Computational Fluid Dynamics

Prof. SreenivasJayanthi

Department of Chemical Engineering

Indian institute of Technology, Madras

Module No. # 05

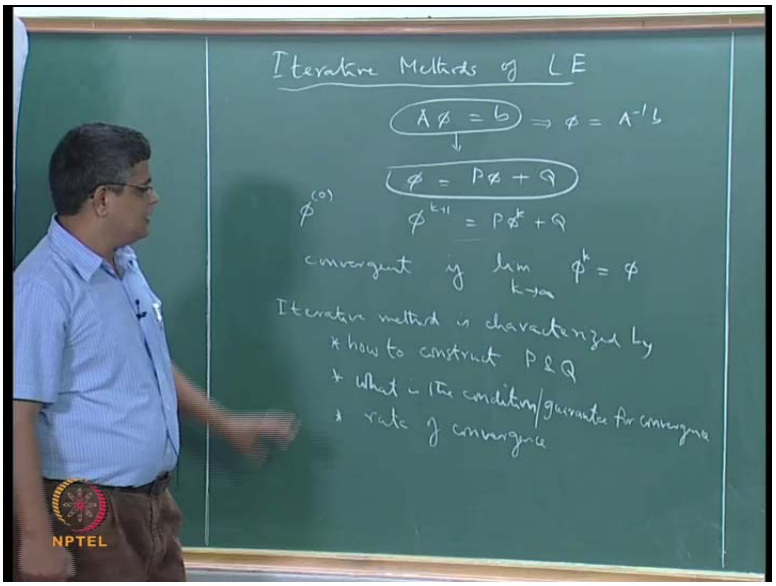
Lecture No. # 25

Basic iterative methods for linear algebraic equations

Description of point-Jacobi, Gauss-Seidel and SOR methods

Iterative methods have a very different way of the solution of the approach to the solution of A phi equal to b.We do not actually solve phi equal to b.

(Refer Slide Time: 00:21)



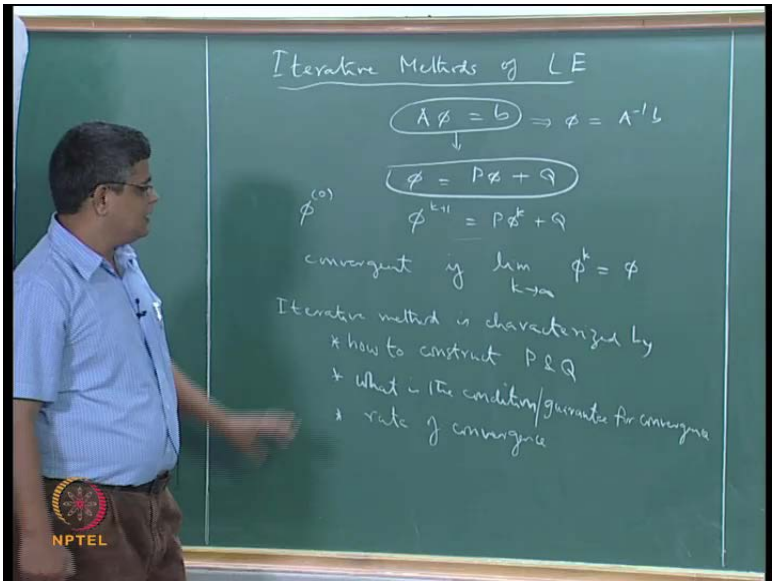
Let us just put a heading Iterative method for the solution of linear algebraic equations.Our basic equation is A phi equal to b.We do not try to say that therefore, phi is equal to A inverse b.We do not solve this equation at all, but you solve another equation at phi equal to P phi plus Q.

So, we write A phi equal to b in the form of phi equal to P time, phi equal to Q and we solve this by successive substitution also known as Picard's method.Therefore, we start with some initial guess for the phi and we write a recursive relation phi equal to k plus 1 is equal to P

phi k plus Q, where the k superscript here indicates the iteration number. So, we start with an earlier value of the phi that is the set of all variables that is the guess values and then we go to the new values.We generate a new guess value, new values from this rephrasing of the original equation and this successive substitution is said to be convergent, if limit as k tends to infinity phi k is equal to phi.

So, that is as the number of iterations increases that is with successive and more and more successive substitutions.If the phi k tends towards a fixed set of values phi, then we say that this iteration is convergent, is a convergent scheme and we we hope if the method is correct that the converged value is a solution to the original equation. So, this particular approach and the typical iterative method follows this way that from the given A phi equal to b, we generate we rephrase it as phi equal to P phi plus Q and from this we generate a recursive relation which enables us to go through from an initial guess, go through a series of ever improving values, which will converge towards the final solution.This method, any iterative method is characterized by three quantities.By how to construct P and Q from given A phi equal to b, what is the condition for convergence condition or guarantee for convergence and what is the rate of convergence?

