Design and Optimization of Energy Systems

Prof. C. Balaji
Department of Mechanical Engineering
Indian Institute of Technology, Madras

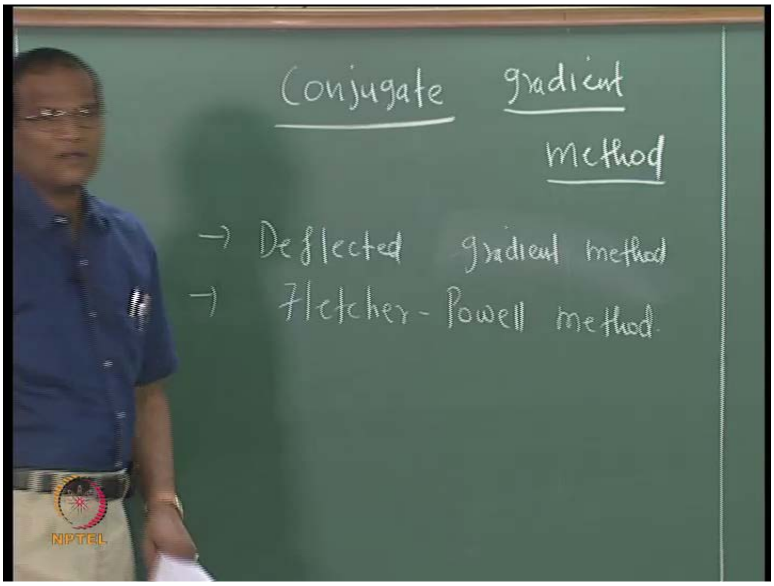
Lecture No. # 35

The Conjugate Gradient Method Contd…

So, we will revisit the conjugate gradient method. So, in the last class, towards the end, we wrote the algorithm very quickly and without paying much attention. So, the first few minutes, we will discuss the algorithm. Of course, I do not have time either to prove the, prove how you got that beta or to solve problems. It is basically an extension of the steepest descent or steepest ascent method; except that, after first two steps, you will correct; you will deflect the steepest gradient and proceed in a new direction. This is quadratically convergence; it can be proved. The proof is available in many books.

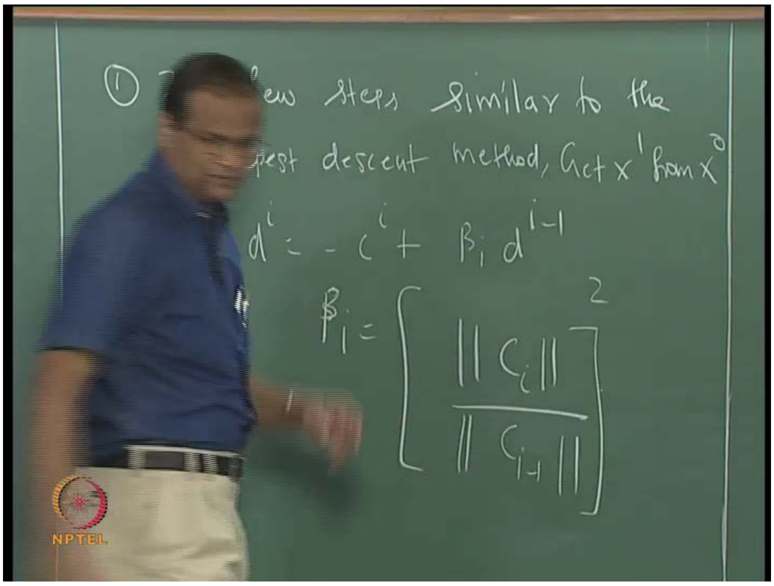
For example, you can look at the book "Engineering Optimization" by Ravindran, Ragsdell and Reklaitis; Reklaitis from Broadwig university; he is one of the stalwarts in optimization; he worked out lot of algorithms; that book is pretty good. Or, "Optimization Engineering Design" by Kalyanmoy Deb, Prentice Hall, that is available; that is a reasonably good book; or, "Optimization Techniques" by S. S. Rao. The conjugate gradient proof must be available in any of these books. We will quickly go through this conjugate gradient.

(Refer Slide Time: 01:18)



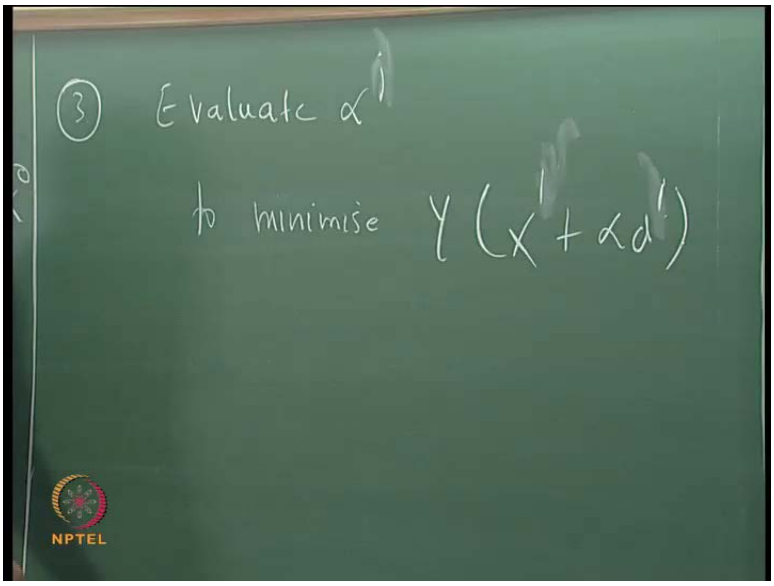
So, it is basically a deflected gradient method, also called the Fletcher Powell. Then, somebody called Davidan, not Davidson, who made some improvements to this, which called Davidan Fletcher Powell - DFP method - which is state of the art, for many of the problems. And, there is no need that your objective function must be quadratic. It is very rarely to encounter a quadratic function. But, what these people have done is, they have shown that the algorithm can work for non-quadratic functions also. So, the DFP method is pretty powerful; you can write papers out of that.

(Refer Slide Time: 02:39)



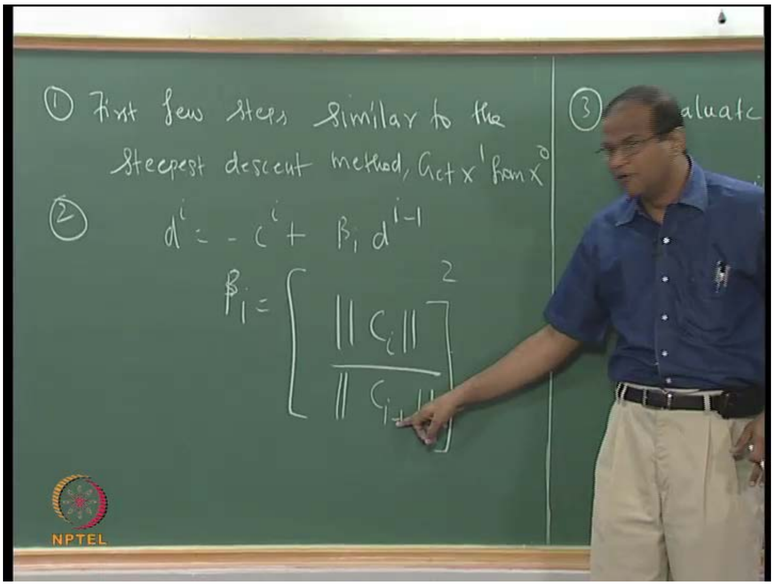
Now, when I discuss this, we said that first few steps, the first few steps are similar; steepest descent or steepest ascent depending upon whether you are looking at the, looking at a maximization or minimization problem. Then, so, you get x 1 from x naught, right. And, then, d i, it should be i minus 1, on the right hand side; that is a correction, you have to make. May be, I wrote it as d i, right; does not make sense; it has to be d i minus 1. So, where beta I; did we put C, subscript or superscript? Does not matter, so long as you understand.

(Refer Slide Time: 04:05)



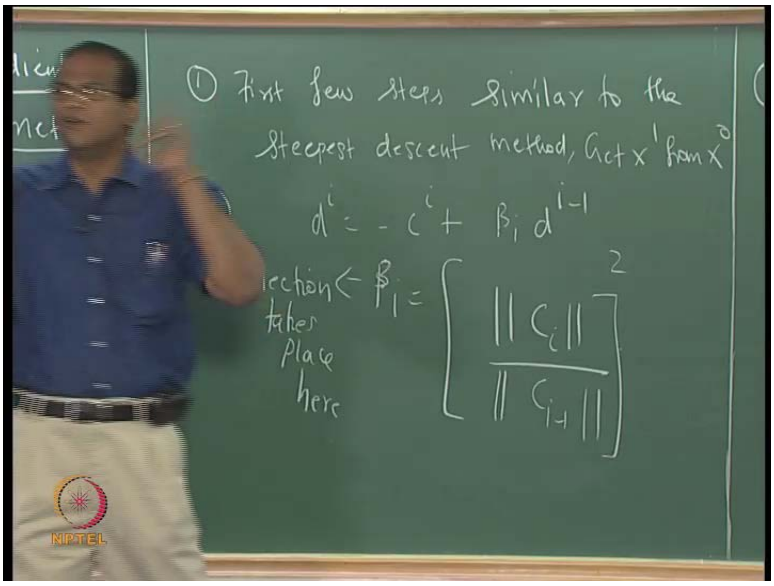
So, now what will be the third step? You get the value of beta, and then proceed to; you get the next value of, evaluate the value of alpha I; or, since, we started with 1 and 0, we can call this as alpha 1. So, the key point is d 1; d 1 is not equal to minus del y; C is equal to del y or minus del y? Student: C is del y. C is del y, d equal to minus C. Therefore, originally in the steepest descent or steepest ascent method, this d was minus del y; but now, this d is the modified d; how was it modified?

(Refer Slide Time: 05:22)



You add the d to the beta i into; beta i, is the deflection parameter. How does it get deflected? Depending on the del y at the current iteration, divided by the del y at the previous iteration; are you getting the point? That is it. So, it seems very logical, it can be proved; it requires 15-20 minutes, I do not have time; if somebody wants, I can take that proof and put it on modal. Let me see, if I have time, I will put it; there is some, I created something for this course; I will try to postulate on that. So, this can be proved.

