Convex Optimization

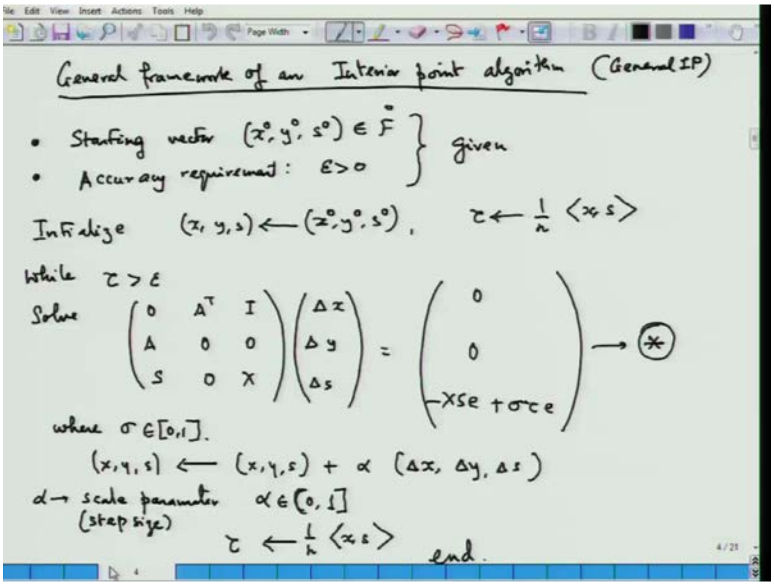
Prof. Joydeep Dutta

Department of Mathematics and Statistics

Indian Institute of Technology, Kanpur

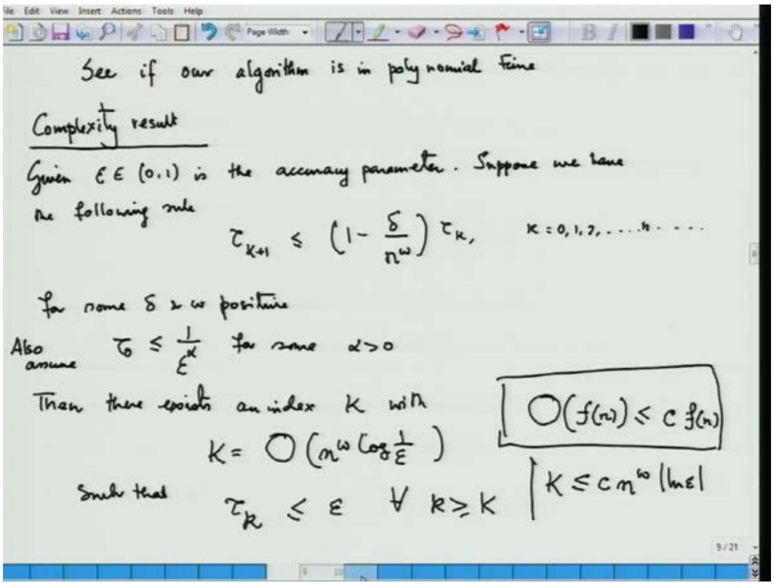
Lecture No. # 31

(Refer Slide Time: 00:28)



So, before we progress to showing that the IP algorithms are actually polynomial time, it is good that we take a short look at what we had done yesterday. So, we will just go over it very fast. So, this is what we did in the last class was to talk about the general framework of an interior point algorithms, where you have a starting point, and then you initialize the duality measure, and you solve not exactly the relaxed k k t condition. But some little modification of it using the centering parameter sigma, and that would once done; this is the way you update from a current solution to a new iterate. And then we did certain estimates which allowed us to show that when I change from estimate x k to x k plus 1, my duality measure is actually going down, because the whole aim of the interior point methods is to pull down the duality measure and force it to 0.

(Refer Slide Time: 01:27)



So, our aim now, is to see, see if our algorithm is in polynomial time. Now it would take quite a while, and it is quite technical to go into the details of what it means by polynomial time. Again I would like to stress looking at the wide range of audience with wide range of capabilities and backgrounds. I would say that I would like to again assert that polynomial time means, an algorithm which works fine; an algorithm working in polynomial time means, an algorithm is working fine. In the sense that I can tell you that the number of steps required to stop the algorithm is bounded by a polynomial, which is based on the number of data inputs or the say x n in the r n plays the quite an important role in the polynomial. So, it is a polynomial in n.