(Refer Slide Time : 00:21)



So the idea is, it is not sufficient to be able to get P and Q from A.For example, we can write A equal to some P plus Q and then try to make an arbitrary division and, but the resulting

iterative scheme must be convergent. So, under what decomposition of A into P and Q will we have a convergent iterative scheme?

It is not sufficient to have a convergent iterative scheme.We would like to have good rate of convergence.If you are dealing with a linear equation, then the choice of the initial guess, the initial vector phi nought is not important, but if it is a non-linear equation that we are trying to solve,we have said that it is a linear equation, but we have non-linear equations as part of our discretization of our governing equation.

If they are not linearized, so if we are dealing with a non-linear equation, then the choice of the initial guess may also be important, but otherwise for linear equations,we it is sufficient to be able to construct P and Q from a given A and b.Also, we need to find out under what conditions this would converge. We should also importantly find out the rate of convergence, so that we have an estimate and also an assurance that we will soon be getting to the converged solution because an iterative method is a way of generating successive improvements and improvement is never final.It will go on till the end, until we are defeated by machine accuracy and things like that. So, in that sense the the method of approach for an iterative solution is different from the method of approach for a direct method and because we are getting an everimproving solution for a convergent scheme, provided we have that come up with a P and Q, which is which guarantees convergence, provided we have that; we are going to get an everimproving solution for phi ok.

So, in such a case, that means that we can stop the iteration at any time.We have already some estimate values for phi and this will be useful in the context of our CFD type of solution where we take account of the coupling between the X momentum equation, Y momentum equation, Z momentum equation and the pressure correction equation or pressure equation, continuity equation.

So, these kinds of couplings are done by effectively by another Picard type of solution, where you solve AU equal to b and then AV equal to b prime and then A double prime, W equal to Q like that. So, each time you solve for one of the variables assuming the values of the other variables. In that case, you are alreadystarting with guest values of V and W, when you solve the X momentum equation. So, you solve the X momentum equation using the iterative method until you get to some reasonable improved solution.You go to the Y momentum equation and solve for V and then you solve it until a reasonable improvement, you go to W

equation and once you get this,then you can come back to the U momentum equation and substitute the improved values for V and W and then go like that.

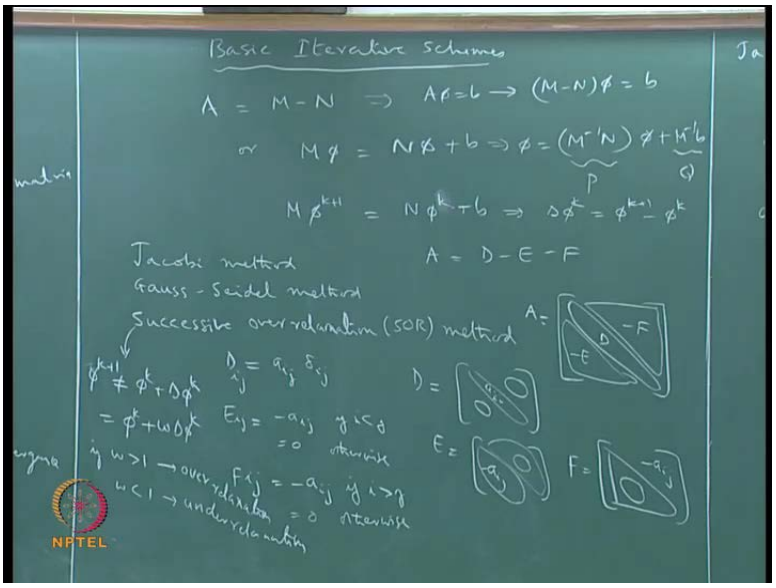
So, in the context of sequential solution of the X, Y momentum, equations and the continuity equations, an iterative method which provides us an improving an ever improving solution as we provide the number of equations is the good choice.For an iterative method to work,it is not sufficient to have a convergent scheme or one that has a good rate of convergence.It is also important that this construction of P and Q should be such that; that does not require too much of mathematical operations ok.

In addition to that, the P and Q should be such that the successive generation of solution that is starting with phi k to get phi k plus 1 should not also involve too many mathematical operations because we need to do this iteration.

So, a successive iterative method is one which enables an easy decomposition of A and Q, P and Q or easy construction of P and Q from given A and one which enables us to go from phi k to phi k plus 1 in a in a relatively easy way without too much of computational time and one which has a wide range of convergence, especially relevance to our type of problems and one which also has a reasonable or good rate of convergence ok.

So, there are Classical Iterative methods like the Jacobi method and Gauss-Seidel method and the Successive Over Relaxation method and we would like tothey share some of these features andwe would like to examine them in this context. So, if you if you take let us call this as basic iterative schemes.

(Refer Slide Time: 10:26)



In all the cases, A is written as M minus N. So, the consideration of P and Q is such that A is decomposed into the summation of M and N.