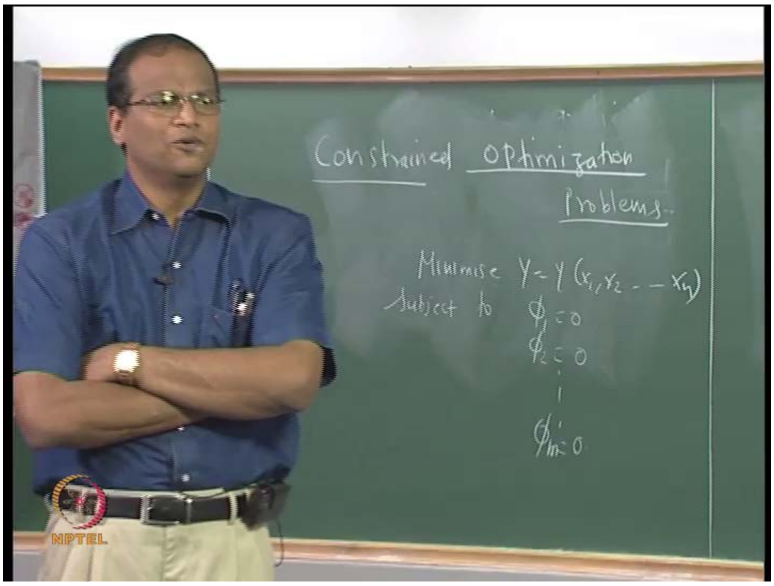
(Refer Slide Time: 05:58)



So, this is actually where, deflection takes place here. So, first to start with, how many times should we get C i to start working with the conjugate gradient? Student: 2. So, minimum you have to start with 2; it cannot start on its own; it cannot start at initio from 0. So, you have to start with the steepest ascent or steepest descent; evaluate the C, 2 times; and then only, you have an opportunity to get the beta and then deflect the d. Is it alright?

Now, so, though we have these algorithms are impressive, we are not able to handle; so far, we have not been able to handle constraints, right; all these are still unconstraint. How do you handle constraint optimization problem?

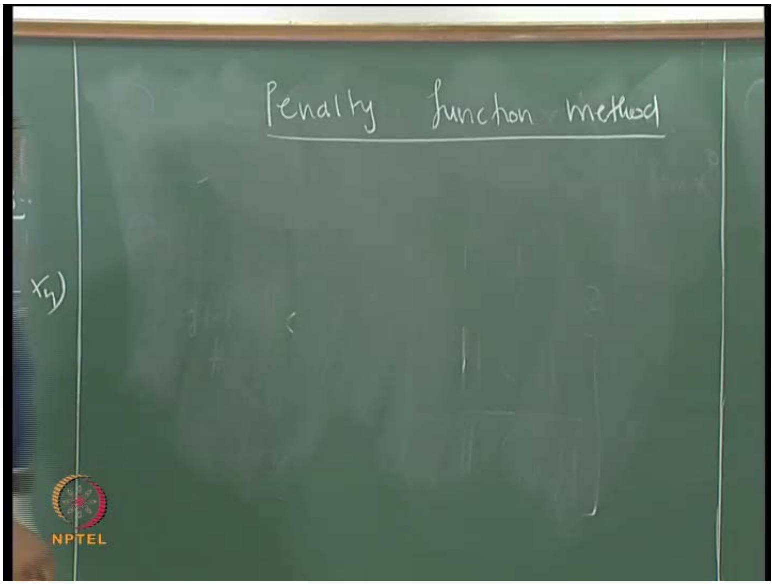
(Refer Slide Time: 07:04)



Needless to say, we are having a multi variable; we are having a multi variable constraint optimization. So, minimize; assume that we are going to handle, we would have only equality constraints. First, we will see how we can develop an algorithm for handling equality constraints; then, we will extend it to handle inequality constraint. Inequality constraints were lot more difficult to handle, Kuhn Tucker conditions and all that, did you remember? It leads to a lot of pain. Equality constraints are easier to handle, except that, the number of equality constraints must be less than or equal to the number of variables.

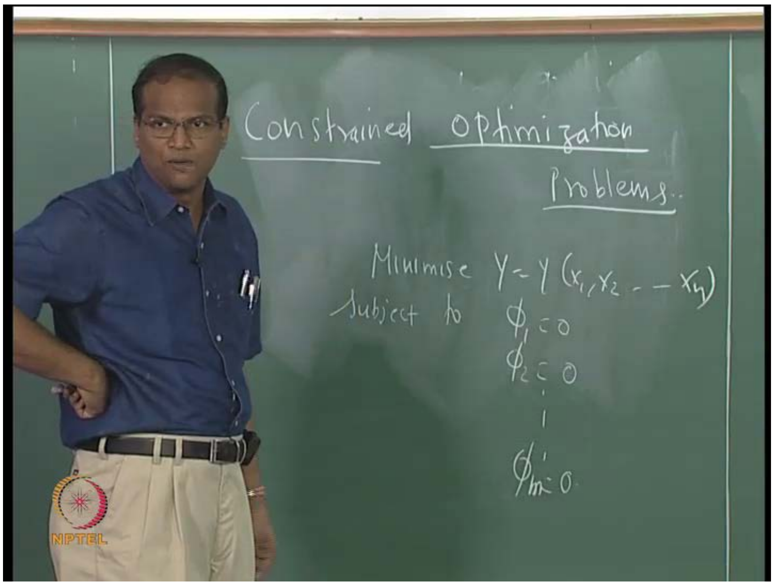
Now, there is a very interesting technique, where we transform the objective function; we transform the objective function and convert it into an equivalent unconstrained optimization problem. I come again, we transform the objective function in such a way, that the constrained optimization problem becomes an unconstrained one, and then we solve the unconstrained optimization problem with the state of the, our technically conjugate gradient or genetical algorithm or whatever; what is this concept of penalty?

(Refer Slide Time: 08:51)



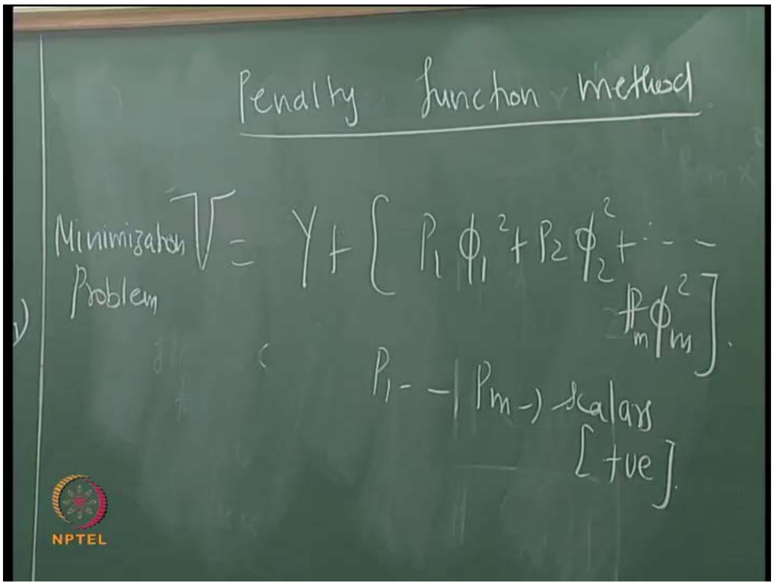
So, we introduce, what is called the penalty function method, is one of the most powerful techniques for handling multi variable constraint optimization problem.

(Refer Slide Time: 09:11)



Suppose, I want to handle, suppose I want to handle this problem, but I want to use a conjugate gradient or steepest ascent, then I will have to transform my objective functions plus constraints into equivalent unconstrained optimization problem. Therefore, I introduce, what is called a composite objective function.

(Refer Slide Time: 09:27)

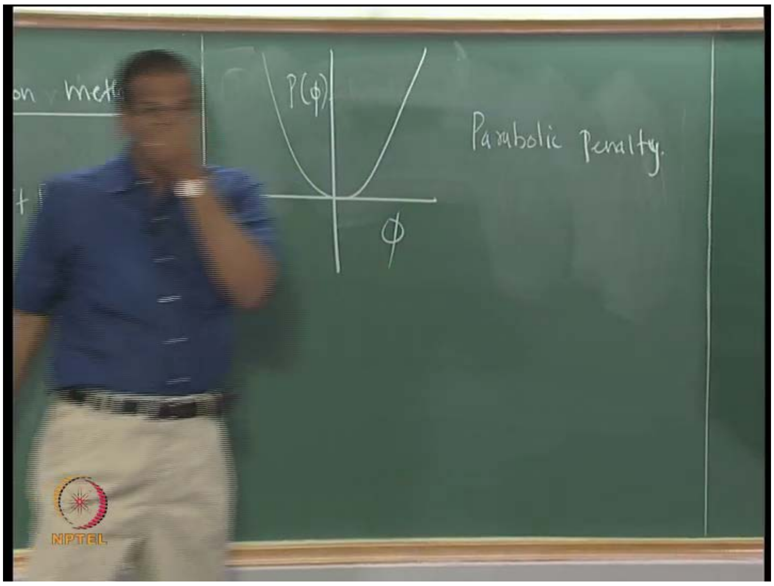


I introduce a composite objective function called V; V is nothing but Y plus P1 to P m are scalars, where positive.

What is the funda here? Each of these constraints is expected to be 0, because it is an equality constraint; if each of these phi takes on a value which is different from 0, I square it. Therefore, whether it deviates from the left side of 0, right side of 0, this phi square will always be positive; if you multiply by a positive quantity, P1 phi 1 square, P2 phi 2 square, all these things, will be positive quantities. These are all added to the original cost, because I am talking about a minimization problem.

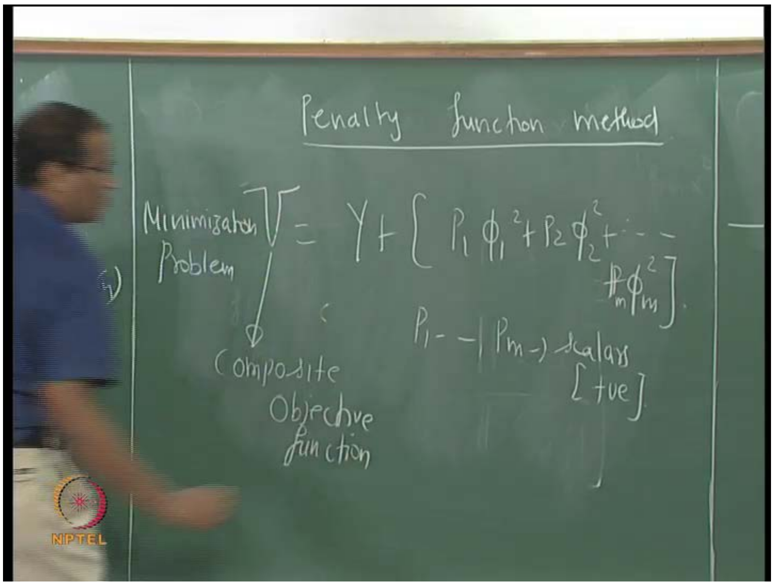
So, these are all adding to the original cost. Therefore, V will be substantially higher compared to Y; if the phi(s) are not equal to 0. Therefore, the objective function is finalized for violation of constraints; this is the penalty. If your violation is too much, then phi 1, phi 2, phi 3, everything will be more, then if you multiplied by P- P is a scalar, which can be decided - it can be 5, 10, 100, 200, 300; the analyst will decide it; the person who was optimizer will decide it; we will come to little bit later. Therefore, you finalize the objective function for violation of constraint. If all the constraints are satisfied, there will be no penalty.

