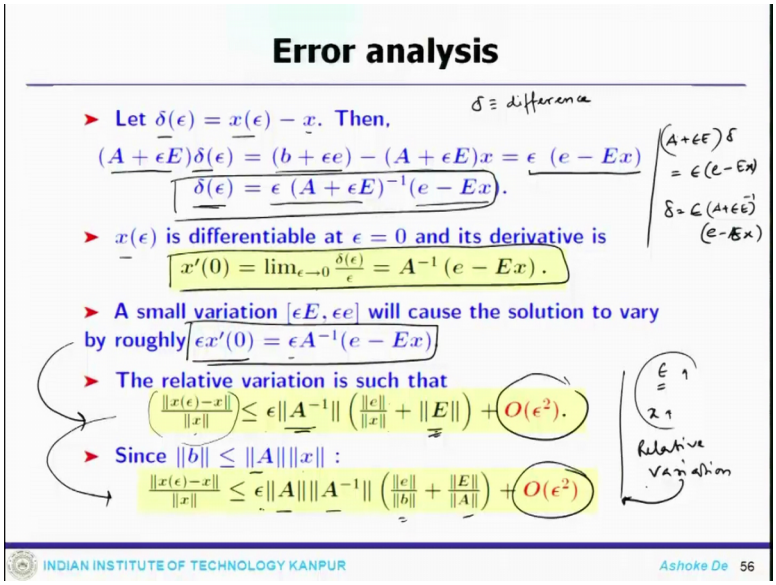
Introduction to Finite Volume Methods - I
Prof. Ashoke De
Department of Aerospace Engineering
Indian Institute of Technology, Kanpur

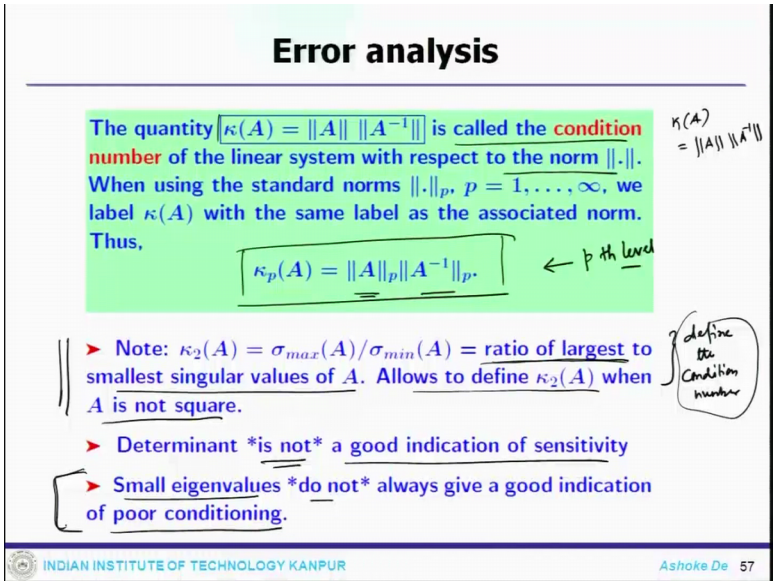
Lecture – 40
Error Analysis-III

(Refer Slide Time: 00:13)



So, welcome to this particular lecture and we will continue our discussion what we have been doing so far. So, these are going to be important, when you actually use the iterative methods to find out the error. Because, the error needs to be I mean decreasing when you go ahead with the iterative process.

(Refer Slide Time: 00:40)

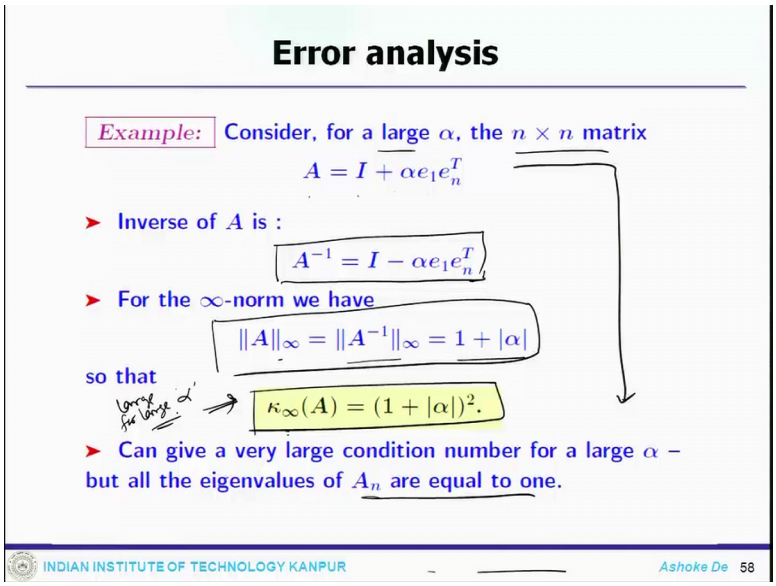


Now, there is a quantity, which you have come across is called the condition number; it is kappa A, it is essentially represented at kappa A is the condition number of a matrix A is the magnitude or norm of A into A inverse. So, of a linear system with respect to norm this.

So, when you use the standard norms the norms, 1, 2, 3 to infinity, we say this kappa A with the same level as p condition number for the pth norm of A can be written as magnitude of A at the pth norm and inverse at the pth level. So, this p could, it is a pth level, so p can go from 1, 2 n like that. Now, one can note here, the kappa 2 is the ratio of the largest to a smallest singular values of A. So, it allows to define when A is not square. So, this is a very important note, when A is a not a square matrix, you can actually define the condition number. And as you see this is the ratio of the largest to the smallest singular values of A.

So, which tells you depending on the condition number, the sensitivity of this matrix for the linear solution because, that is what is getting associated with the condition number. So, from the condition number also one can anticipate the convergence level of your linear system. So, when you say the determinant is not good, a good indication of sensitivity. Small eigenvalues do not always give a good indication of poor conditioning. So, that is the point, once you find out the eigenvalues from there, if you try to anticipate whether, the conditioning is good or bad that is not an good indication, but the determinant can be a good indication for getting the condition number.