So, it is not a multiplicative decomposition like LU decomposition.It is the summation type of thing and this implies that A phi equal to b will get converted to M minus N phi equal to b or M phi equal to N phi plus b.From this, we can construct, we can also write this as phi equal to M inverse N phi plus M inverse b. So, when you put it this becomes RP and this becomes RQ. So, when you do decomposed M into M minus N, then we have it like this.Typically, we do not construct P like this and Q like this because that involves finding M in this.Instead of doing the M inverse and then doing the product of these and these like this, we go directly here and then we can say that M phi k plus 1 is equal to N phi k plus b.

So, in this **B**asic Iterative Methods, we do not construct M inverse and then we do not multiply this;we write it like this. So, that and the solution of this is also very easy and we will see how easy it is,so that not constructing M inverse and then putting it like this, is not a disadvantage.

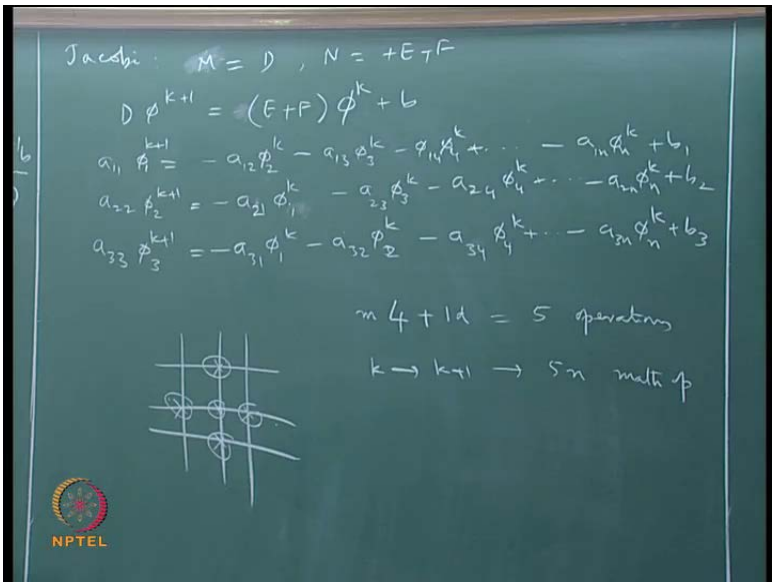
Now, what are this M and N?In the Jacobi method and also the Gauss-Seidel method and the Successive OverRelaxation; this is also known as SOR method.These are classical methods for iterative solution of a matrix equation.In these all the three cases, A is written as D minus E minus F and what is the what are the three here.If you take, this is the A matrix all the

diagonal elements together constitute D and all the lower triangular things except for the D constitute minus E and and all the upper triangular things except E will constitute minus F.

So, A is broken up into D minus E minus E.What this means is, that finding the breaking up the decomposition here is straightforward. So, you can write D to be equal to Eij equal to Aij delta ij. So, this identifies this delta ij is Kronecker delta when i is equal to 1,j equals to 1, then this equal to 1. So, this obviously identifies only the diagonal elements, all the rest will become 0.So, our D is all the Aii,this is 0, this is 0, this is our D and E is all the diagonal elements in upper triangular matrix is 0 and here it is minus A ij in this part. So, we can write E ij is equal to minus Aij, if i is less than j and is equal to 0 otherwise and F ij is minus A ij, if i is greater than j and is equal to 0 otherwise.

So, our F here is 0 here and this is minus Aij. So, this decomposition is straight forward.It is just minus 1 times this and even this we do not have to do this multiplication by minus ok.

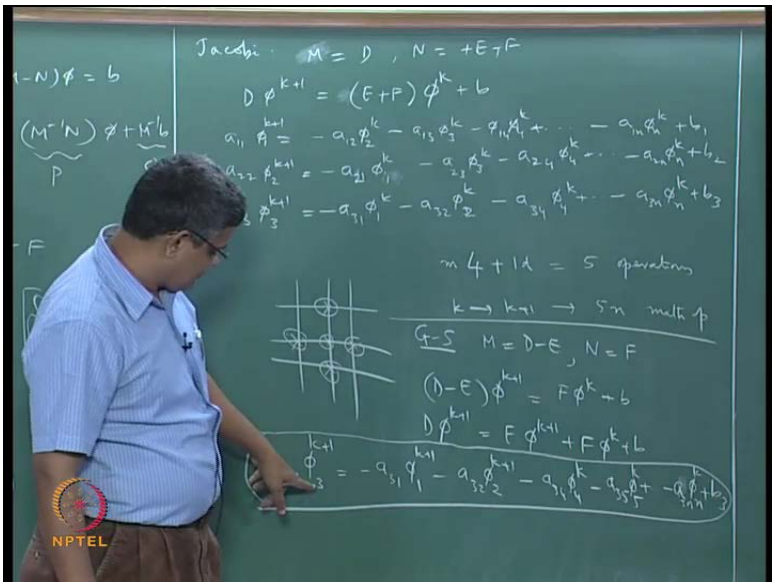
(Refer Slide Time: 16:25)



So, with this, the Jacobi method to be this basic thing, we can say in the Jacobi method, M is equal to b.This M here is equal to b and N is equal to minus E minus F,when we put it like this. So, N is equal to E plus F,so that when you put M minus N this becomes D minus E, D minus F and that is what and this makes the equation like D phi k plus 1 equal to minus we take it to the right hand side,so E plus F phi k plus b. So, this looks somewhat strange, but when you actually write down the equations, then what this one is saying is that this is a11

phi 1 is equal to, this we are taking it k plus 1 is equal to this is minus of this. So, this is minus a12 phi 2k minus a13 phi 3k minus a14 phi 4k minus a1 n phi nk plus b plus b1 ok.