(Refer Slide Time: 11:31)



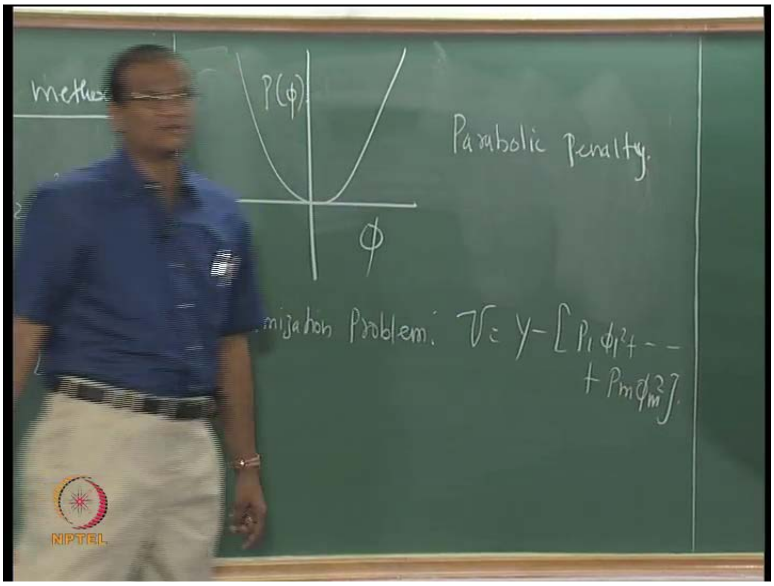
So, this is called a parabolic penalty, because a penalty is like this. Agreed? Now, you minimize V, instead of minimizing Y; that is, if you minimize V, by assigning different values of P 1 to P m, and then keep changing these values of P 1 to P m, and find out, and reach a stage where regardless of the values of P 1 to P m, you are getting more or less the same value of V, that means, it attains stationarity, with respect to this scalars also; are you getting the point?

(Refer Slide Time: 12:22)



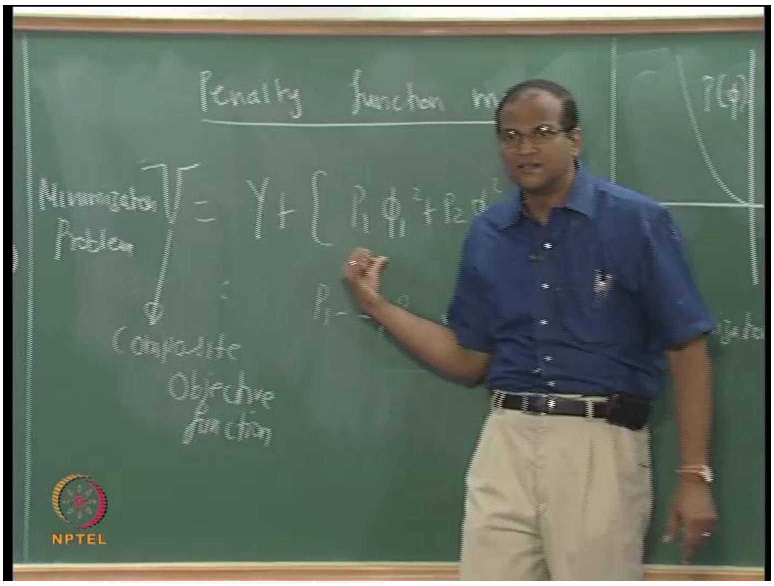
Now, we will write. So, this is called the composite objective function.

(Refer Slide Time: 12:41)



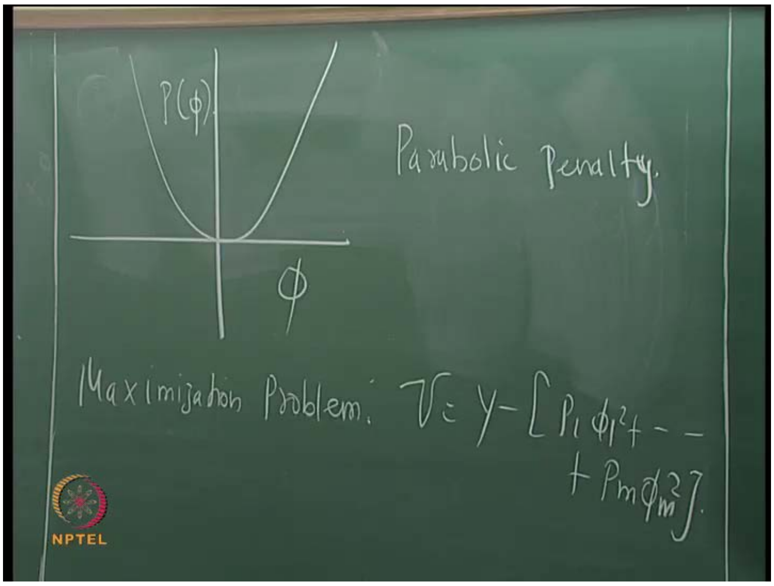
For a maximization problem, though, it may appear, though the minus sign and then positive sign appear to be incongruous, with the maximization-minimization problem; because for a minimization problem, intuitively you feel you have to put a minus, but actually we are finalizing, the cost has to increase.

(Refer Slide Time: 13:32)



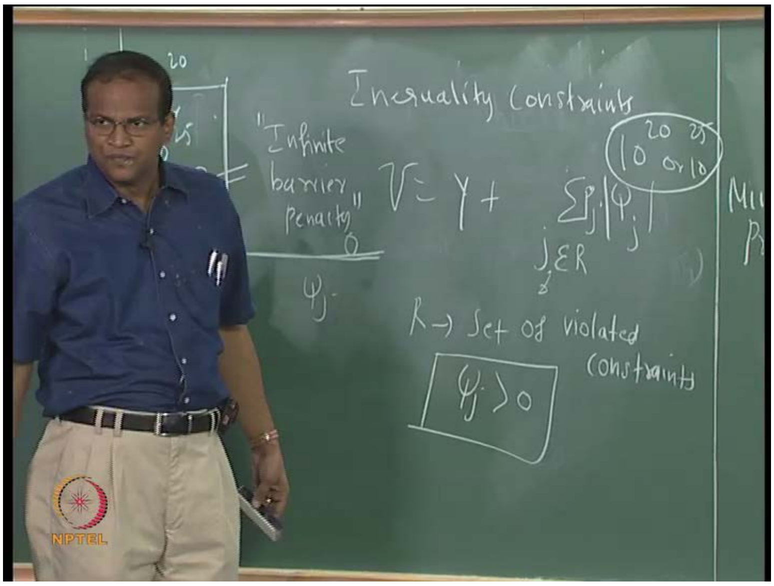
Therefore, for a watch out, for a minimization problem, it is Y plus; for a maximization problem, it is Y minus; because there is a penalty on the profit, therefore, it is minus; there is a penalty on the cost, therefore, it is plus; because it is minimization, your mind somewhat tells you that it should be Y minus; is it not so? If you argue out, it will become clear to you. Upon first glance, it seems incongruous; it seems that we have made a mistake; not so; are you getting the point? Did you confuse, or did you clarify?

(Refer Slide Time: 13:57)



So, there are other penalty parameters, which are available; penalty functions, which are available for handling inequality constraint. For example, phi of x must be greater than 0. Suppose, there is a constraint like that, for inequality constraints. Now, we are going to solve problems with inequality constraint.

(Refer Slide Time: 14:23)



Inequality constraints. So, V is equal to Y plus, let us say, P, sigma. So, j is a part of R, where R is a set of violated constraints, for example. You can put a, P j, or you can put universal P, it does not matter. If you want to be this thing and put a P j also, which means, you take the modulus of this, and if it is getting violated, then you sum up all the violations, multiplied by; this will be very huge quantity, like, 10 to the power of 20 or 10 to the power of 25.

So, this constraint could be psi j greater than 0. So, if it is like this, the penalty P; the penalty will go like this; the penalty will be 0 here and it becomes negative; it will be 10 to the power of 20, 10 to the power of 25. So, the penalty will be selectively applied, it will always look at the value of psi, whenever it goes; whenever it becomes negative, it will be applied, it will be added to the cost. So, this is called an infinite barrier penalty.

Of course, you can immediately conjure up with new ideas, where you can say that, p instead of mod log of psi x, log of mod of psi x, and so on; you can come up with your own; so, your imagination is the limit; so, you can come up with your own concept of penalty or barrier, which will prevent. See, basically, all this is, basically, you are adding to the cost, you are ensuring that the psi never become a negative; if the psi becomes negative, this fellow will, 10 to the power of 20, it will, it will become huge, and you will automatically correct it, such that psi is satisfied; are you getting the point?

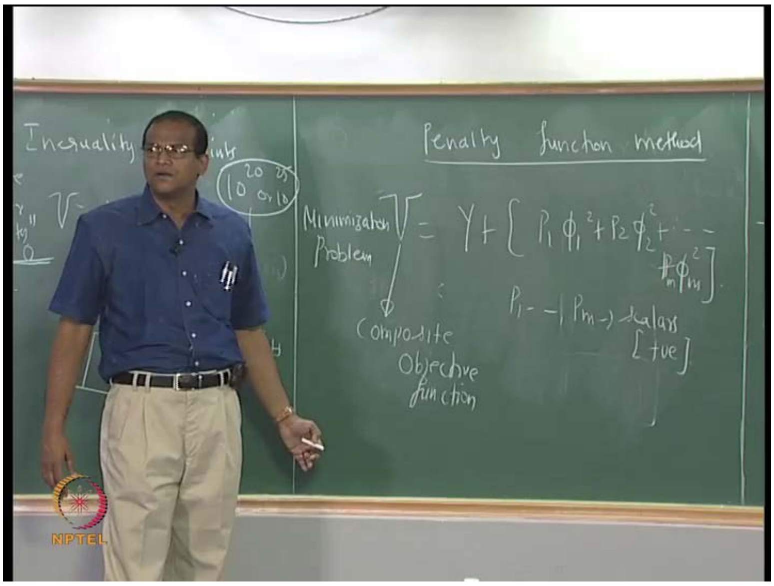
Student: You know, 1 P for each psi, right.

Yeah, yeah so,

Student: Here, we have only 1 P.

Yeah, because I am giving a very huge value of; if you do not like it, you can take the P j inside; I do not have any objection, but generally it is very high, what you are saying is all right, fine? So, but this could be; so, this represents the set of violated constraints; you ensure that, it is not violated; instead of checking each time, whether, it is violated or not, you put it as a part of the objective function, and then put, set up an infinite barrier; infinite barrier penalty, so that, it is not violated.