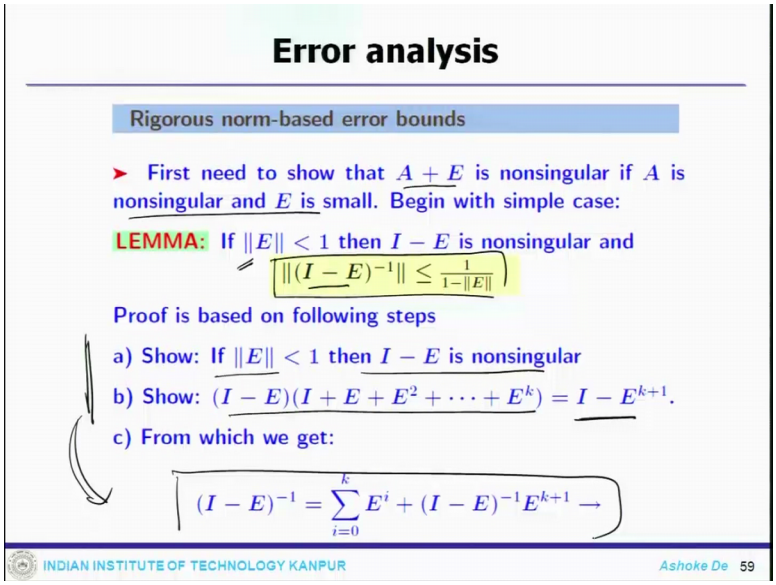
(Refer Slide Time: 03:14)



One can again check an example, simple example and try to see what it means. You take an large alpha for n by n matrix such that, equals to I plus alpha into epsilon 1 and epsilon 1 n transpose. So, the inverse of A if you find out, it will remain I minus alpha e 1 e to the power transpose n. For the infinity norm we have A infinity norm equals to A inverse infinity norm equals to 1 plus mod alpha so, that is what you get. So, this leads to the calculation of the condition number at the K infinity level which is 1 plus mod alpha and completely square of that.

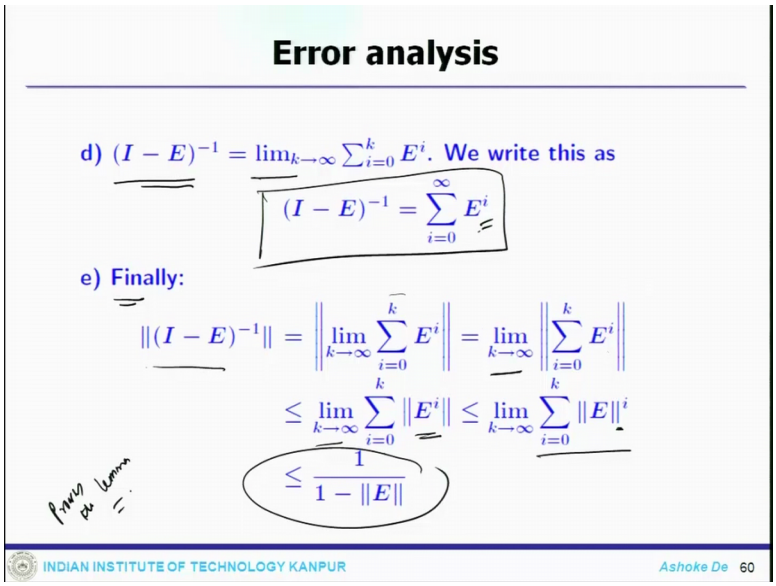
So, that means, it can provide you back a very large condition number for large value of alpha, but all the eigenvalues of A n are equal to 1. So, this is the check that, what we made a statement here the eigenvalues are not the good indicator for condition number, but once you look at here, the eigenvalues are nicely behaving, but the condition number is absolutely large for large alpha so, that is not a good marker for condition number.

(Refer Slide Time: 05:02)



So, now if you look at more detailed error bound, so one needs to see what happens to A plus E, when it is non singular, if A is non singular and E is small. So, we can take an simple case as in lemma, the magnitude or the mod of E less than 1 then, I minus E is non singular and it satisfied the norm 1 minus E inverse. So, magnitude of that less than equals to 1 by 1 minus E bound. So, this can be proved based on this kind of theorems. So, you have first you can show if mod E than 1 then E is non singular and second step you show I minus E and expand this to the kth level which is identity matrix minus E to the power k 1, then you can from their combining these two you obtain this.

(Refer Slide Time: 06:12)

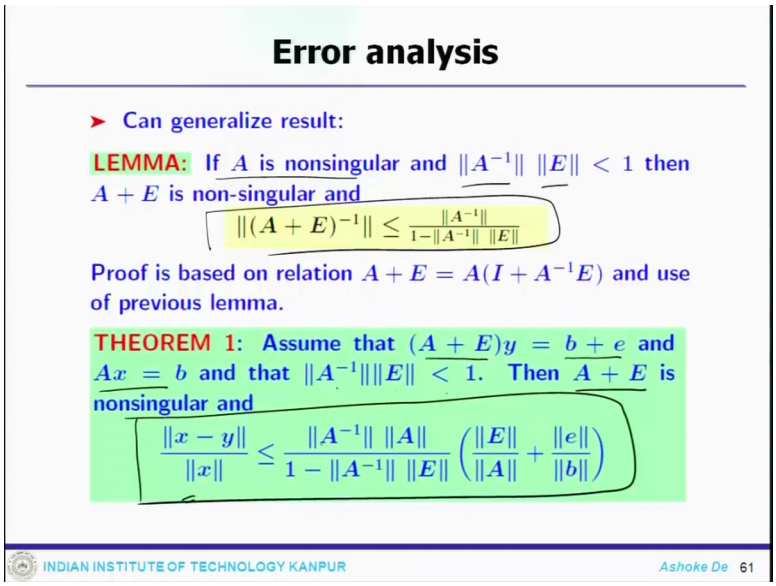


So, one can actually get and then once you obtain this take the limit at k tends to 0, we get 1 minus E inverse, which is i to the power 0 infinity the smallest value.

So, the lemma says identity minus E inverse mod should be less than equal to 1 by E. First we showed that that, E is small so, magnitude of that less than 1, 1 minus E is not singular that is the first part. Second part we show that, 1 minus E the power inverse. So, basically you have expanded this term and you get 1 I minus E to the power K plus 1 and when you take the limit this becomes the sum of E to the power I.

And the final step, the magnitude of that equals to limit K tends to infinity, i 0 to k E i and the limitic sense of k tends to infinity one can write i 0 to k E to the power chi then, which is less than equals to k tends to infinity i 0 to k magnitude of E, less than equals to E mod to the power i, which is nothing, but 1 by 1 minus E so, the it prove the lemma. So, that sense sort of in very strong error bound in terms of computation.