So, you know that if I know what is n? I can say that within this number of steps my algorithm is actually going to finish up. The lot of issues here, but we are not getting into this too much of detail. So, let me start with this complexity result: The statement goes as follows, and this is from these book primal dual algorithms for interior point methods by Stephen J. Wright. It would be good that if you have a look at the book I am showing, it is I think this is much better technique then writing it down. You could see the whole title, the publishers Siam and the author

So, let us take epsilon between 0 and 1. This epsilon as I told you before, this is if you look at the frame work, what we intend to achieve? We intend to achieve, as we work on till epsilon is less than equal to epsilon; till tau is strictly bigger than epsilon, we continue doing this; repeating the courses and tau is less than equal to epsilon we will stop the algorithm.

So, is the accuracy parameter? So, we are driving tau to 0; we need not reach 0, but if we come very near 0, that would almost all almost acceptable for us. So, as I want to assert you every time that optimization is not really the science of finding the exact solution, but it is really a some sort of science, rather than mathematics of finding something which is satisfactory and how much satisfactory? That is a question. That brings into brings into in front of us whole gamato questions that need to be discussed. But for us by viewing a good or a happy thing or something, which I rather, I have been satisfied with is that, if tau is less than epsilon; that is what I want; nothing more. So, that is my requirement.

So, this is my parameter. Suppose we have the following rule, that is you are generating a sequence, where you have the following rule that the duality parameter at the k plus 1 iteration is satisfying this expression; so, this recurrence relation for some delta and omega positive. Now, also assume that when you start the algorithm mu naught sorry tau naught, the starting duality measure. So, you have the starting point x naught, y naught, s naught; so, x naught, inner product s naught divided by n.

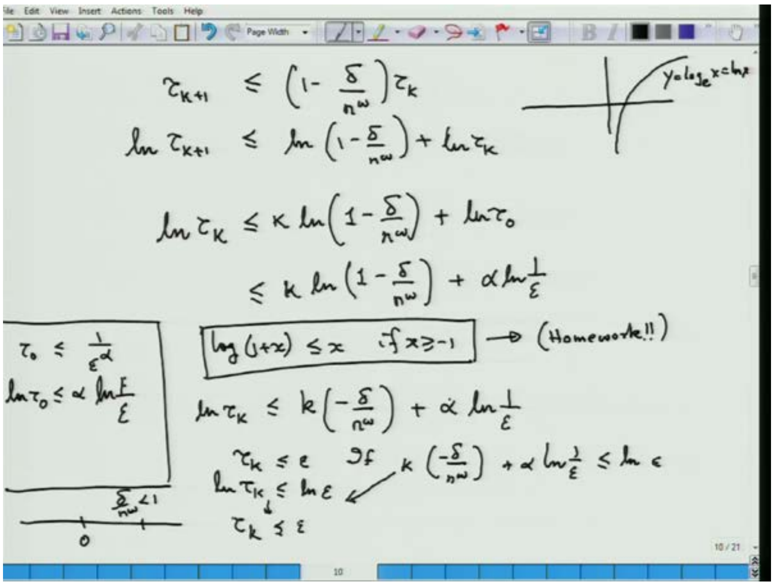
So, starting duality measure is this has satisfies this property, for some may be its greater to give a different notation. for some I guess I would say some alpha or some alpha greater than 0. So, we were assuming this. We can device alpha algorithms, which will actually follow this sort of property, and that is what we are going to show next, that we can device an algorithm, where these properties are followed.

So, mu naught is less than equal to 1 by alpha, for some alphas strictly bigger than 0. Then, we are telling, then not mu naught, I am making a mistake tau naught. Then there exist an index k; k is natural number; so, x of k, that is where you stop; there exist an index k, with k equal to big O of n omega natural log of epsilon. So, you see such that tau of k is less than equal to epsilon for all k, which is bigger than this k. Now, this capital O; so, capital O of function say f n means this is nothing but a constant times f of n; it is less than a constant time f of n. This is the meaning of big O. So, these are actually asymptotics, which tells you, how when how fast how fast this is moving with respect to this first; how fast this quantity is moving with respect to f n.