(Refer Slide Time: 16:25)



Similarly, we have a22 phi 2k plus 1 is equal to minus a11 phi 1 phi 1 k minus a21 phi 1 k minus a23 phi 3 k minus a24 phi 4 k plusso on,minus a2 n phi n k plus b2 and so on.So, let us just put one more a33 phi 3 k plus 1 becomes minus a31 phi 1 k minus a32 phi 2 k minus a34 phi 4 k minus a3n phi n k plus b3. So, these are the equations that we are actually solving here. So, we can see that in order to get phi 1 k all we need to do is that we do these computations plus b and the whole thing is divided by a11.In order to generate phi 2 k plus 1, we do each of these multiplications, add to that b2 divide by a2 and again here, these multiplications and then divided by a3. So, if you want to generate phi 1 and phi 2 and phi 3 and so on 1 2 3 4 so many multiplications and one division for this and again 1 2 3 4, so many multiplications and one division for this.Now, the advantage especially in the context of the CFD solution is that, if this is an equation involved in all of the equations, but we know that in a discretized scheme only a few of these are non-zero. So, if you retain only those things that are non-zero, then you can write only these things very compactly,so that the number of multiplications here becomes very small.

We can take advantage of the sparseness of the matrix. So, typically in a two-dimensional flow kind of situationwe have we have if we imagine this is a point here,then this is

influenced by this point, this point, this point and this point. So, this is what is coming on the left hand side and these four are what are there on the right hand side.

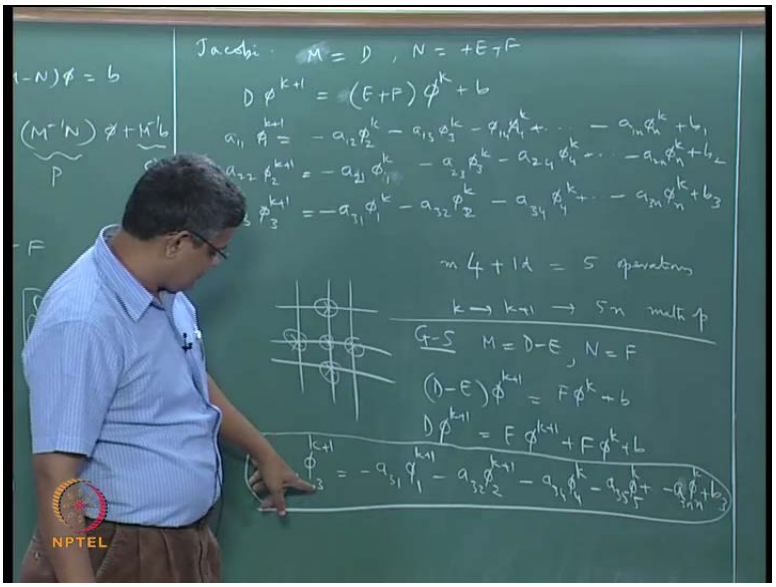
So, the solution for this at K plus 1 is given by 4 multiplications of these 4 coefficients with the current values, the estimates at the current values plusthe b1or b2 like that and division by this point here. So, that means that each evaluation of the 5k here, so it requires 4 multiplications and 1 division 4 multiplications and 1 division. So, that is 5 operations and we have n such equations because you have n variables, so that means to go from k to k plus 1,we need only 5n number of math operations. So, we can we have made the point that you want to go from k to k plus 1 very efficiently and we can see that for a typical two-dimensional flow with this kind of discretization and computational molecule, we require only five n number of mathematical operations to go to k to k plus 1.