(Refer Slide Time: 18:02)



Now, who tells us the value of P? How do we choose the value? Without, already, we do not know many things; we do not know what are values of x 1 to x m, which optimize; now, you made the problem more complicated; we introduced additional scalars like P 1 to; for each of these constraints, there may be a Lagrange multiplier also. Is the solution worse than the problem?

Student: When the solution, which will behave, P equal to 0.

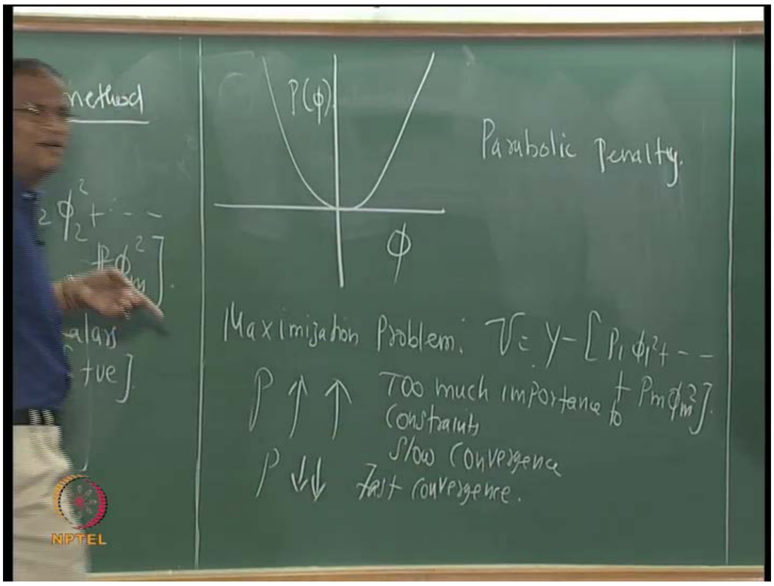
What is that?

Student: Sir, even for small evaluations, the penalty will be very huge.

Yeah, yeah, that is; that is, that is the intention; that is a intention, in the case of a, in the case of a inequality, because, I mean, it is non-negotiable, we do not want it to go less than this, some pressure limit or temperature limit; that is why, it is an infinite barrier is set. Here, if there is an equality constraint; equality constraint is also used. In fact, your c f d software, they use this; they use this penalty function method. For example, there is a penalty on not satisfying the continuity equation, how does your fluid work? Dou U by dou x plus dou V by dou Y equal to 0; dou U by dou x plus dou V by dou Y equal to 0, is a two-dimensional equation, continuity equation, for incompressible flow. This will never be equal to 0.

So, you set the limit, and then dou U by dou x plus dou V by dou Y, set it as psi, and thus, this psi whole square multiplied by P. So, this formulation is routinely applied in the petro of Galerkin finite element formulation, for solving Navier Stoke’s equation. So, the continuity equation itself is used to introduce a concept, by way of introducing penalty, or putting a penalty, assigning a penalty to the continuity equation, CFD is, you can develop your CFD algorithm.

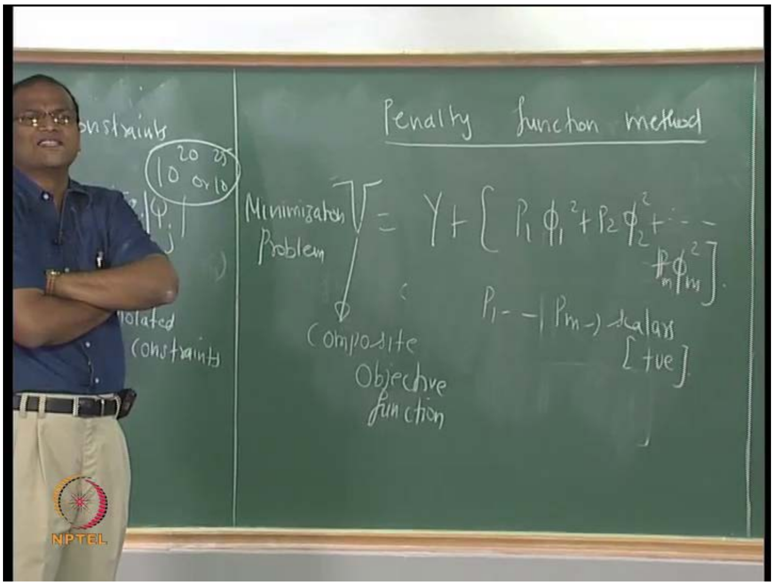
(Refer Slide Time: 19:41)



Now, who will tell me the value of P? Now, P very high; P very high means, we give lot of respect to the, we give lot of respect to the constraint; so, it will be very slow convergence; too much importance, slow convergence; this is fast convergence, but less importance to constraints. So, what you do is, in order to accelerate the convergence, I will start with small values of P. I will get some solution, then I will keep on changing the value of p; I will increase the value of P. And then, the solution will proceed in such a way, that the phi(s) are not violated too much; a stage will come when the phi(s) are satisfied more or less exactly, and regardless of the value of P, you will get the same value of V; that means, you have hit convergence.

So, it means, it is lot more painful to solve it, using a penalty function method, because you are going to try with different values of P. But, if it is a computer, which is going to do it, then it is very eminently and elegantly done on the computer. So, penalty function method can be ideally used for solving practical optimization problems, instead of, solving problem; I mean, for solving problems in the class, you can use any technique you want, but if there are many variables, you are looking at a real practical example, the penalty function method is a God send for us.

(Refer Slide Time: 21:40)



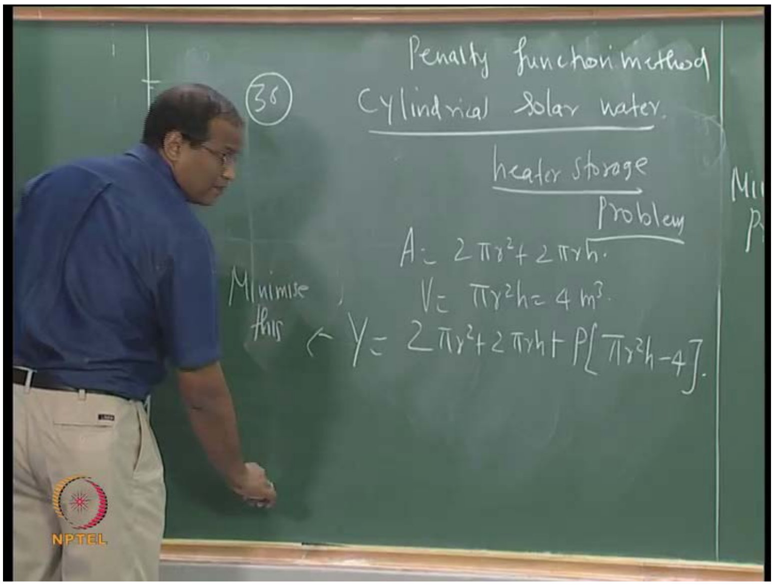
Now, we will revisit that cylindrical solar water storage heater problem. But now, I am not going to solve it with hand. So, we will try to demonstrate it with MS Excel. What we will do is, we will assign, there is only one constraint; it is a two variable, one constraint problem - radius and height two variables; one constraint is, volume is 3 meter or 4 meter cube.

Student: 4 meter cube.

4 meter cube. So, we will do that. What is the problem number? 37 or 38?

Student: 38, Sir.

(Refer Slide Time: 22:00)



This already, it is volume, right. So, I will not use V, I will use Y; minimize this. I want a volunteer now. You will come on NPTEL, man; or, you will be volunteered. Do not worry, if we get a wrong answer, we will correct it, the class will correct it; somebody you can copy, paste and this thing and you can do it fast, somebody.

Student: Blade.

Who is blade? Come, you are affectionately called blade.

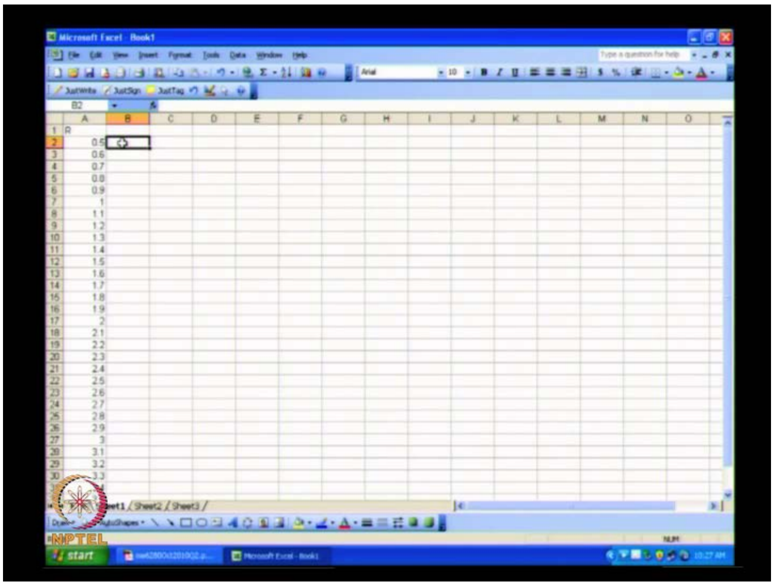
Student: Nickname.

Why, you are such a blade or what? Sit down-sit down, take the seat of honor. So, what was our range 0.5 to?

Student: 3.5.

So, Abishek, 0.5 to 3.5; you go in steps of 0.1 and then, repeat it 10 times.

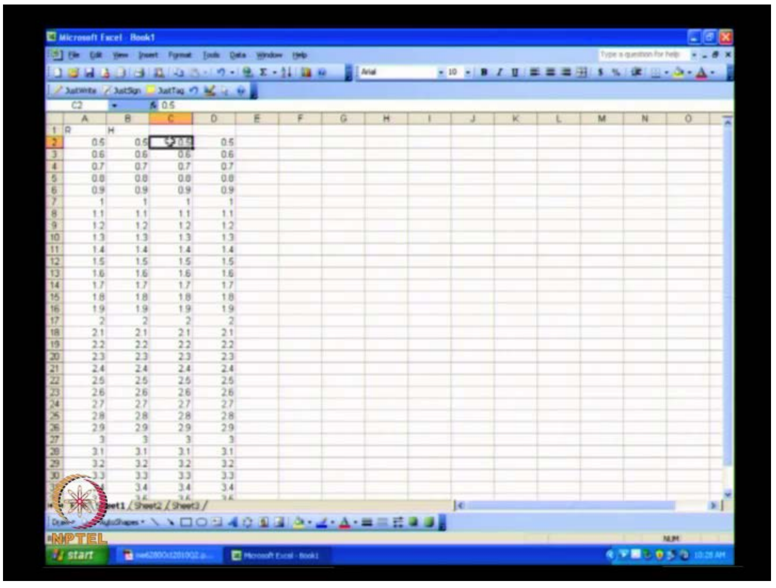
(Refer Slide Time: 24:08)



Student: 0.5 to 3.5.