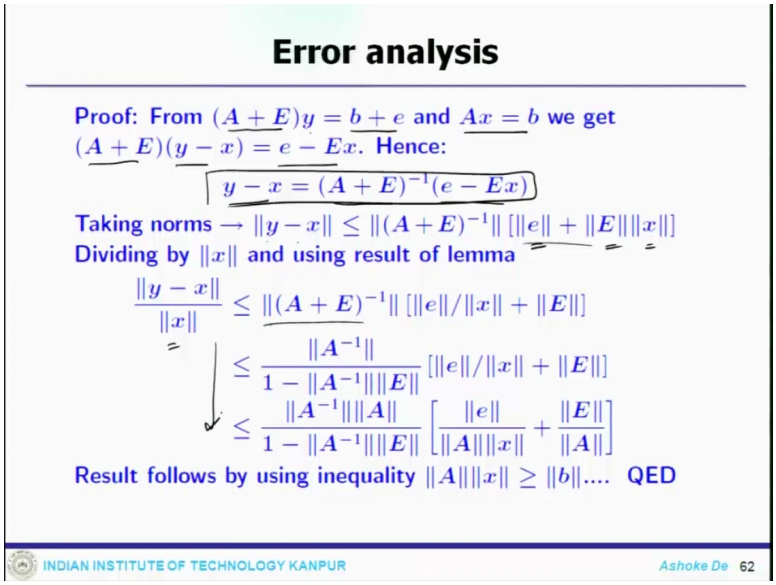
(Refer Slide Time: 07:49)



So, the thing is that if we can generalise that lemma if A is a non singular and A magnitude of A inverse and norm of E less than 1 then A plus a is also non singular and that satisfied this so, this one can also prove based on the previous one. So, this can be similarly proved like that. So, the theorem says assume A plus E y equals to b plus epsilon and A x equals to b and that norm of a inverse or magnitude of E less than 1 then A plus E is non singular and it also satisfied magnitude of x minus y divided by x. So,

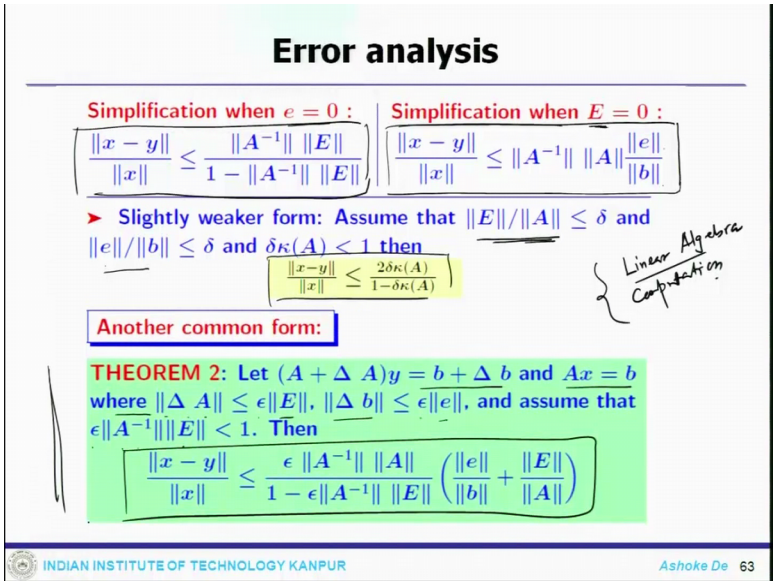
this is the normalisation then, you get magnitude of A inverse magnitude of A by 1 minus A inverse E E by A epsilon by b.

(Refer Slide Time: 09:02)



So, how you prove that? so you consider A plus E y equals to b plus epsilon and A x equals to b So, what we get when you write A plus E y minus x, equals to e minus Ex y minus x equals to A plus E inverse e minus E x so, that is what you get for this. Now, taking the norm y minus x the norm of that must be less than the inequality says A plus E inverse the norm of that and norm of E norm of E norm of x. Now, you divide that by the norm of x, so the inequalities stands and the complete 1 is divided by x and if you do the algebra you can prove that, that this happens.

(Refer Slide Time: 10:00)

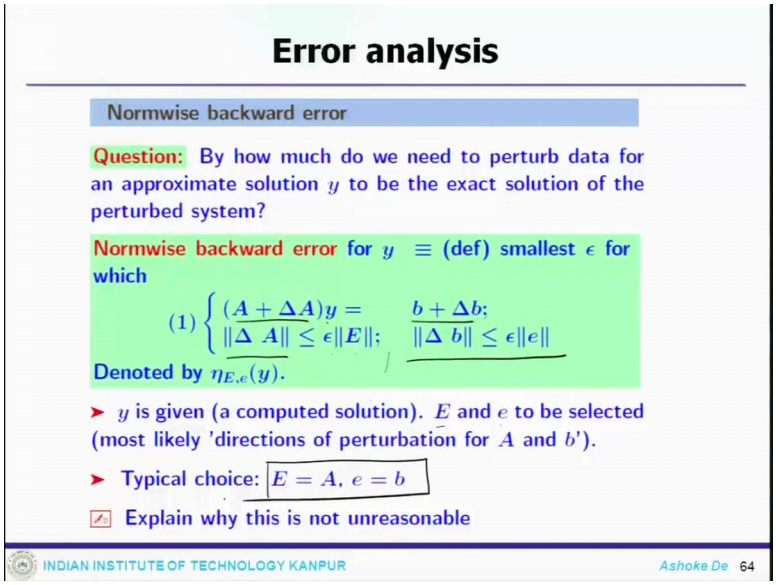


Now, one can do some sort of an simplification to this kind of system. One immediate simplification is that, you say E is 0, so then norm of x minus y divided by norm of x less than equals to this, so it boils down to a another inequality which is A. So, from the generic case, you can always simplify for the special cases and simplification when capital E 0 x norm x minus y by norm x less than A inverse A e by b. So, that is simplification, which you can do.

Now, if you write in shorten or slightly weaker form and assuming that, norm of E by norm A less than equal to delta and norm of epsilon by b or e by b less than delta then I can write, norm of x minus y divided by norm of x less than 2 delta condition number a by 1 minus delta condition number A. So, it can be also written in another form which leads to second theorem. So, that these are the part where you can find in the any linear algebra calculation or book with linear algebra computation. So, that provides those information very nicely.

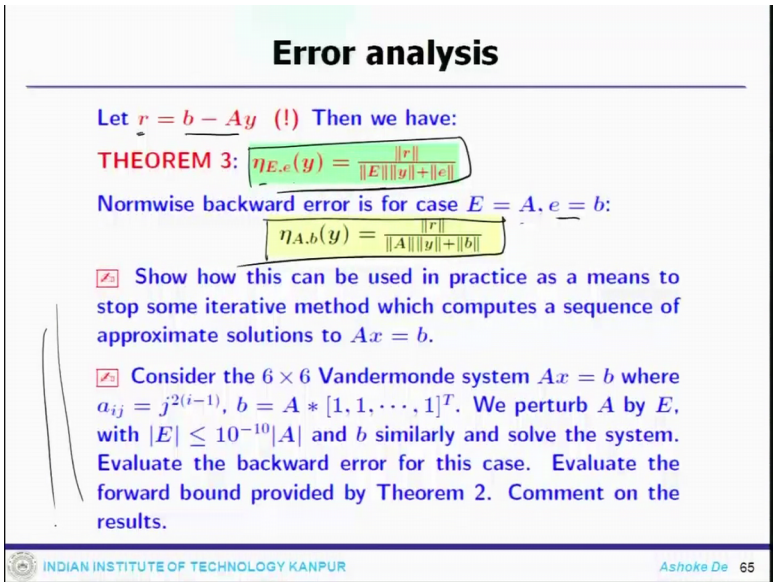
So, let A plus delta A y equals to b plus delta b and A x equals to b where norm of delta A less than equals to epsilon norm of E and norm of delta b less than equals to epsilon into e and assume that epsilon norm of A inverse norm of E less than 1 then it satisfied this inequality. These are hard core matrix computing theorem, which typically if one wants to have a very efficient algorithm for the linear system, he needs to keep checking all these details and then put the norm.