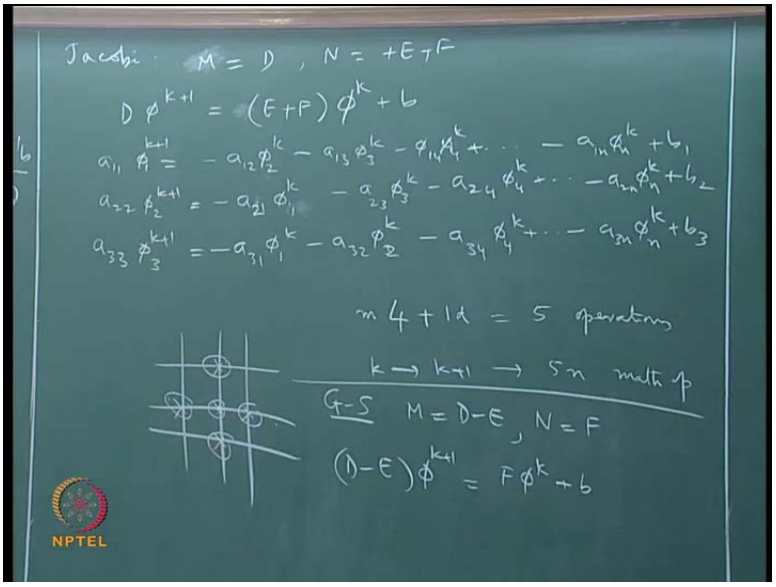
So, if it is a three-dimensional flow, wewill have two more terms two more things that becomes only seven times. So, to go from two-dimensions to three-dimension, it hardly takes another 2n number of operations, so it stays linear.This is as far as going from k to k plus 1 is concerned. So, we see that the implementation of the Jacobi scheme is such that the decomposition of a intoM and N is straightforward and the generation of phi k plus 1 from phi k is also very simple, because it requires only 5n number of mathematical conditions.

So, the next condition that we have to satisfy is,is it convergent?Under what conditions it is convergent?We will see that when we will do the convergence analysis and we will see if this is convergent if we have diagonal dominant for matrix A.The rate of convergence is also important and we can look at the rate of convergence from when we look at some ideal cases and from that we can get some kind of estimated cases ok.

Now, what about the Gauss-Seidel method? This is a Jacobi method, where M is equal to D and N is equal to E plus F.

(Refer Slide Time: 16:25)





In the case of Gauss-Seidel method, M is D minus E and N is F, so that this equation gets written as D minus E phi k plus 1 equal to F k plus b and we can write the same thing as D phi k plus 1 equal toE phi k plus 1 plus F phi k plus b.There is only a small difference between the Gauss-Seidel method and this method.When you write down an equation which is similar to this for example, if we write down the evaluation of a33 phi 3 k plus 1 in the Gauss-Seidel method, this will be equal to minus a31 phi 1 k plus 1 minus a32 phi 2 k plus 1 minus a33 phi 3 k plus 1 minus a34 phi 4 k plus 1 minus a35 phi 5k plus 1 minus a3n phi nn plus b.

So, the evaluation the implementation of the Gauss-Seidel method and Jacobi method,they look very similar.Here also you have to do 1 multiplication 2 multiplication third fourth five etcetera. So, if you have 5, 4 non-zero components coming in the right hand side, we have to do 4 multiplications and division by a33. So, in terms of number of mathematical operations,to go from k plus k plus 1 for a particular variable it is still 5 operations, 5 multiplications and divisions in 2 operations and 7 in the case of 3 dimensions and there are again n equations like that. So, to go to k to k plus 1, we again require 5n number of mathematical operations and n number of operations as in the Jacobi method.

So, even though we have made a different decomposition into M and N, it has not resulted in any increased number of mathematical operations per iteration. So, that is the generation of k plus 1 solution from k solution does not require additional number of mathematical operations.

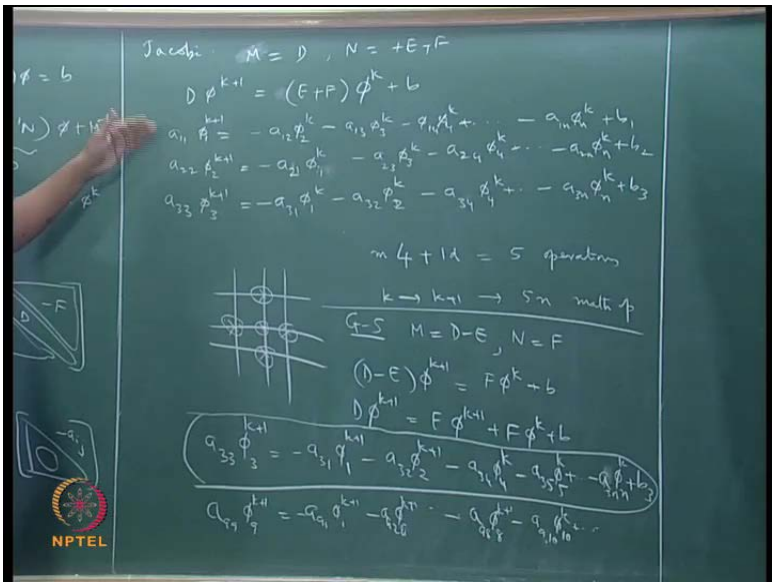
The difference that we see is that if you are solving the third variable, the previously computed variable that is phi 1 and phi 2 are approximated using the latest value that is by the time we in this sequence of solution we have completed phi 1 and phi 2 now we are at phi 3. So, we have already got improved estimate for phi 1 and phi 2 that is; we already have phi 1 k plus 1 and phi 2 k plus 1. So, we make use of those values here and for those succeeding ones that is from 4 2 n those values are still the old values, so we use them as k. So, that is in the Gauss-Seidel method, we make use of the improved values wherever they are relevant.