0.5 to 3.5, yeah; repeat it, up to 3.5 you have to go.

(Refer Slide Time: 24:49)



Now, copy paste this, another 8 or 9 times, because this is, here, you can write it as r; a 1 you write it as r. We are trying to search out the solution. You know what I am doing? I am just putting various values of r and h; third column, we will put P equal to 1; fourth column, we will put P equal to 5; fifth column, we will put P equal to 20, and we will look at Y and, in 10 minutes quickly, we can; r; second one is h; no, no; now, the problem is, for each value of r, we should get each value of, you should get, you do that.

Student: For each value of r.

For r of 0.5, I want h varying from 0.5 to 3.5. So, how do you do that?

Student: Voice not clear.

Not 4 by, I do not want the constraint; I want it to, you first, Abhishek did you get it?

Student: Yeah, I get it.

Do not get h from that volume constraint that...

Student: You want, for each r, you want different kind of h.

Yes. So, I want at least;

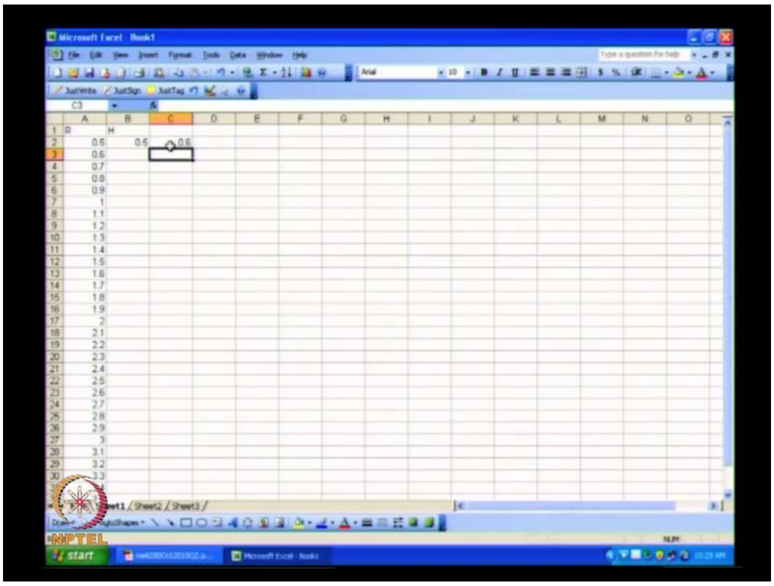
Will have some…

The best way is,0.5 to 3.5, you cut and paste, some 10 times; right side you vary it from 0.1 to 0.1, 0.2; is there any other way of doing it?

Student: Voice not clear.

All values of h.

(Refer Slide Time: 25:54)



You go ahead.

Student: h from

H is also 0.5 to 3.5.

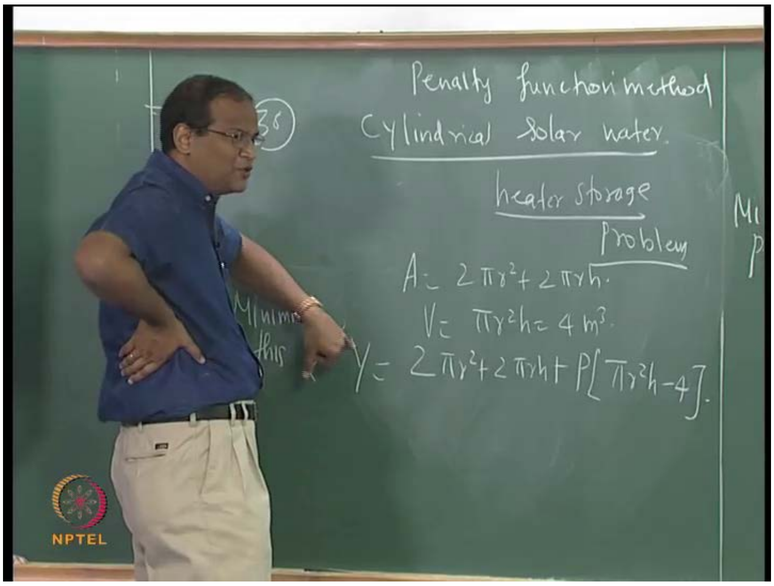
Student: Voice not clear.

No matrix, then it is difficult to do.

Student: No, I can do.

Then, how will you calculate the function?

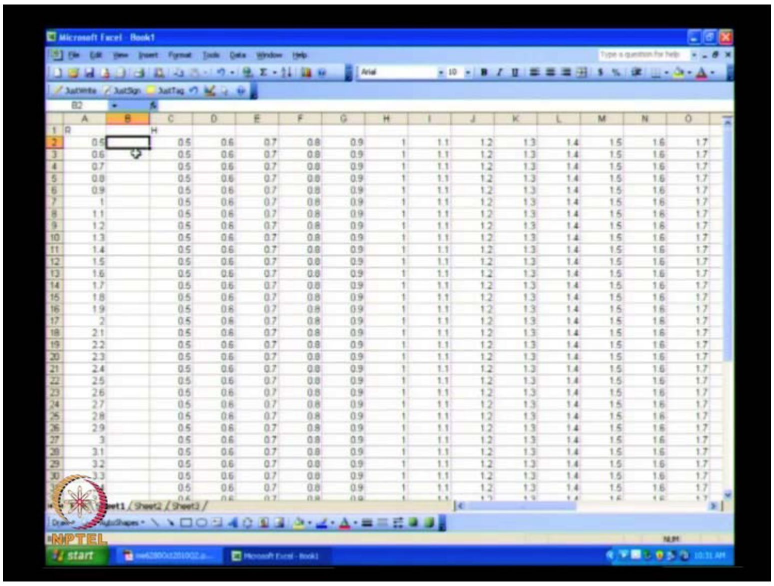
(Refer Slide Time: 26:23)



You should be able to calculate the value of y.

Student: We can, we can generate, and that is not a problem.

(Refer Slide Time: 26:28)



No time.

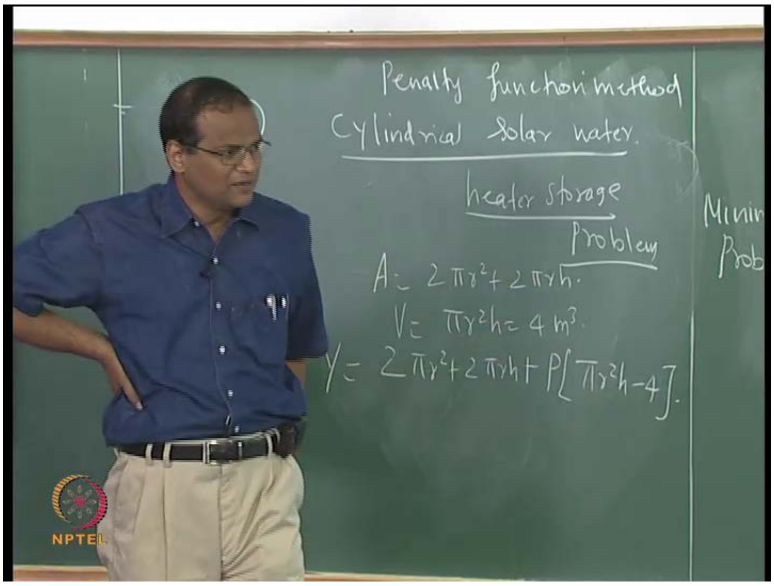
Student: No, it would not take time.

You got it, know? Yeah, now, this has to be completed. You have to calculate the function, Abhishek.

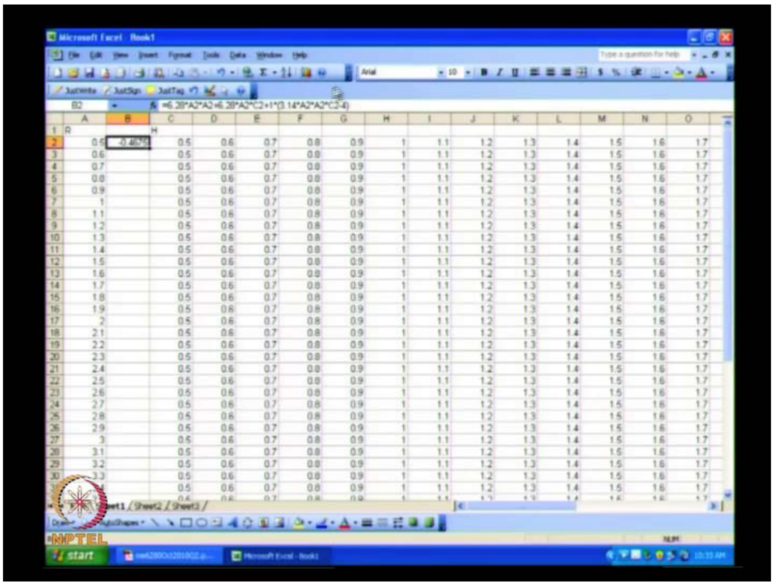
By the matrix, the function will be there? Then how will you search for the?

Student: You want each value for this.

(Refer Slide Time: 27:42)



(Refer Slide Time: 27:50)



I want you to do this, now; 6.28 a 1 square plus, 6.28 a 1 into whatever, plus P equal to 1, first. How do you do C1, d 1? How will you do that?

Do it, let us see, equal to…

6.28 times a 2 square, a 2 that thing also, plus 6.28 into a 2 plus, how will you do that h?

Student: That can be

You have to put that into C, what is it into? Not plus.

No, you did not do that.

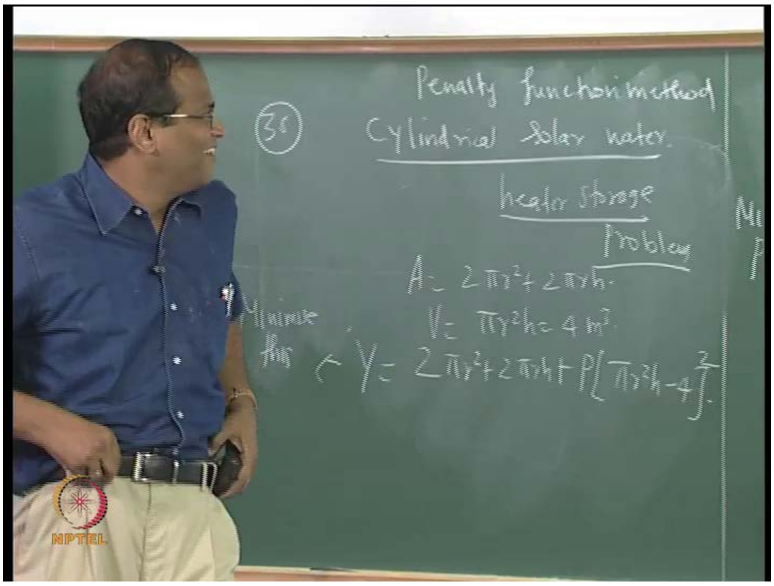
H is there.