(Refer Slide Time: 12:33)



Now, backward norm if you put, so it can tell you how much we can need to perturb the data, so, that we can still exert solution of the perturb system. So, backward error of y is smallest epsilon for which, A plus delta A y equals to b plus delta b and norm of delta a less than equals to delta E, where delta b less than equals to epsilon e. So, y is given at the computed solution and E and small e to be the selected for the perturbation and typical choice is that E is A and small e is b.

(Refer Slide Time: 13:29)

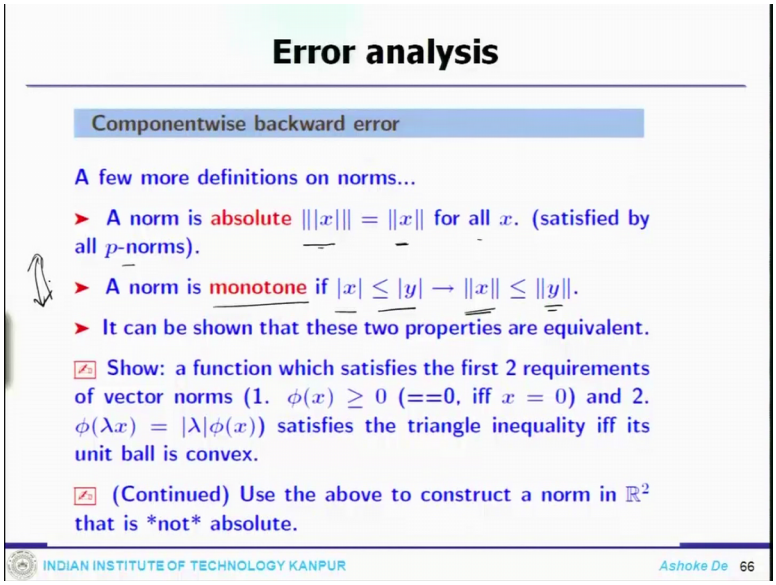


Why that is not an unreasonable choice, let us see, you define r equals to b minus A y then we have eta E e y equals to norm of r by E y norm and epsilon norm. So, norm wise

backward error is for case E is equal to A and small e equals to b. So, that shows this expression.

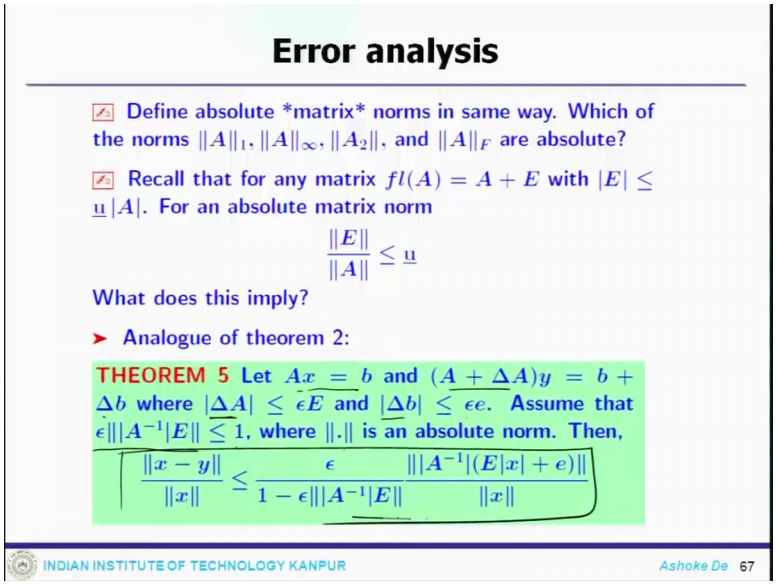
So, one can so, you can find out by considering some system and then estimate this kind of norm or one can take some examples and put it some code in the MATLAB and find out that.

(Refer Slide Time: 14:15)



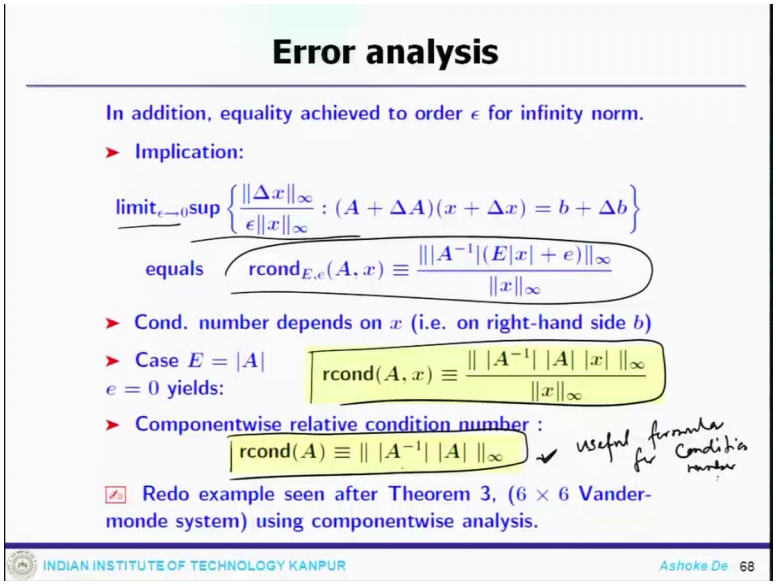
Now, if you go by component wise then, some more definitions which would be handy to have. The absolute norm equals to the norm of that, for all x which satisfied the p norm. Secondly, the norm is also monotone, if mod x less than equal to mod y which leads to the norm of x less than equals to norm of y. One can show that, these properties are there are some equivalents to that. So, you can show easily these things.

(Refer Slide Time: 15:00)



Now, there are another theorem, which is theorem number 5, which is analogous to theorem 2, which state that, if you have A x equals to b and A plus delta A y equals to b plus delta b, where delta A or the mod of delta A less than equals to A epsilon E or mode of delta b less than equals to epsilon e, assume that epsilon a inverse this magnitude less than equal to 1, where dot is an absolute norm. Then, the linear raised definition would be x y norm, x minus y norm divided x norm less than equals to this expression. So, one can also prove this.