So here, it means this is nothing but some constant times, this into this; that is what it means. And you see n to the power omega; so, omega is some positive number, and n is obviously natural number. So, this is actually a polynomial in m. So, if I know m, and if I fix my accuracy constant, I can put this and I can tell you that within these steps say some. So, I can say that, k is actually like this; some constant which is which let us for the time being of course, you do not know the constant.

So, k is not more than this. So, within this number of steps at the max, you will reach this conclusion. So, that is what it tries to say. Now we are going to try to prove, what is in here; how do we prove these thing? Now, our we will start with our initial expression; the tau k plus 1 is less than 1 minus delta n omega tau k. I would take logarithms on both sides; l n means natural logarithm log two base e. So, you can do, take the log, because log is in increasing function. Look at the those, who have forgotten about this poor function called log. This is the graph of y equal to log of x log of e x, which is same as l n x.

(Refer Slide Time: 09:21)



So, here if I take the log, log is increasing function. So, x less than y means, f x less than f y; log of x y is log x plus log y, standard result. This is what you have. Now, if you repeatedly apply this formula, non log tau k; tau k is again less than 1 minus delta n omega tau k minus 1; and so you repeatedly apply, so, if you do that you will get a general formula of this form, that log tau of k is less than equal to k times, log 1 minus delta n omega plus log tau 0. Now, what was tau 0? Tau 0 is less than 1 by e to the power alpha. So, you can take log here also. So, let us see, what would what would happen here? So, I will just do this calculation on the side in case, but this is simple calculation, but still I will just do this calculation. So, tau naught is less than equal to 1 by… So, l n tau naught is less than equal to alpha times log 1 by… I can write also log of minus e, but this is nothing but 1 by e to the power alpha. So, the alpha comes in here, and so standard log formula.

So, this is becoming kappa. I will just maintain the same symbol symbols. So, these are all to the base e log base e; please remember that. Nowhere in our discussion, we are using log 2 base 10. Those who are thinking that is log 2 base 10, please note this is log 2 base e. In mathematical analysis, it is only log 2 base e. (No audio from 11:55 to 12:07) this is what you have.

Now, once you have finished that calculation, we will use the standard result; log of 1 plus x is less than equal to x, if x is greater than equal to minus 1; this is a very standard result, once you know how to write the logarithm series, logarithm as an infinite series. So, you can take up this as the homework, and try it out. And you can find it in any standard analysis books. So, here I can take 1 plus x; x could be this 1 minus delta y n omega.

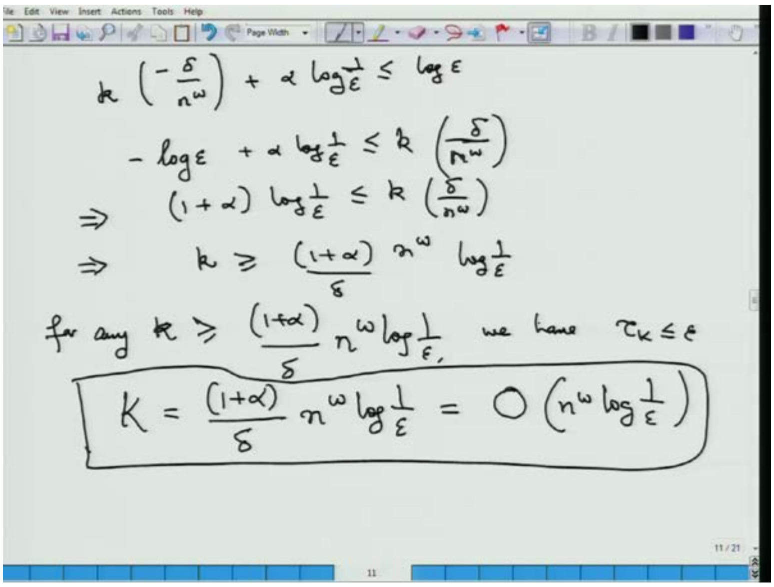
So, what would happen is the following: log of tau k is less than equal to k times; x is here, minus delta by n omega; we are see this is bigger than minus 1. You can choose it in such a way that this is bigger than minus 1 delta y n omega. So, you choose your delta n omega in such a way that this can be made bigger than minus 1; these are positive quantities. But if I put a negative 1, I can choose in the sense that, I can choose delta n omega in such a way that delta y n omega has to be strictly less than 1; otherwise this has no meaning; 1 minus delta n omega has no meaning.

