions, while in regions where the solution is not of much interest. I can mesh it I can replace my grid points in as course the manner as possible and get some sort of approximate solution, but I am not really interested in that I am interested in the part that the solution shows sharp gradients as large variations.

So, it gives the lot of freedom basically it allows use of lower order polynomials, it assures the use assures that it is the best possible solution in that space and it gives. So, it is gives overall it gives a lot of freedom it is applicable to a very wide range of problems because and that is why it is so very attractive.


(Refer Slide Time: 50:40)



**GENERAL APPROACH**

This allows  $H^n$  to be chosen as the space of linear functions with local support, as described earlier - and enables finding the solution of Laplace's equation with basis functions which are low order (e.g.) linear polynomials

Similar conversions from "strong" to "weak" form for other partial differential equations (eg. the diffusion equation or the wave equation) exist and enable the finite element form of the weighted residual method to be used for a variety of applications in solid and fluid mechanics.



So, this allows the  $H^n$  to be chosen as the space of linear functions as I just said with local support, and enables finding the solution of Laplace's equations with basis functions, which are low ordered linear polynomials. Similar conversions from strong to weak form for other partial differential equations, for example, the diffusion equation and the wave equation exists, and enable the finite element form of the weighted residual method to be used for a very wide variety of applications and solid and fluid mechanics.



But it is important to understand that the reason why the finite element method works is because it is a Galerkin method, most importantly and why the Galerkin method is. So, very good is because it has this best approximation property right, it has this best approximation property. And more over for a particular class of problems for instance for self adjoint linear operators, self adjoint operators the Galerkin method and the variational method are identical.

So, if you can solve a problem using a Galerkin variational method you can you are guaranteed that you can solve it using the Galerkin method means you can solve it since you can solve it using a Galerkin method, you can solve it using the finite element method. So, we will end our lecture here next class we are going to start talking about integral operators.

Thank you.