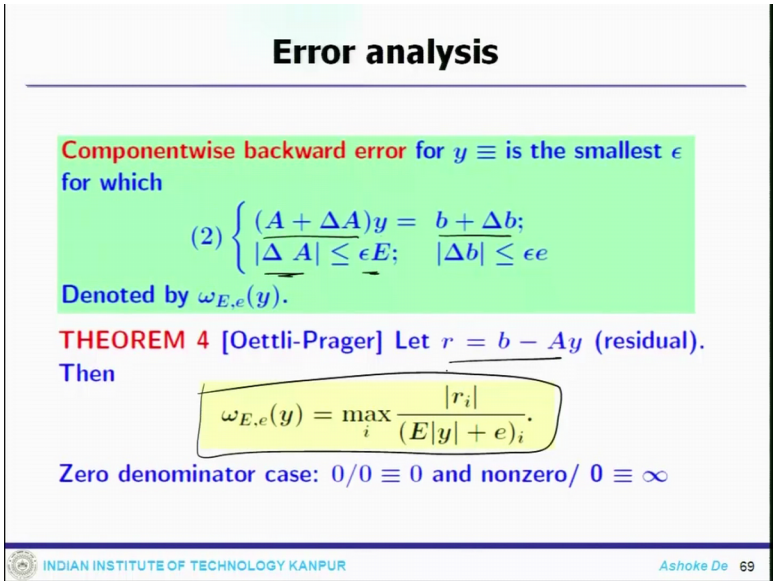
(Refer Slide Time: 15:53)



Now, in addition to that, finding that infinity norm, the implication is that if you take the limit of that whole theorem you get back this condition system. And the condition

number, recondition number depends on x and all these things and the relative condition number would be define that A inverse magnitude, A magnitude and the norm of that at the infinity level. So, this would be an very useful formula for condition number, and where you need to take care of all these while you will be programming your linear solver.

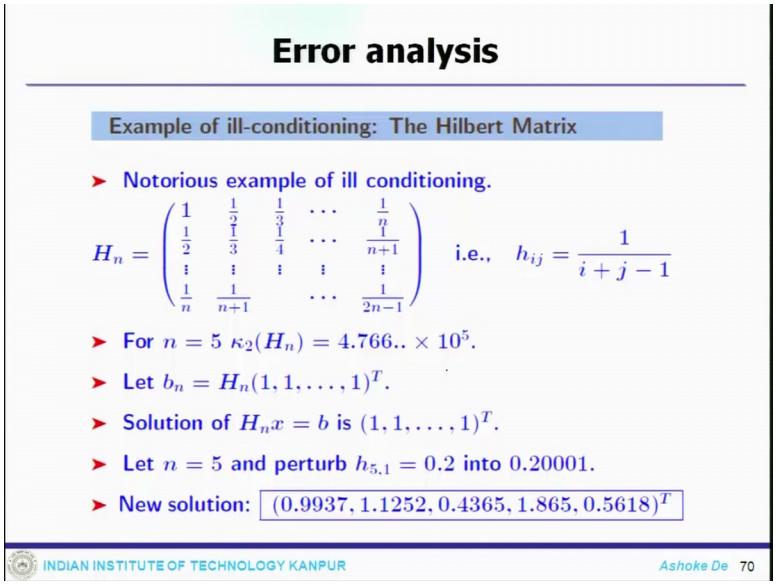
(Refer Slide Time: 16:45)



And the final component twice backward error for y equivalent is the smallest epsilon then, A plus delta y, y equals to b plus delta b and the magnitude less than this.

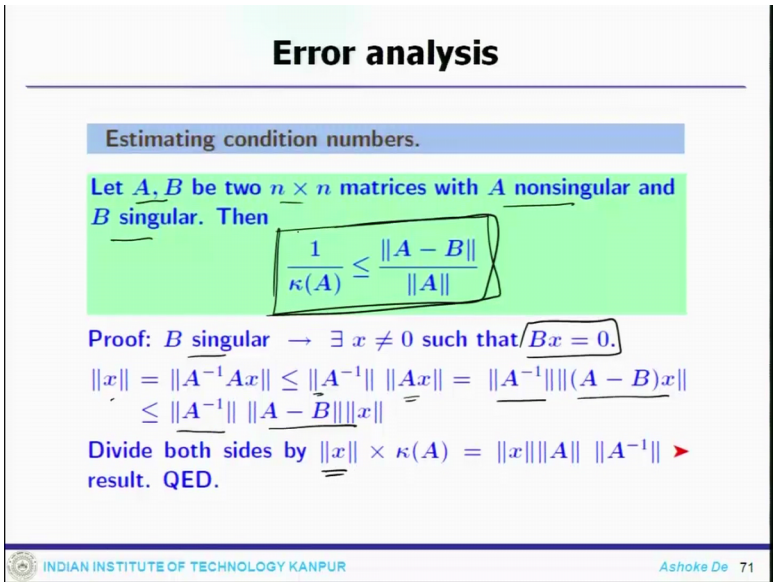
And the theorem says that, if r is b minus A y which is the residual then, w E y satisfied the maximum of this calculation.

(Refer Slide Time: 17:13)



Now, you can have an example of the ill condition system, but we may skip that.

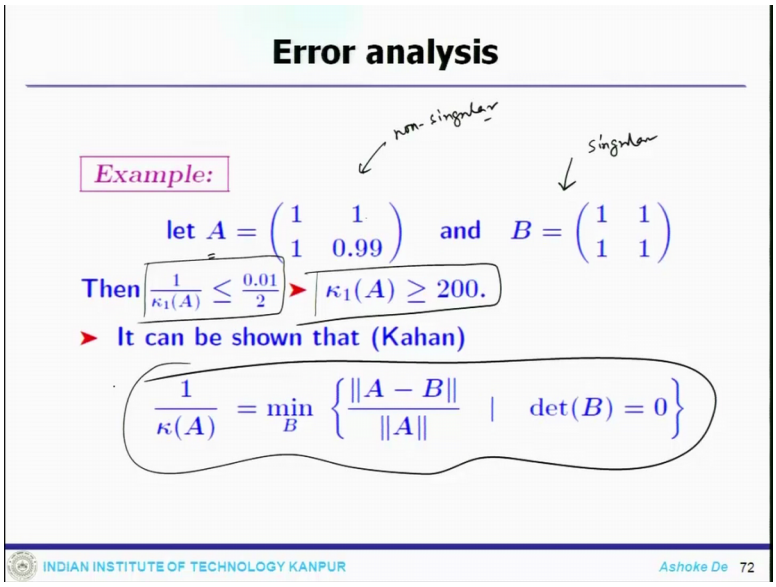
(Refer Slide Time: 17:17)



Now, this is an estimation of the condition numbers if, you have A and B two n by n matrices and A is non singular and B is singular so, that means, you have 1 singular matrix and 1 non singular matrix then, one can prove 1 by condition number is less than equals to A minus B norm divided by A norm. So, B is singular for x not equals to 0, such that B x equals to 0. Now, the norm of x is A inverse A x that, norm which is essentially less than equals to norm of A inverse and norm of A x.

So, one can rewrite norm of A inverse and norm of A minus B x. So, if I write down that inequality, it is A inverse norm of A inverse A minus B delta x. So, if you divide by both sides the norm of x into condition number it gets you this results. So, that shows the proof for that thing.