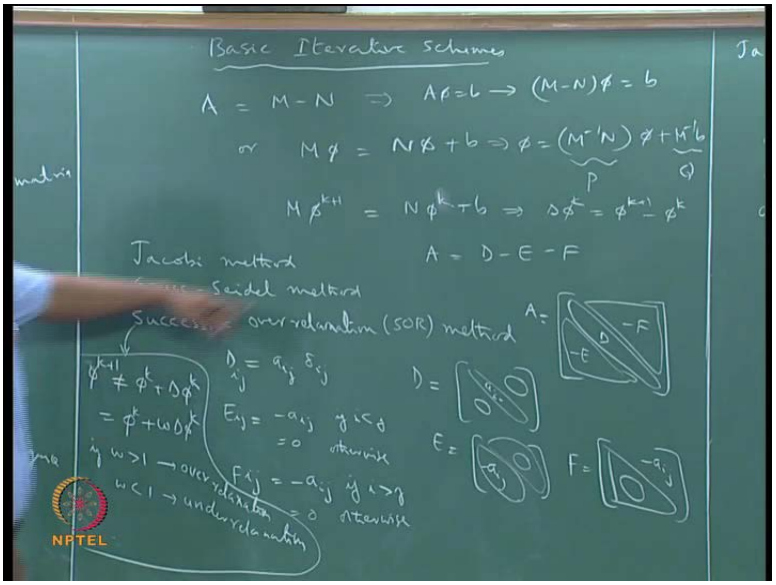
So, if you look at for example, the 9th variable a99 phi 9k plus 1 will be given by minus a 91 phi 1 k plus 1 minus a92 phi 2 k plus 1so on minus a98 phi 8 k plus 1 minus a99 n phi 10 k plus so on like this. So, up to 8 we make use of new values and after 9 that is from 10th onwards, we make use of whole values. So, in both the methods number of operations required to go from k to k plus 1 is about the same, but the actual number of solutions to go to the converged solution will be different for the Jacobi method and the Gauss-Seidel method because in the 2 cases, the P matrix that is the Iteration method,the Iteration matrix is different for the Jacobi method and the Gaussian method.Everything about this Iterative scheme that is; will it converge or not, is it a convergent scheme or not and is it what is the rate of convergence is determined by the characteristics of this Iteration matrix.Once this Iteration matrix changes, then these properties of the Iteration scheme may change.

So, it is not necessary that both these will be convergent for all cases and both these will have the same rate of convergence. So, we can expect different rate of convergence for these two methods and different condition of convergence for these two methods, but the end solution because if both of them are convergent, we expect the end solution to be the same irrespective what the initial solution is.That is the advantage of this. So, the two methods are very similar and very easy to implement.

In the case of the SOR method Successive Over Relaxation method, what we will do is that when you go from k to k plus 1, this you have in this Iterative method this gives you this. So, we have delta phi k, this is the improved value minus the old value. So, this is the improvement in the value of phi that you have predicted using this iterative scheme. So, in the SOR method you add more than the predicted improvement than in the case of over relaxation. So, you actually say that phi k plus 1 is not equal to phi k plus k delta phi k like this.You actually say that phi k plus some omega times delta phi k and if omega is greater than 1, then this is called overrelaxation and if omega is less than 1, then this is called under relaxation. So, this is the nature of this successive over relaxation.

(Refer Slide Time: 32:25)



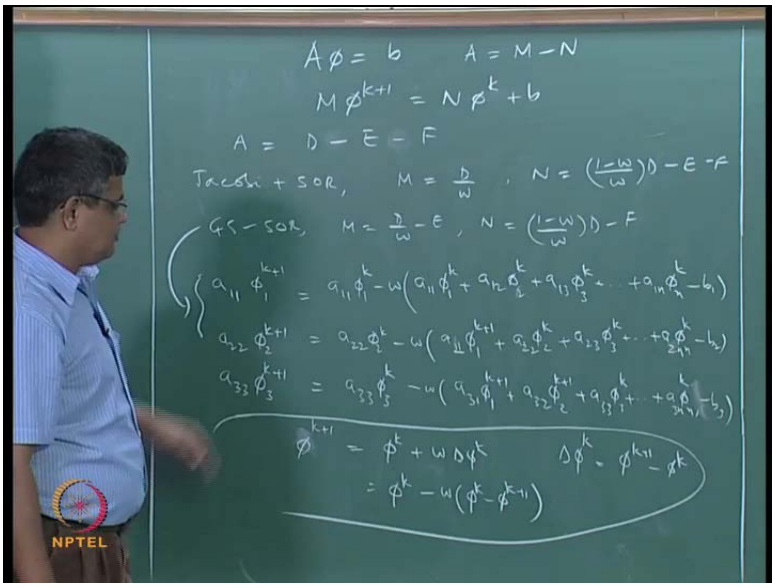


One point to note about this Jacobi method and the Gauss-Seidel method is, in the Jacobi method, we are making use of the old values for all the computations and here, we are incorporating the new values as and when they are available. So, what this means is that in the Jacobi method, the order in which these equations are solved does not matter because the phi k plus 1 is generated from known values of phi k.Whereas, in this particular case by the time you want to do phi 3, you have to know phi 1 and phi 3.Then in order to know phi 9 in order to phi 9 from k plus 1, we need to have solved phi 1, phi 2, phi 3,phi 4 up to phi 8 using this.