Plus, one times, star open bracket; 3.14 star a 2 square, then you have to put brackets, it will work?

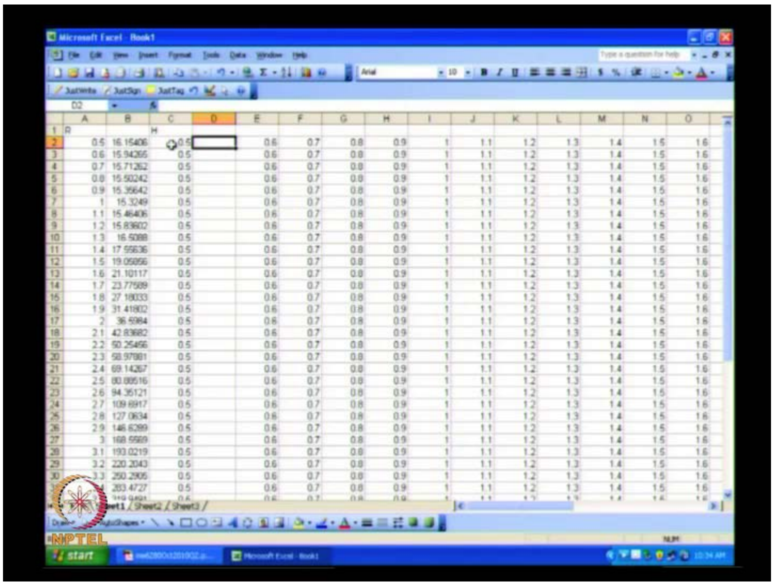
Into C2 minus 4.

Student: Minus 4.

(Refer Slide Time: 29:29)



(Refer Slide Time: 29:56)



Minus; into square; finally into; hang on; last term, here; I argued you; t into square of; that is why, it takes a long time to solve the quiz problems. Raised to the power. No, that raise is there, this one. Now, how do you find the minimum or maximum? No, you are copying it in all the columns? It is not a very efficient way of doing.

Student: Only one.

What happened there?

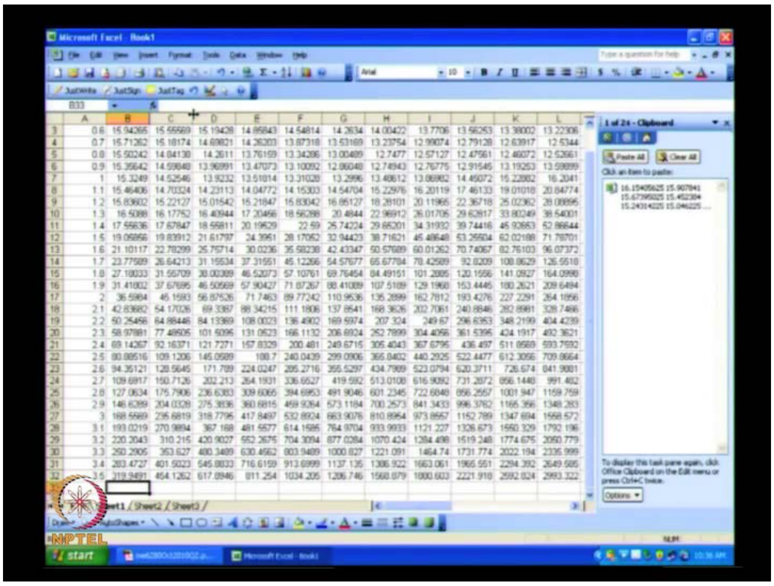
Student: No, just again, the values are same right, so the matrix will have that value.

B dollar 1 is correct.

Student: Yes, sir.

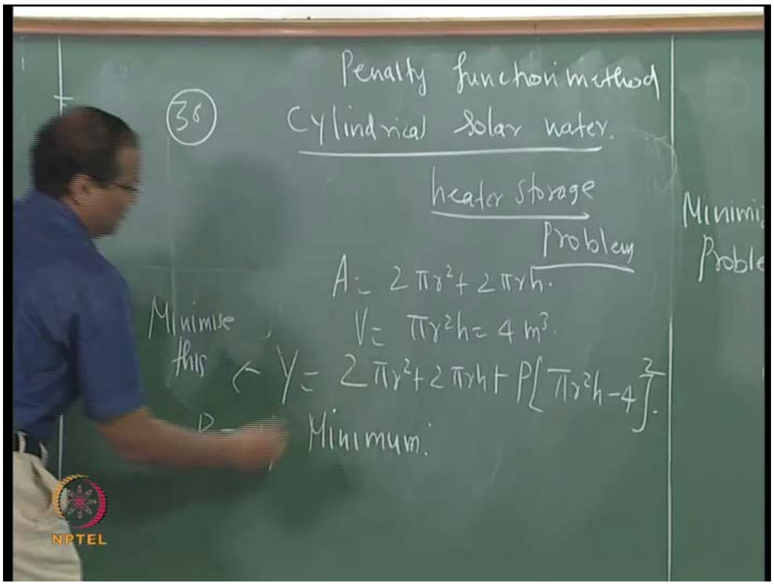
I have very basic technique, which works; so take 0.1, 0.5 to 3.5; keep on copying it, then right side, you just 0.1, 0.2, 0.3 then…

(Refer Slide Time: 32:15)



It is coming know; that is good; now, we again; now, tell me where is the minimum?

(Refer Slide Time: 32:56)



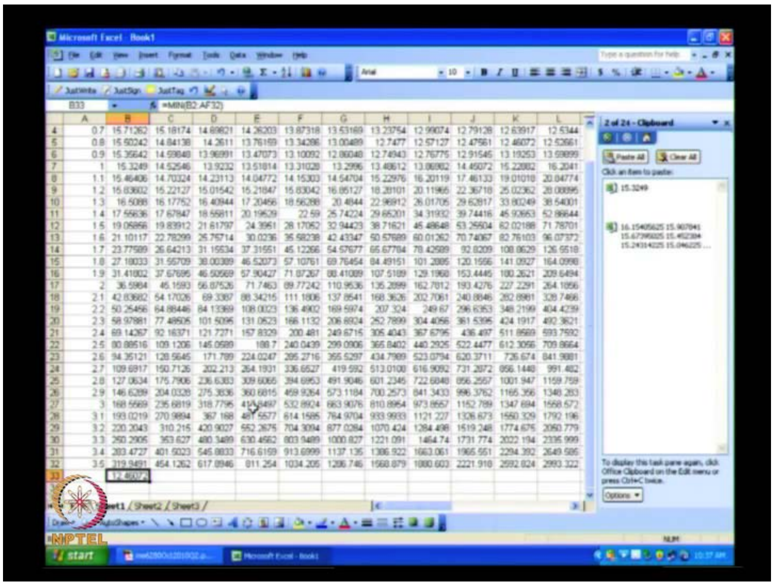
Student: You can put a

You can show, close up.

Minimum

Student: Voice not clear.

(Refer Slide Time: 32:03)



Anyway, we generated all the values.

B 2 to…

Student: B 32.

32?

You can do it only column wise.

Student: Voice not clear.

L 32, no.

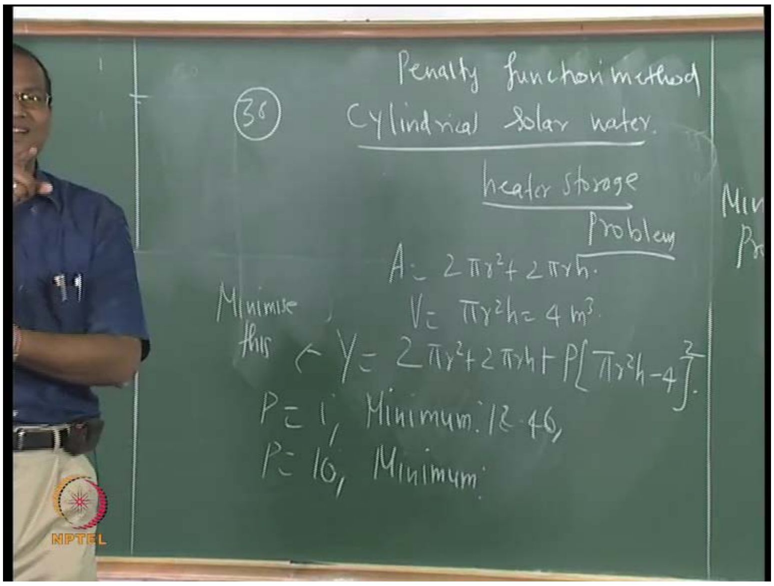
A of 32 name.

12, where?

There is the minimum.

Where?

(Refer Slide Time: 34:21)

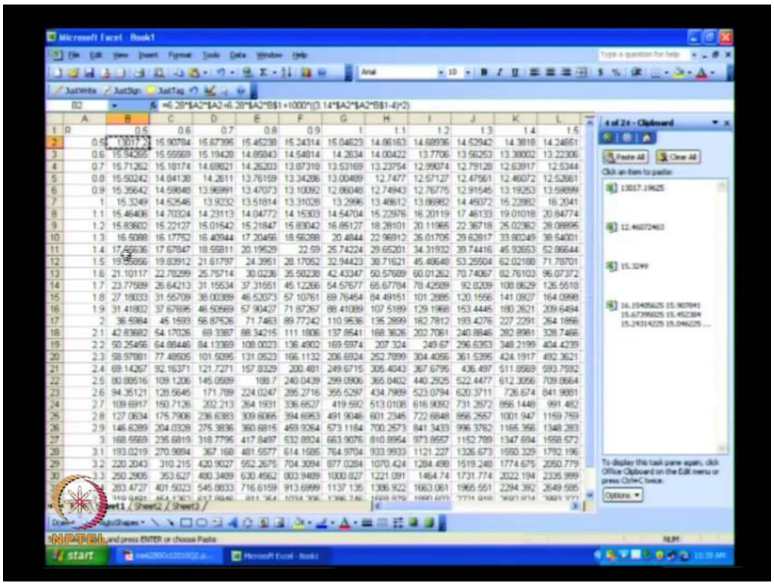


12.46, how is it possible? What are the solution to the problem?

What a…?

That is not possible. I think...

(Refer Slide Time: 34:34)



What are the solution? 0.866 and1.732, right. What is the value of, a minimum, we got?

What are the value of, a minimum?

You cannot get this.

Even otherwise, it is square know…

No, a itself is 14.9, you have to get something more than that.

It is not correct; something is wrong with the formula, just go to the formula.