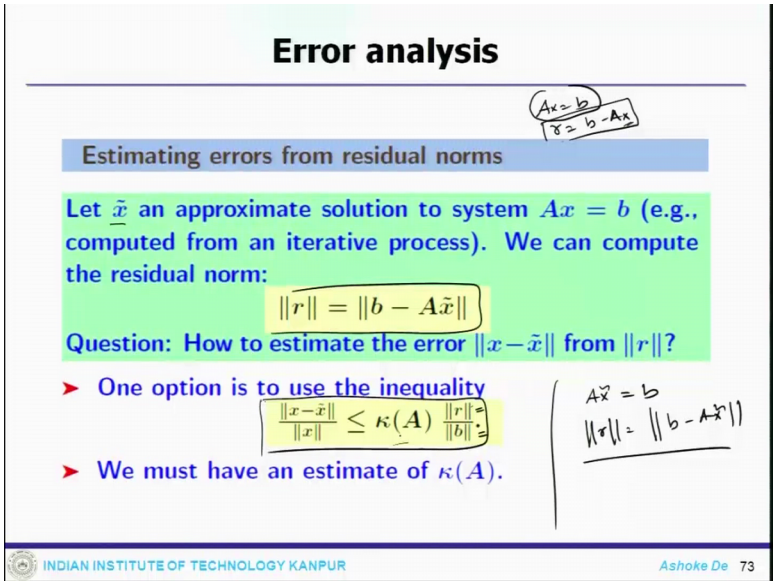
(Refer Slide Time: 18:28)



Again you take an example, let us say A equals to 1 1 1 0.99 and B has 1 1 1 0. So, this case it is singular and this is a non singular system. Then, I have to show or one has to show that 1 by condition number is less than that or other condition number of A greater than equals to 200.

So, you can show that, this is the minimum of this and find out those numbers.

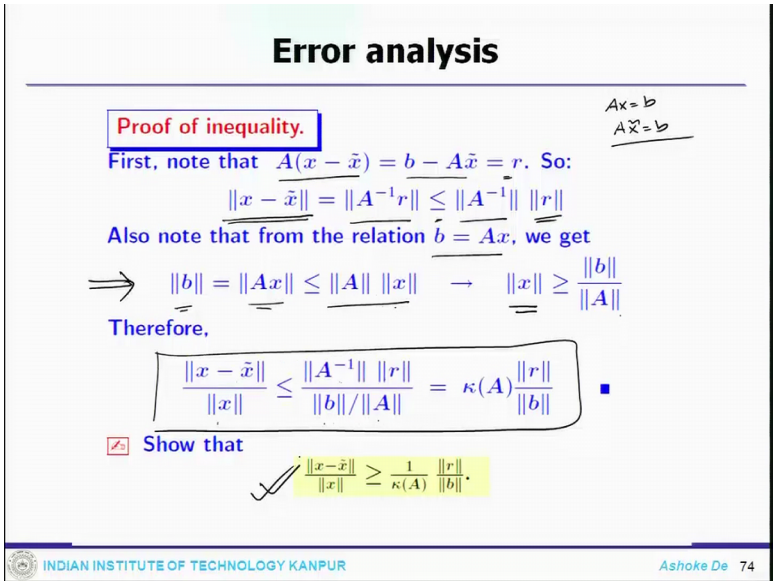
(Refer Slide Time: 19:10)



Now, it comes down to the stage where, now you get the errors from the residual norms. Residual means when you are solving A equals to b, the residual vector would be b minus A x. And this information is quite handy, while you will be dealing with different kind of iterative solvers where, you need to check the convergence level and this error becomes or slowly reducing, so that you get close to the exact solution. And how do you estimate that error, from the residuals norm let us say x prime is an approximate solution what we are trying to find out that A x equals to b and x would be the exact solution.

So, let us assume x prime is the approximate solution and then that satisfied the equation of the solution to the system A x equals to b then it should satisfy A x equals to b. And then, we can compute the norm of the residual which would be norm of b minus A x prime ok. So, how to estimate this, from this one point is to use the inequality of the norm of x minus x divided by x less than the condition number of A and divide by the norm of r norm of b. So, this plays an crucial role and along with that, the condition number which also comes into the picture to find out that system.

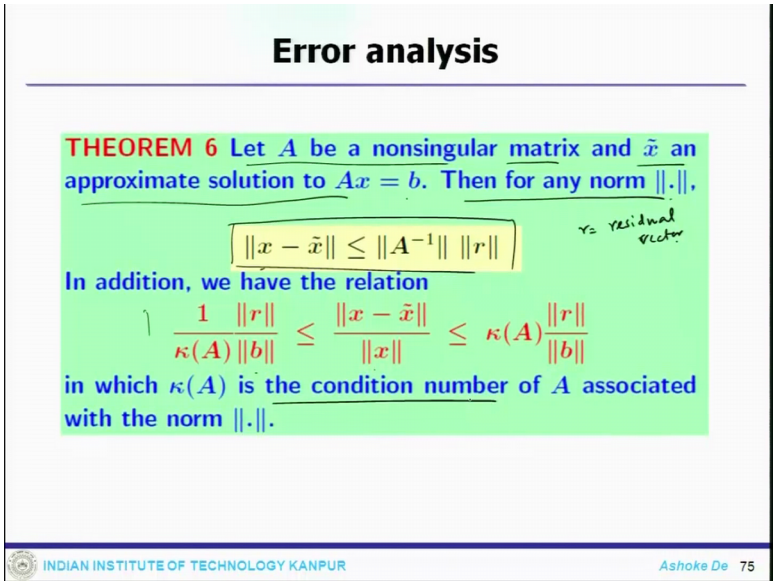
(Refer Slide Time: 21:03)



And as the proof follows you see we have now, the A x equals to b and the approximate solution satisfied A x binary to b, then one can have A x binary x prime equals to b minus A x prime which is r. So, the norm of x minus x prime is going to be the norm of A inverse r, which is when you put that equality to inequality it is norm of a inverse norm of r. So, from the relation of b equals to A x you get back, the norm of b equals to norm of A x, which is again less than norm of A and norm of x. So, what one can write norm of x from here, greater than equals to norm of b divided by norm of A.

So, you just use this simple relationship of the norms for vector and the matrix. So, now, I can write for norm of x minus x, divided by x, which would be less than A inverse norm of A inverse and norm of r divided by norm of b by norm of A. So, which nothing but the condition number of r and b so, one can show that it is greater than this factor.