So, this is Sequential Successive Calculation whereas, this is a simultaneous solution. So, you can say that this is a Simultaneous Evaluation whereas, this is a Successive Evaluation and this Successive Over Relaxation comes from the application of the over relaxation parameter omega to this particular method ok.

So, typically when we say Successive OverRelaxation, we talk about the Successive Over Relaxation or over relaxation of the Gauss-Seidel method. So, this is also this is GSSOR effectively it is equivalent to. So, let us see what form the SOR method takes in terms implementation.

(Refer Slide Time: 34:05)



We are solving A phi equal to b, using an Iterative method, we write M minus N and construct an Iterative scheme like M phi k plus 1 equal to N phi k plus b and then the SOR method is, this is the general decomposition of the diagonal elements and lower triangular elements except the diagonal upper triangular elements except the diagonal limits here.This is the standard decomposition; we take D only in the case of M,in the case of in the case ofJacobi method and D minus E in the case of Gauss-Seidel method.

If you are writing a Jacobi method plus SOR, then in that case M is written as b by omega, where omega is the Over Relaxation Parameter which has a value between 0 and 2 and N will be 1 minus omega times omega times D minus E minus F.In the case of Gauss-Seidel SOR, M is D by omega minus E and N is 1 minus omega by omega times D minus E. So, that when we now put M minus N, then this D by W here, that is cancelled like this and we get omega by omega times D. So, everything becomes the same as this.

So, M minus N here is still the same as this one, except the fact that in the implementation part here, we are putting the value of omega which is between 0 and 2. So, if we write like this we can substitute these into the corresponding equations, but ultimately we can show that in terms of the computation of phi 1k plus 1 is written as a11 phi 1k minus omega times a11 phi 1k plus a12 phi 1k phi 2k plus a13 phi 3k plus a1 and phi nk minus b1 and a22 phi 2k plus 1, we are looking at the Gauss-Seidel based SOR.In this case,this is a22 phi 2k minus

omega times a11 phi 1k plus 1 plus this will be a21 and a22 phi 2k plus a23 phi 3k plusso on upto a2 nphi nk minus b2.We will just write one more before we write this as a general expression.This is a33 phi 3k minus omega times a31 phi 1k plus 1.

We already have, by the time we come to the third row third variable, we already have phi 1 at k plus 1 and we also have phi 2 at k plus 1 plus a33 phi 3k plus so on upto a3n phi nk minus a3. So, this is this is how we can do this.

Is this equivalent to the SOR method?If we put, let us just check the limits.If we put w equal to 1, there is a this thing becomes D minus E and this thing becomes 0,so N becomes minus F. So, when you put in this SOR Gauss-Seidel based SOR method, if you put w equal to omega equal to 1, then you recover the Gauss-Seidel method. Now, what will happen to this if you put omega equal to 1 here, this and this cancel out and this becomes b1 minus all these things and here, this will cancel out with this and this becomes b2 minus all these things and that is exactly the way that we compute these variables using the Gauss-Seidel method.

So, in that sense this 1 when you put omega equal to 1 here, then this goes back to the standard Gauss-Seidel method.Now, when you put, we have also said that in an SOR method, we write xk phi k plus 1 equal to phi k plus omega times delta phi k and this is actually the form that we have used here and we know that delta phi k is equal to phi k plus 1 minus phi k.

So, if we substitute that here, we get phi k minus of phi k minus phi k plus 1. So, this phi k plus 1 is minus phi k plus 1 for example, is this whole thing and this is the phi k that we have here and this is the minus omega that we have here. So, in that sense this whole expression within the brackets is phi k minus phi k minus 1 that is equal to minus of delta phi k and then we have the minus omega parameter here. So, this particular way of writing is consistent with the definition of the SOR method, although it is a bit hidden when you express it like this.

Now, why do we write it like this because, in terms of computation; if we look at these things the number of additional computation that we have to do is very little.We have an omega, which is multiplying all these things and anyway, we are doing all these things.It is only that a11 times this, but this anyways is going to be divided by this.

So, this equation can be written as phi 1k plus 1 equal to phi 1k that is when this divides this, you get 1. So, you do not need this multiplication; it is only one additional multiplication of omegathat is required.

So, implementing the SOR method into the Gauss-Seidel method does not increase the computational terms significantly to go from k to k plus 1. So, that is why we have written it in this nice form, so that we can take advantage of limited number of computations that we have to do. So, the point that we want to make here is the number of mathematical operations in terms of multiplications and divisions for Gauss-Seidel method and SOR method or Jacobi method and the corresponding SOR method is virtually the same.