So, one has to observe that this expression is meaningful, if I have this to be positive. So, 1 minus delta omega is strictly bigger than 1 minus delta omega is strictly bigger than 0. So, 1 is strictly bigger than delta by n to the power omega; so, minus delta is strictly bigger than minus 1. And so, I can apply this on this part with x is equal to minus delta by omega, and I will get this result (no audio from 14:23 to 14:31) Now, I want tau k; what I actually want? I have to find the k, for which tau k is less than equal to epsilon.

So, if tau k is less than equal to epsilon, then log of tau k is less than equal to log of epsilon. So, we would we would like to see, if this is less than equal to this part, if this expression, so, if k minus delta n omega plus alpha l n 1 minus epsilon. If this expression is less than equal to log of epsilon, then it is guaranteed that log tau k is less than epsilon, and tau k is also less than epsilon.

So, this will at least give me, what is the number of steps required to reach this. So, if this happens, this would imply this and that. And that these would imply tau k to be less than epsilon. So, what I would now have is, k times minus delta n omega plus alpha times log 1 by epsilon is less than log epsilon. Now, I will take k on the other side. You have to find this k. So, k times minus delta n omega minus alpha log of epsilon are less than log of epsilon. So, you see what it tells me that for any k on effort, the k for which this is less than equal to log epsilon. So, if you take any k which is bigger than that k, for which this inequality holds. Then of course for all such case, this will be satisfied

(Refer Slide Time: 15:35)



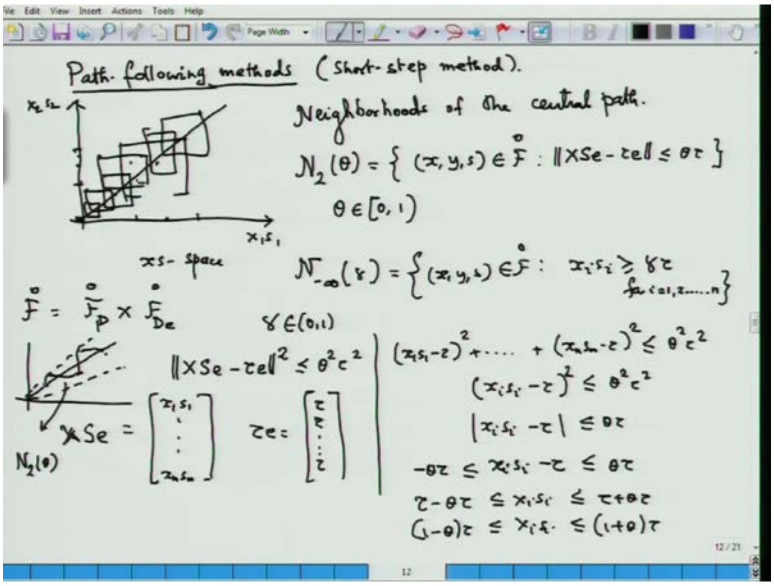
So, this is the minimum value of k, for which this will hold, basically this expression. That is what I want to find out. So, what does it say that, if this would happen, then I would take it to this side. So, I will have k times; no it is much better, and this is not the way. It is much more simpler to do it in this way. So, I will put log minus log epsilon plus alpha log 1 by epsilon is less than equal to k times, minus delta by small k. I am writing minus delta y n omega. (No audio from 17:37 to 17:48) Which implies that, I can make log minus log epsilon is log of 1 by epsilon.