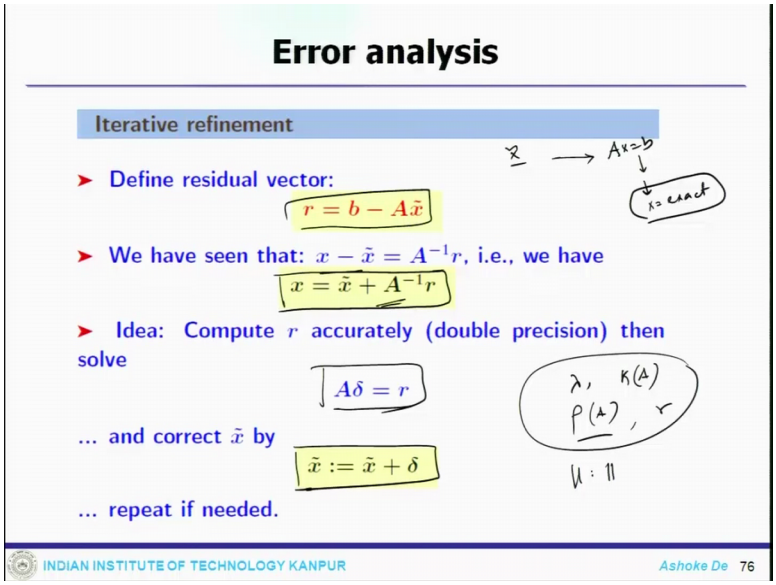
(Refer Slide Time: 22:47)



So, that takes you to the theorem number 6, which is let A be a non singular matrix and x prime an approximate solution to A x equals to b. Then for any norm x minus x prime, the norm of that less than equals to A inverse norm and the r norm.

The r is the residual vector. Now, we have the relationship 1 by condition number A and magnitude or norm of r norm of b less than equals this. So, which so, k A is the condition number and A is associated with the norm.

(Refer Slide Time: 23:49)



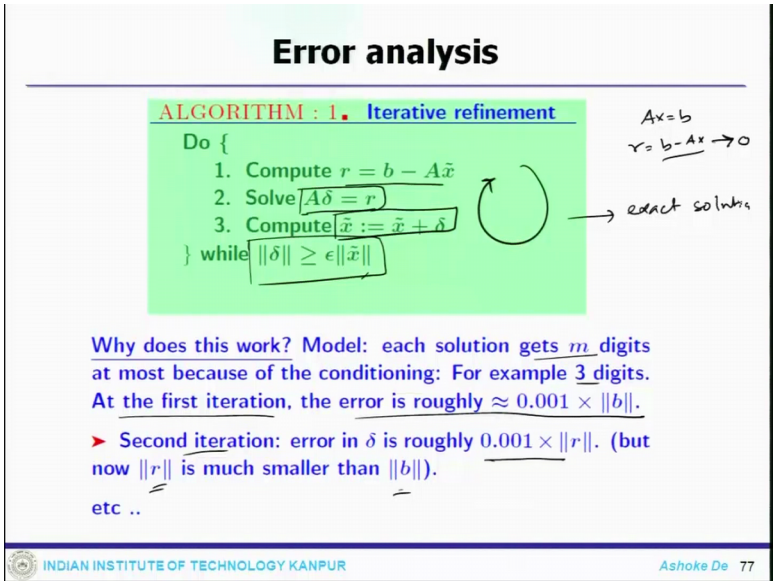
Now, when you go to the iterative solution procedure because finally, you are solving A x equals to b and slowly approaching to the exact solution of a through iterative procedure.

What you need to do, you can say that, x prime as we said it could be the approximate solution to this particular system then, the residual vector would be r equals to b minus x prime and we have seen that x minus x prime equals to A inverse minus r. So, we have x equals to x prime plus A inverse r.

Now, what is the idea? The idea behind this process is that, you compute r accurately, that means: precision wise accuracy, then you solve for a delta equals to r So, if r tends to 0, then delta should approaches towards the solution and the correct x bar as x tilde plus delta.

So, what happens if residual vector tends to 0, that means, solution is going towards the convergence then this delta would be also tending to 0. Then the solution the approximate solution will reach towards the exact solution and this is what it is done in the iterative process. And when you do that, actually you have some terms which are associated with A inverse and all these calculations that is taken care of through all these and while doing those things the properties and the theorems that we have discussed they become important.

(Refer Slide Time: 25:56)



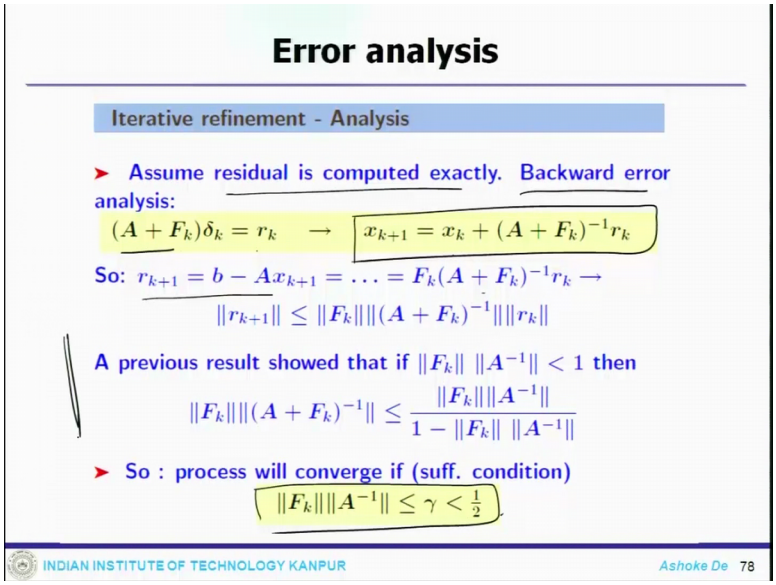
One has to look at lambda, one has to look at the condition number of A, one has to look at the spectral radius of A all these parameters, residual vectors, norm so becomes important. Now, what is that refinement algorithm if one has to see through the pseudo algorithm? You compute first the residual vector b minus A x prime then, you solve for A

delta equals to r. So, now, if you see your original system was A x equals to b and your r is b minus A x So, theoretically when you get the exact solution this should approach towards 0.

Now, what you are solving here, the approximated solution a delta equals to r then you once you get delta you update your approximated solution by x tilde plus delta. While delta or the norm of delta will remain within this bound, which is greater than equals to epsilon and mod of x tilde. So, this is where so, why it works actually. So, you get some digit, at most because of the conditioning for example, 3 digits at the first iteration, the error is roughly, this second iteration the error delta would be like that, but now residual or the norm of r is much smaller than norm of b.

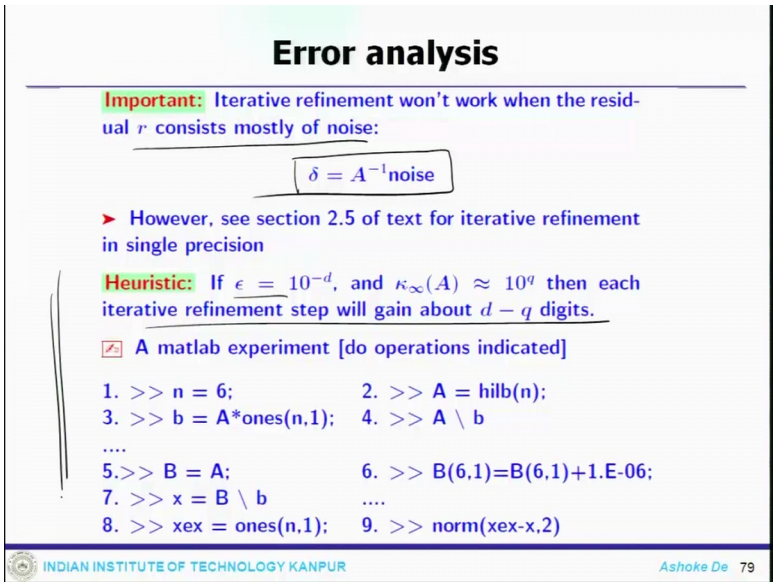
So, that is why slowly, while you keep doing this process, you move towards the exact solution and you can reduce the computational over rate by not doing the direct approach.