What makes the difference is, as you go from k to k plus 1, you have the same number of operations, but how many such iterations do we have to get?We have to go from k plus 1 to k plus 2 and then we put this back here,we go through this loop. So, how many times we go through this loop depends on what is the rate of convergence and that depends on specifically the parameter omega.

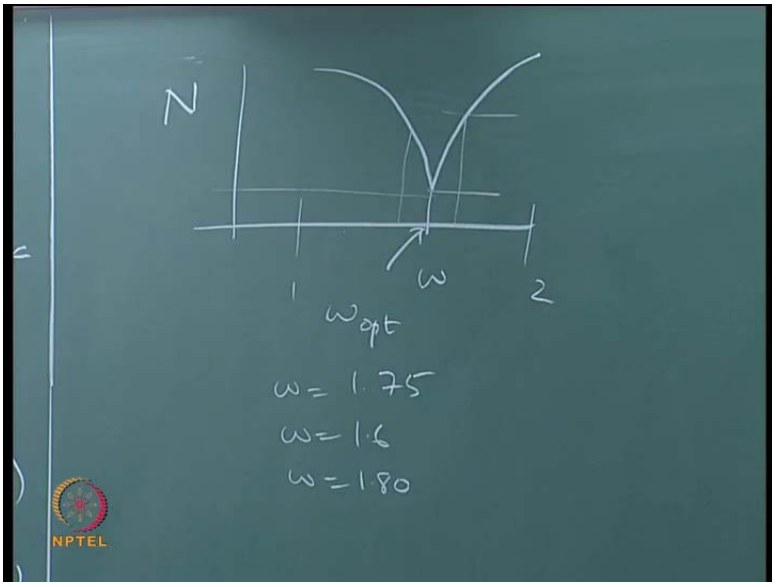
So, we chose the parameter omega such that the number of these successive approximations that we have to make becomes lesser and lesser,so that we can gain an overall advantage in terms of computation of the solution for A phi equal to b. So, the SOR method will work if the number of iterations, number of successive approximations is decreased and that by how many decreases depends on the choice of the omega parameter.

So, that is why one has to use the optimal value of the omega parameter and at this stage before we do anything formal, we can say that we have seen that the Jacobi method and SOR method and the Gauss-Seidel method,the standard Jacobi and Gauss-Seidel have a number of mathematical operations, capital N for a solution of the order of n square is what we have mentionedand SOR method with the optimal parameter of omega.

So, that is that parameter that value of omega here, which will reduce the number of these operations to the minimum that is necessary in order to achieve an accuracy desired accuracy of evaluation of these is such that this will vary as 1.5.The exponent is reduced from 2 to 1.5, then small n is large when you are dealing with large number of equations, this change in exponent from 2 to 1.5 can be humungous.

For example, if you say that is N is 10 to the power of 4 that is 10000 equations, which is not much, then SOR gives you a fact of 10 advantages in terms of computational term as compared to the standard Jacobi method and Gauss-Seidel method. So, an order of magnitude increase in the speed of computation is possible if we use SOR method, but this is true only for optimal values of omega.

(Refer Slide Time: 46:07)



If we are sub-optimal, as we go as we increase the, let us just make a schematic, N is a total number ofmathematical operations to get a solution and omega is a parameter here, typically N may go down like this, may be between 1 and 2 is typically the over relaxation range and at the optimal, so this is the optimal value of the parameter.

So, at the optimal value we have a large decrease in the number of mathematical operations, which are required and if you are this is typically steep,so that if you are sub-optimal, ifyou took a value of omega either like this or here you would have much higher value of the number of operations.What is important is that we have to specify what the value of omega is in order for this scheme to work. So, the Parameter Omega, the Over Relaxation Parameter is an input to the scheme and we have to choose it for some theoretical cases.We can determine apriori what the optimal value of omega is, but for the general case, when you are talking about a general A phi equal to b type of solution, then we do not know the value of omega. It can be anywhere between 1 and 2 and it there may be a difference between omega at 1.7 and omega of 1.6 and omega of 1.8.

There may be a difference between 1.75 and 1.80.So, how steep this valley is depends on the specific case that we are looking at and the more the number of grid points, the more the number of N that we have typically, the closer towards 2 that it goes and the steeper the value of this valley is, the shape of the valley is.

So, when we are dealing with a large number of parameters large numbers of equations, then the determination of the optimal value is not a triviality.One has to do,one has to search for it,one has to implement a method of looking for the optimal value into the scheme of solution itself andthat will take extra time, but in many cases that extra effort to find the optimal value is very well worth the effort because this variation is quite steep. So, you could have a factor of 2 or a factor of 3 in terms of the speed of computation when you go to the optimal value as compared to some things which is suboptimal ok.

So, this is something that we have to keep in mind when we are trying to implement the SOR method.