So, 1 plus alpha times log of 1 by epsilon is k times delta y n omega. So, it implies that k is greater than equal to 1 plus alpha times n omega by delta into log of 1 by epsilon. So, what you can write? Simply, you see this is now; this just tells you it is in polynomial time; you can also write minus, and write log mod e; that does not a very big problem. Here, if you do not write in this way, here in this case, instead of this, you can just also write log of 1 by epsilon. It does not matter; it is same thing.

Now, what is this? What does it tell you? It tells you that, for any k which is bigger than this, this expression would be less than log epsilon. So, for thus tau k would be less than equal to epsilon. So, for any k bigger than equal to 1 plus alpha delta by n omega, log 1 minus epsilon, we have tau k less than epsilon, because what we have shown that if this is less than this then k must be bigger than this 1. So, when k is bigger than this we can just work out in the reverse version, and show that this is happening.

So, once this happens the log of tau k is less, than log of epsilon. So, tau k is less than epsilon. So, if I say k is equal to 1 plus alpha times delta, n omega log 1 minus epsilon; this is nothing but if I take this as the constant C, and this is nothing but order of n omega log epsilon. So, with. So, here I have expressed the whole thing as a function of the problem data. So, in that shows, that the problem is polynomial type. Now, can we device a algorithm can we device an algorithm which will actually follow all this property, and follow these sort of polynomial time framework. Can I device something? So, that would lead us to the study of path following methods; obviously, all the methods are path following, except the potential reduction once.

(Refer Slide Time: 20:40)



Path following methods: So, we are first talking about the short step method. (No audio from 20:54 to 21:02) So, what does this short step method do? In the short step path following method, we have to realize one thing; we know that, if I project, say, my central path on to the x space. In this case, just we are looking at two variables, because that is the way we can visualize it; and the projection here is nothing, but the central path projection is nothing but straight line, but straight line through the origin, because x 1 s 1 equal to x 2 s 2. This is straight line which is dividing the archon into equal halves.

Now, once I solve the equation which is here, once I solve this equation, I am not exactly getting the solution of a point, which will take me to the central path or I am not exactly solving the relaxed k k t condition. So, as the result of which, what I would get is not a point on the central path, but some point near it. But we cannot allow these points to wonder of too much; that is we cannot allow this point to come near the boundary, where one of them x 1 s 1 can drop to 0; x 1 s 1, x 2 s 2 can drop to 0.

So, that is what we cannot allow. You cannot allow the points to come too far from the interior, and more towards the boundary. That is what the barrier does; it stops you from going to the boundary. Now, we have to then somehow force these points these solution points; this approximate solutions actually to stay in a region, near the central path.

So, we have to define, what is called the neighborhoods of the central path? (No audio from 22:46 to 22:58) So, that two types of neighborhoods, and one is called shorter neighborhood; the neighborhood which allows shorter steps; there is a short step method. And another one, which allows you more flexibility, and that would be used for the long step. So, what we will use now is the following neighborhood. So, we will write down the two neighborhoods. Now this is an element of F naught, and I want to again remind you that, F naught is nothing but F D cross F naught D cross F naught D e, which you already have seen.

So, this tells you, that the standard Euclidean norms: Finding the distances; standard distance; finding thing; distance finding; computation. So, once you know the duality measure at a particular point, find all such points which are of this form, that corresponding to this tau is the duality measure; that is x s, x in a product x by n; this has to be bounded by theta times tau, and e is obviously, 1, 1, 1, 1; the vector of ones. So, theta is a quantity, which could be 0, but cannot become 1. So, I am defining a neighborhood.

So, one another neighborhood which we will also use, which is called the infinity neighborhood or large neighborhood; what they are trying to do? They are trying to see this much; how much you are form your duality measure, how much you are far from the duality measure? And that is exactly what they are trying to see; how much you are far from the duality measure? To what extent you are violating this thing, this duality measure?

So, let me write down the second one. So, you take a gamma which is between 0 and 1. We will now really be concentrated on this neighborhood, rather than one I am just writing, but this is just for your information at this moment. So, when we do it, we will not mention it again. (No audio from 25:36 to 25:50) So for us, it is important to know, how this sort of neighborhoods looks like. If you look at this, how this sort of neighborhood looks like?