That is, ok; but, that does not matter, there is something wrong with the; 3.14 a 2 square is;

Even, if penalty is relates always; adding know, it should be more than 12.96; even, if we hit the optimum; optimum was, a was 14.96.

It cannot be less than this.

It is straightaway, it is wrong.

Penalty does not matter man.

Student: Sir, at which are not feasible, I might have a very low value, which is...

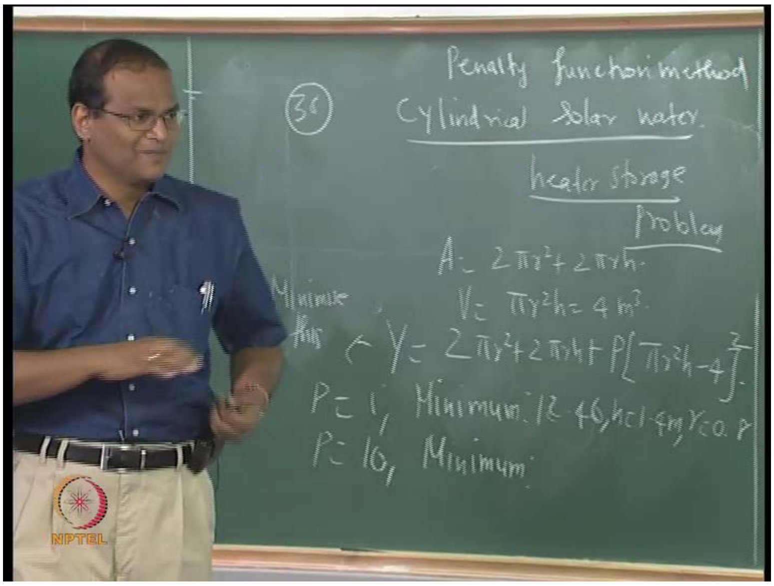
Well taken. What is, you can violate the constraint and still have low, because the penalty is low.

No, tell me the value of r and h?

No, for this solution, before getting confused, for this 12.46, what are the r?

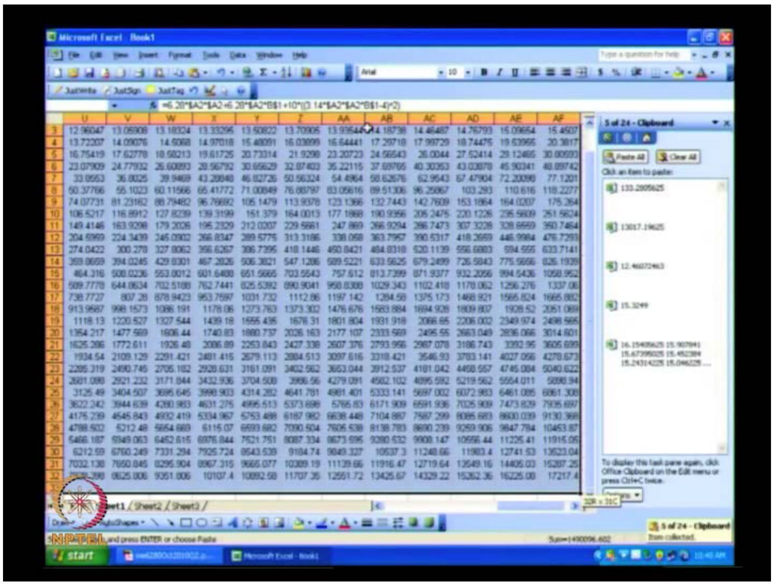
And, this is.

(Refer Slide Time: 36:19)



H is 1.4; r is 1.8.

(Refer Slide Time: 36:19)



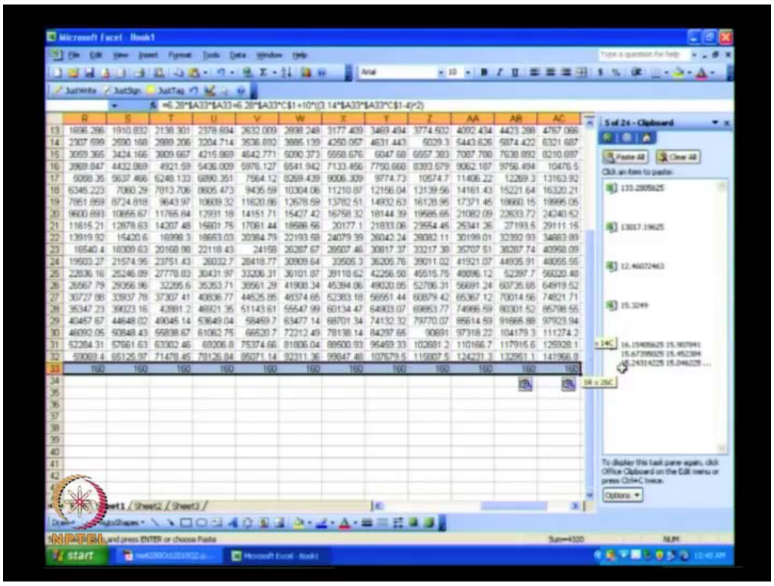
But Sampath, tell me the value of volume. Now, you change the penalty to 10; open it, can you do it with the same sheet? Change the penalty to 10; you make it 1000.

Student: 2.81.

Make it 10.

Recalculate.

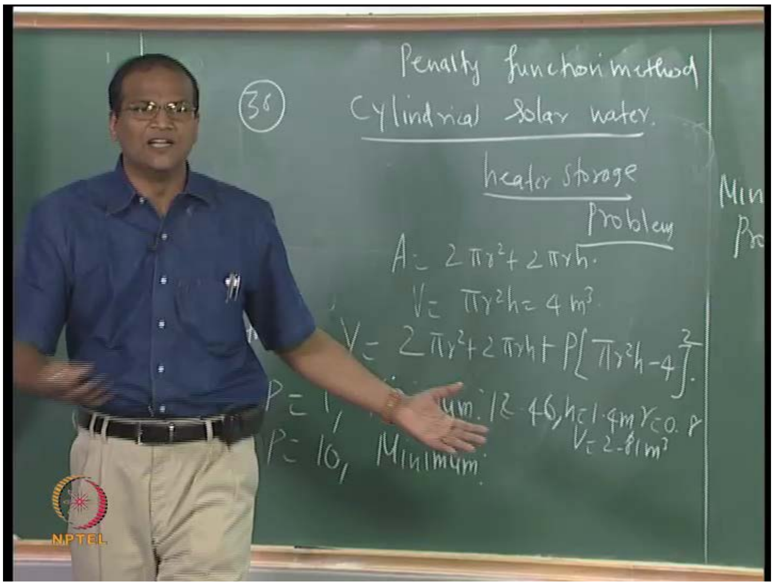
(Refer Slide Time: 36:53)



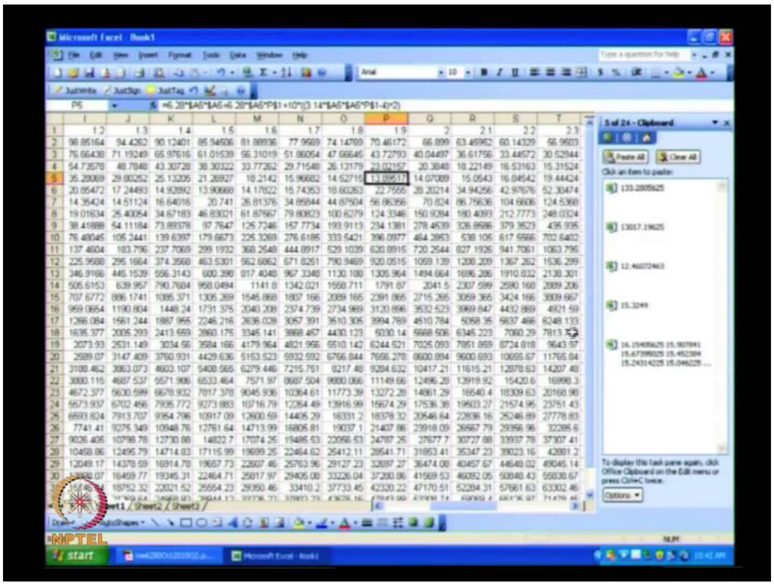
What are the values of the volume?

Student: 2.81.

(Refer Slide Time: 37:01)



(Refer Slide Time: 37:17)



So, the volume was 2.81; so, the penalty is, the FUNDS is, the penalty is so low, that you are deliberately choosing smaller values of r and h, but the volume constraint is not satisfied; this tank will not hold 4000 litres, when you can get lower surface area, right. If, 13.89, ok. So, now, this is, what is the value of r and h?

Student: Move to the right, go to 1.6 roughly.

How do you get?

Go to column n.

It it is row.

No, you no cheating.

You cannot search it out.

13.

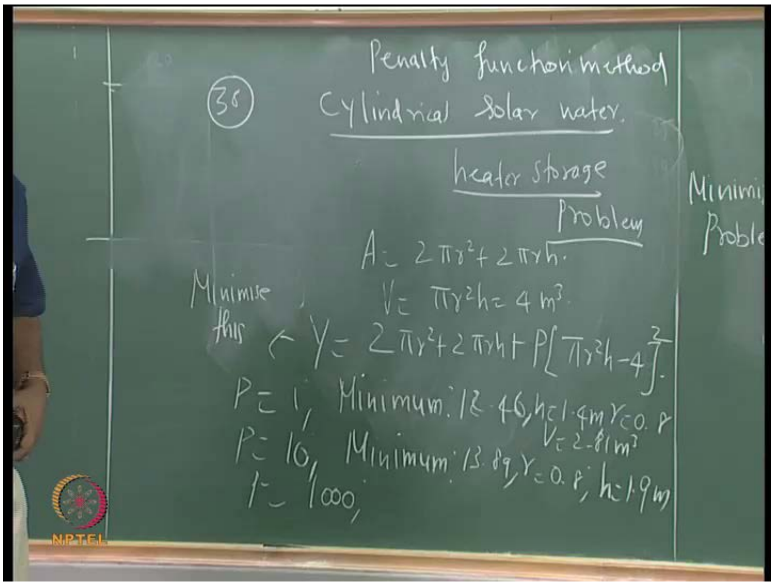
P 5 5.

You got this.