(Refer Slide Time: 27:39)



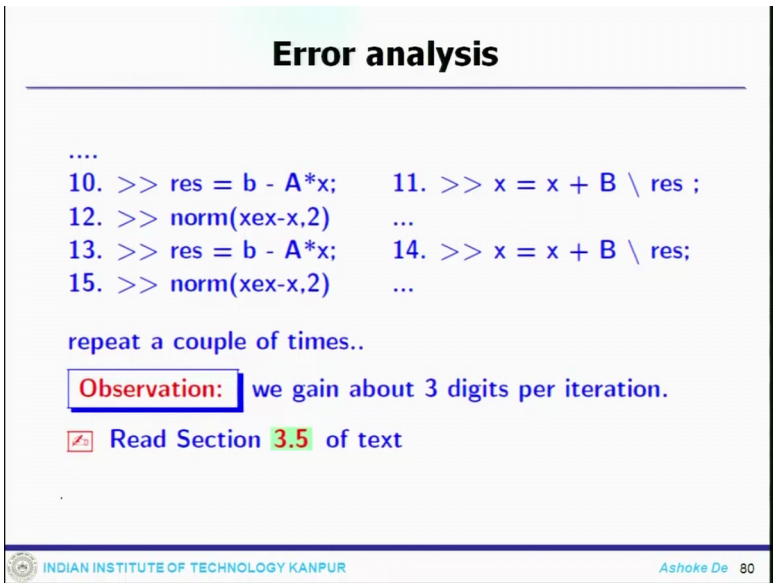
Now, once you do the refinement analysis, so what we have done that residual is computed exactly, so A plus so the backward error if you want to estimate A plus F k delta k equals to r k. So, at the second level of refinement x k plus 1 would be x k A plus F k inverse r k. So, one can see that r k plus 1 is b minus x k plus 1 and similarly if you write. So, the convergence is this sufficient condition for the convergence that one has to achieve or these are the similar kind of prove that we have done.

(Refer Slide Time: 28:25)



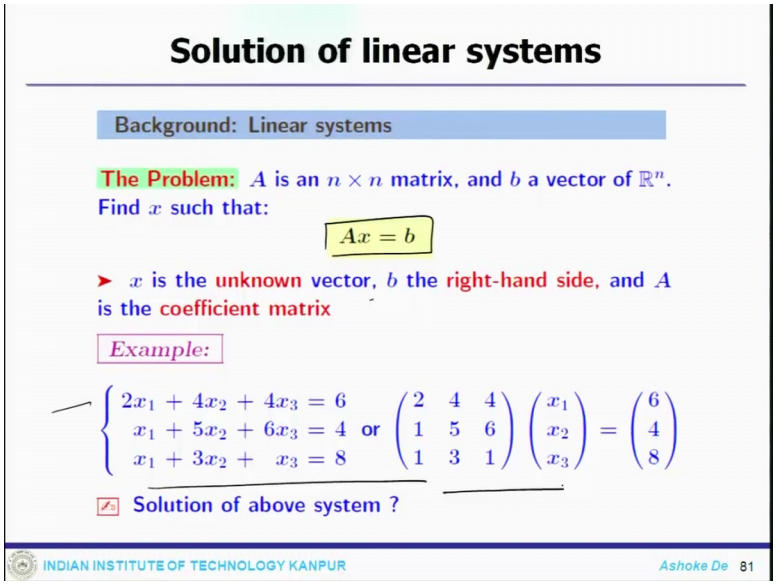
So, the now what is important here is that, iterative refinement would not work when the residual are consist mostly of noise; that means, delta equals to a inverse noise. However, when you heuristic is that epsilon is 10 to the power minus d and condition number of these. So, the iterative refinement we gain by this much of digits. So, these are the things one can calculate.

(Refer Slide Time: 28:53)



And these are the observation what one can obtain.

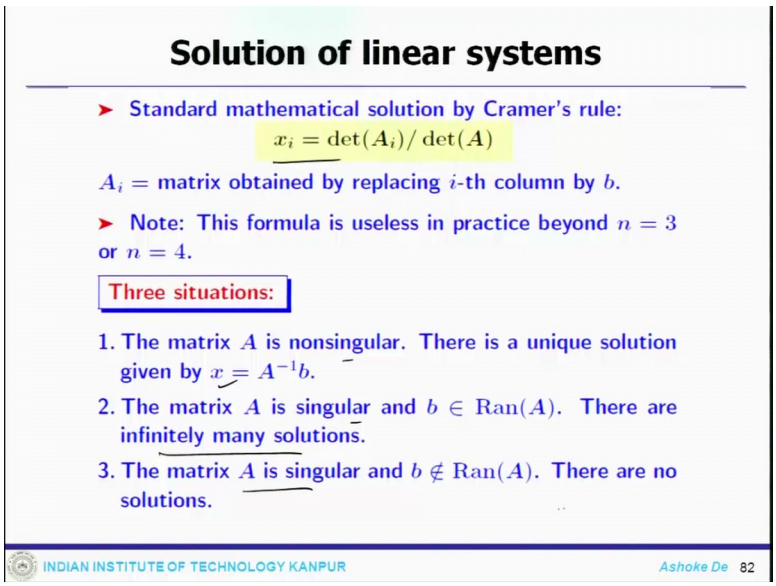
(Refer Slide Time: 28:58)



Now, the linear system the background is that, you have A n by n system or with the b vector.

Vector b right hand side, so you have to solve, so this is a system for example, you get 2 x 1 x 2 and you get A x equals to b.

(Refer Slide Time: 29:17)



So, A get you the coefficient matrix and standard mathematical solution by Cramer's rule will get you like that. And the situations the matrix a is non singular, you get a solution if matrix A is singular, there are infinitely many solution of matrix A is singular there are no solutions. So, this is where, you can see when you are trying to solve A x equals to b

system and these are the properties one has to keep in mind. So, now, once we talk about all these properties and these are the basic important properties, one has to know regarding a linear system or the matrix.

And once we know all these properties then, it will be easier to talk about out or discuss about the linear system and how we get a solution procedure and the linear solver, different kind of linear solver. This actually concludes the content of the finite volume 1 and it will give you an jump start or get you started with a finite volume method to going on. And I hope you have enjoyed this lecture series and continue in the next part of this finite volume series.

Thank you very much.