So, what is this? This is Euclidean norm. (No audio from 26:05 to 26:15) This square is less than theta square tau square. Now what is the meaning of this? So, what is x s e? Let us look at again, just to remind you this is nothing but this vector; (No audio from 26:32 to 26:42) and e is 1, 1, 1; and tau is nothing but tau, tau, tau. Tau is nothing but all components are tau. So, basically if I look at it, this would give me nothing but x 1 s 1 minus tau whole square, plus x n s n minus tau whole square, which is less than equal to theta square tau square.

Now, each of them has to be less than this quantity, because these are all non negative quantity; it is a positive quantity; it is adding up to the quantity which is less than a positive quantity. So, each of them must be less. So, what I am having is, say tau i s i minus… x i s i minus tau is less than theta square by theta square into tau square. So, this would just give me that which simply means we can just directly write down from this. This is what it means. Which means that, tau minus theta tau is less than x i s i is less than tau plus theta tau, which means x i into s i for every i is 1 plus theta into tau; 1 minus theta into tau.

So, if would each i x of here x 1 s 1, x 2 s 2 is both within these two limits. So, it changes; so, what does it says? That x 1 s 1. So, x 1 s 1, this is say x 1 s 1 .and this is x 2 s 2. So, s 1 x 1 is within certain limit; x 2 s 2 is within certain limit. So, I want all points within this; so, at particular choice of x s. So, if we take the cartesian product of these two, this is what sort of a for a so any point xys, which is here has to satisfy something here. So, any point of this form x 1 s 1 which is say, I have given the theta, and this is my tau. Tau is tau also keeps on changing.

So, what you are having is that, at every point you take around that you are basically creating this sort of small squares (No audio from 29:23 to 29:30) of this sort of lengths; the squares the maximum length that you can have. So, it is between 1 minus theta into tau into 1 plus theta into tau. So, within that limit, everything should lie. You can keep on changing x and s s; the tau will keep on changing. So, those lengths of the squares would change. So, basically you would have something like this. So, if you would really draw this up a bit, so, this will be some if you as your theta tau to become 0, the size of the squares will keep on decreasing.

So, basically if you look at if you take the edges of the square, and try to draw some sort of a tube. So, it will be something like this.

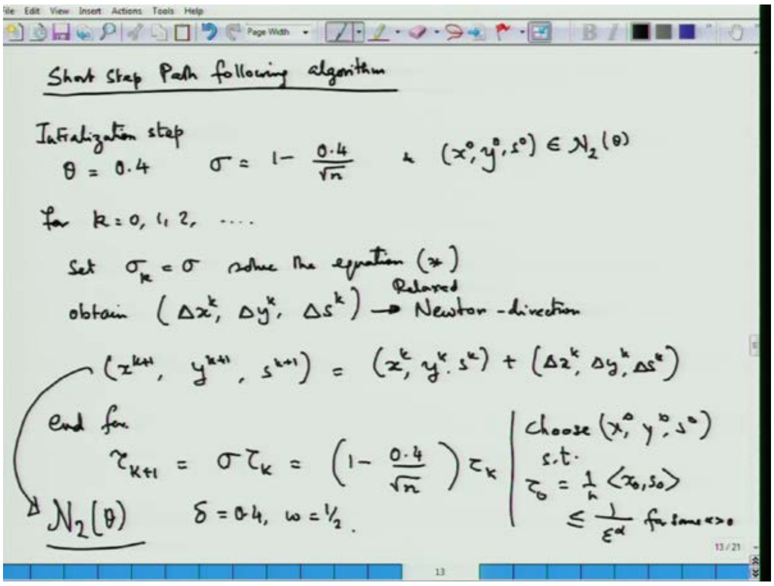
(No audio from 30:16 to 30:30)

Now this is your N 2 theta. So, what is happening with this N 2 theta is that, you can now start from a point here; move to a point here; but you the algorithm should be such that.

That is what I will show you that; short step method would be such, that you are always remain with this N 2 theta, but you cannot make very large steps, cannot take large steps, because you will then there is a chance of you getting outside N 2 theta.