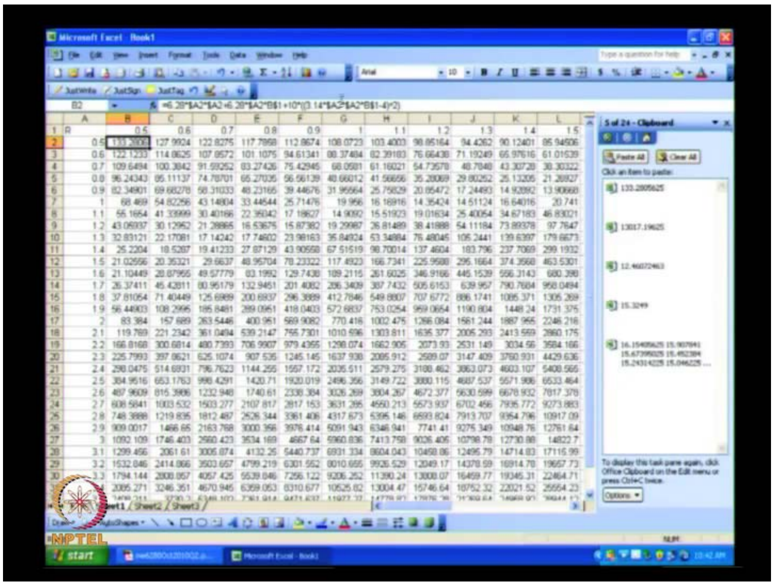
No, that is 13.90; it is here, 1.99 and, Go to the left side.

0.8

(Refer Slide Time: 38:01)

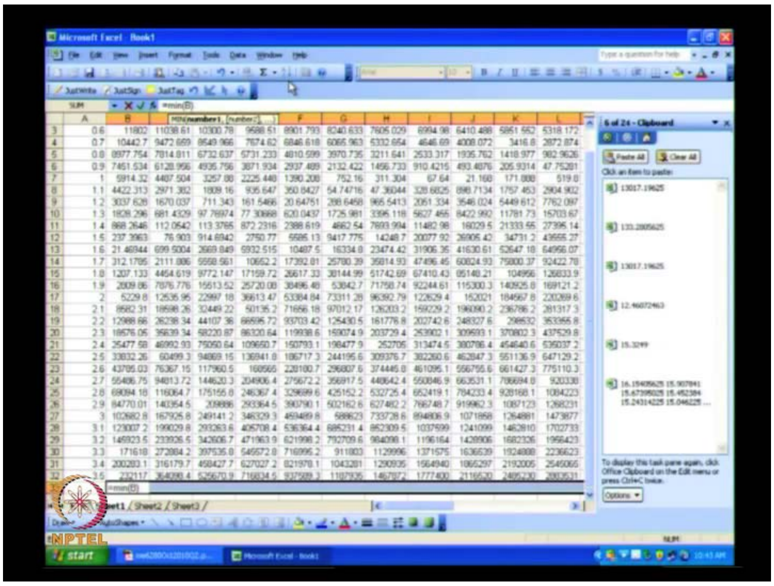


(Refer Slide Time: 38:42)



So, r equal to 0.8; h is equal to 1.9. Now, P equal to 1000; make it 1000. I hope, everybody able to follow the arguments; so, if we keep changing the value of P, finally the constraint, when the P so high, that pi r square h must be equal to 4; otherwise, you will get a very high value of V.

(Refer Slide Time: 39:00)

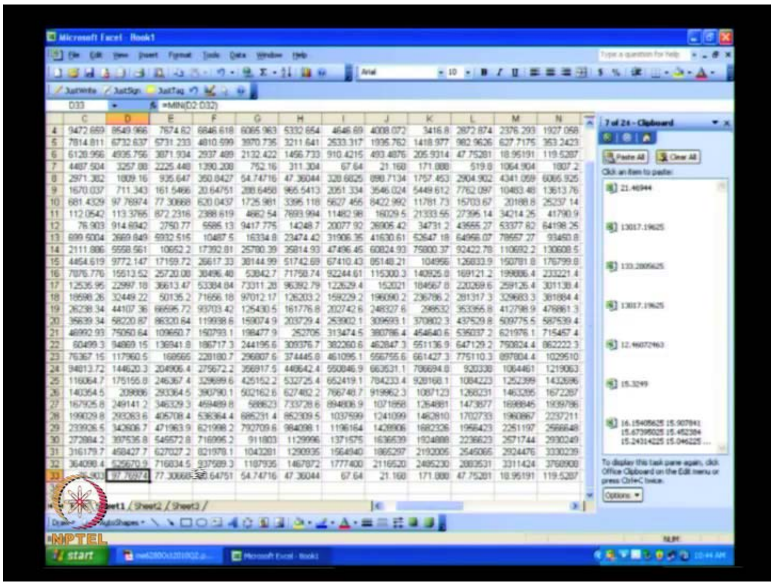


So, did you get the answer now? So, penalty functions is very difficult; obviously, very difficult for me to ask in the exam, you require computer. I can ask only the formulation, that is all, right. I again, just give the papers; we have to discuss the quiz solution, only on Monday, that you have well taken, know; cribs, after the…

Student: I can find it easier if I move it

But, 21 is not correct.

(Refer Slide Time: 39:52)



Student: I am doing it for each column; then I will do it for all the columns; that way I can find it also.

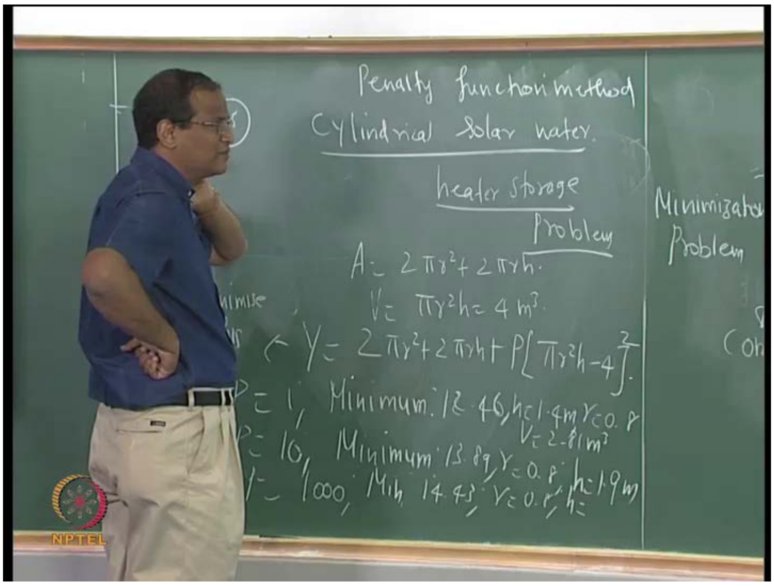
But, last time you are able to do.

No, no. Allow him to, he has some, he has way of doing it, that is good; so, you will find out the, you, will take the minimum of the row now.

Student: Yes.

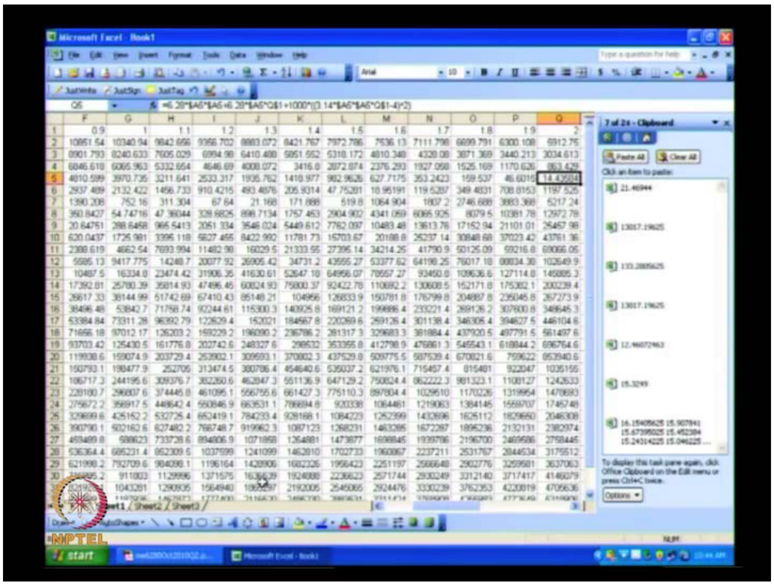
Not bad.

(Refer Slide Time: 40:54)

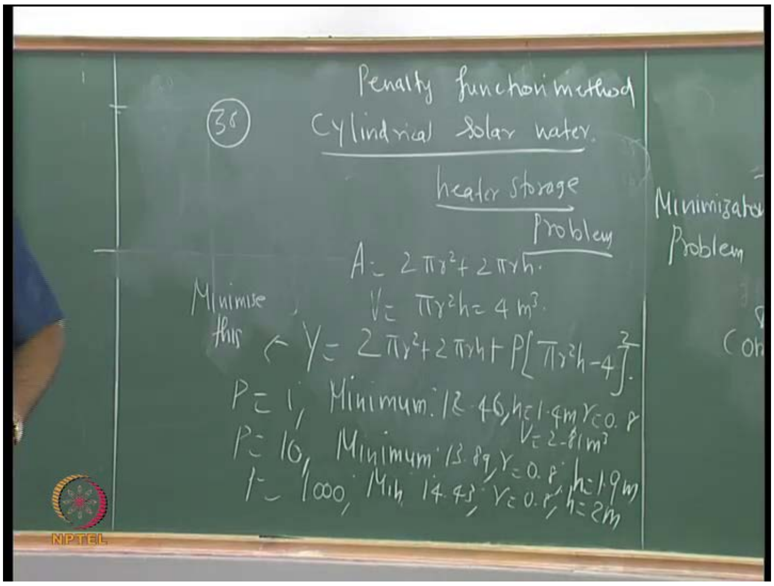


14 points. That is good; r equal to, it will stay at 0.8, because we did not use that 0.86; h is 1.7.

(Refer Slide Time: 41:05)



(Refer Slide Time: 41:12)



The actual answer is 0.866 and 1.7.

The problem is not with a penalty function, because we had a least count of 0.1. If you have a least count of 0.01, which may makes a matrix much more bigger, you will, you can resolve it; the problem is not the least count is only 0.1; the answer is 0.866, then the 2 will change to 1.7; this is how the penalty function works.

We did an exhaustive search, after doing this, right; you can do much more efficient searches, or you can solve it as a Lagrange multiplier. Using the Lagrange multiplier method, at each and every iteration, that is, for each value of P, and get the solution. Thank you. You want to do anything more or that is, fine.

So, this how the penalty function method works. So, it is pretty impressive. So, now, finally, we have come to a stage, where we can handle multi variable constraint optimization problem; not only equality constraints; inequality constraints can also be handled using infinite barrier penalty.

So, with that, we are more or less come to an end to conventional optimization techniques. There are small, there are other techniques like dynamic programming, linear programming, which are used basically by management people, operation research people and so on.

In the next class, we will look at the graphical method of linear programming and the dynamic programming. And then, the last 4 or 5 classes, we will just devote to unconventional or nontraditional optimization techniques. Two hours, we look at genetic algorithm. I will give a heat transfer example, so that, you get an idea of what genetic algorithm is all about. I will teach you simulated annealing, yet another powerful technique. And, if time permits, we will do neural network or basin inference and so on.