So, because you cannot do that, this is called the short step path following method, when you use such neighborhoods. This neighborhood would allow you to be much more flexible. So, let us just for the time being written down the short step path following algorithm. short step path following algorithm. So, initialization step, (No audio from 31:41 to 31:51) you will create the neighborhood N 2 theta. So, that force all the point; that is this sigma; the centering parameter is 1…

(Refer Slide Time: 31:23)



You see the these choices are helpful, because with these you can choices you can show that the whole algorithm within polynomial time. And you have a starting point you have to guarantee; not sorry not lambda is y. This starting point has to be in these guarantee has to be taken. So, of course, you might have I cannot really find out such a thing; I can just find out a point which might be infeasible; then there is something called infeasible you have interior point methods, where you start with the initial infeasible point, but after certain success successive steps you get a feasible point.

So, just for his simple explanation this of this tau algorithm, we are essentially taking this fact that we are already in the feasible set. So, now. So, do while loop for you set. So sigma is k is not adaptive; it is not changing; it is fixed. And solve the equation solve the equation star. So, this is what you have to solve the equations star. Obtain delta x k, delta y k, delta s k and write. So, you are doing for loop for k equal to this to this to this. So, it is the next.

Now, here what you take is a pure Newton's step; you do not have here as scale, which is less than one; scaling parameter order control parameter with controls, how much you move from iterative x k, y k, x k in the given direction. So, along the Newton direction… This is the newton direction, of often this, yes you have to relaxed Newton direction; maybe I should write relaxed newton direction. So, you add these two…

So, this is how you keep on doing, and then you have to… This is your basic algorithm, and you end for loop. So, this is you keep on doing this, till you have tau less than epsilon. Epsilon has to be put in as per as your requirement and for. So, this is the short step path following algorithm. Now, tau k plus 1: we have shown earlier; if you remember this estimation that we did; what did we do this is the estimation. If tau k plus; 1 in this case, of course, is depending on alpha; here alpha is nothing but 1. Here, if you look at this, I have no alpha here corresponding to what was here; there was an alpha here between 0 and 1; my alpha f chosen it to be 1 in the short step path following method.

So, you have alpha equal to 1. So, tau k plus 1 if you again go back, and see the calculation, tau k plus 1 is 1 minus 1 minus sigma into tau. So, in this particular case, just let me write down. So, if when I take alpha equal to 1, this 1 and 1 will cancel, and you will have sigma into tau. In this particular case tau k plus 1 is sigma into tau k. And what is sigma? 1 minus 0.4 by root n tau k; so, here my delta is 0.4 which is same as theta. My delta so my So, here I am having everything in the form this, and my delta is 0.4, and my omega is half. Now, what does it show? I am actually having a scenario of this result; I am actually having the basic requirement of proving the complexity or the polynomial time.

Only if I choose my x naught, y naught, s naught in such a way, such that tau naught is less than 1 by epsilon by alpha, then I am done; then I already put the sequences in this form; then I am just already, I am just straight done. So, I just have to choose x naught, y naught, s naught such that, tau naught which is nothing but 1 by n inner product x naught, s naught; this tau naught is less than 1 by epsilon to the power alpha for some alpha greater than 0; epsilon is by chosen parameter, that depends on what you choose.

So, then you immediately know that you would have a polynomial time algorithm, because the basic equation is already satisfied by your choice of delta and omega. So, we have learned quite a bit of stuff today, that the general IP method; that the general format, under certain mile condition can be shown to be polynomial time. We have created an algorithm called short step path following algorithm, where all the points that we generate by solving the equation star is forced to remain in a very narrow cone, and that those points do not allow you to move very far off from the central path, but keeps you near the central path.

So, that is why it is called the short step path following algorithm, and we have showed that this short step path following algorithm can be easily made to follow the polynomial time pattern. And now, we have done our first step; polynomial time game is over. We have proved that, it is we can show it is in polynomial time. But what important stuff that we have to show, is that, when I compute x k plus 1, y k plus 1 and s k plus 1 from x k, y k, s k, then I must be sure that, this x k plus 1, y k plus 1 and s k plus 1 is also in N 2 theta. And that is what we will discuss tomorrow, that how can we show that this is in N 2 theta. Tomorrow's discussion will concentrate on that, and then we will show some more, little bit more algorithms, before we wind up this discussion on interior point methods.

Thank you very much.