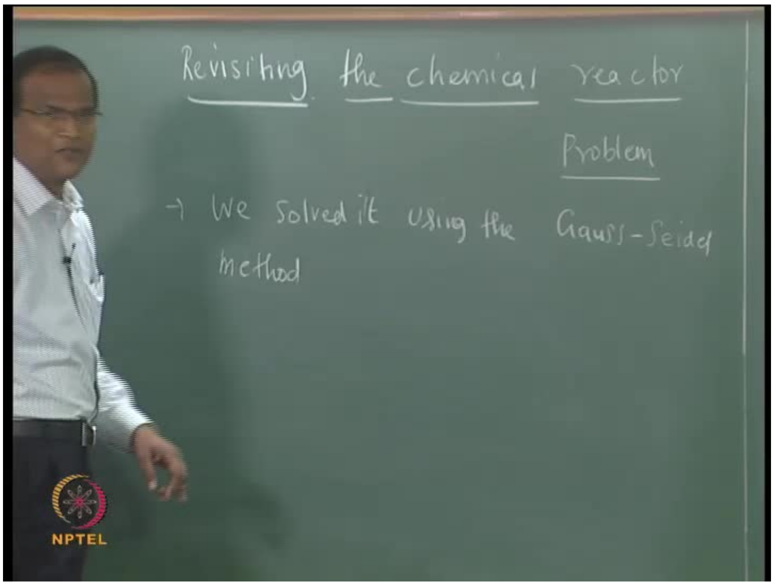
Design and Optimization of Energy Systems
Prof. C. Balaji
Department of Mechanical Engineering
Indian Institute of Technology, Madras

Lecture No. # 13
Introduction to Curve Fitting

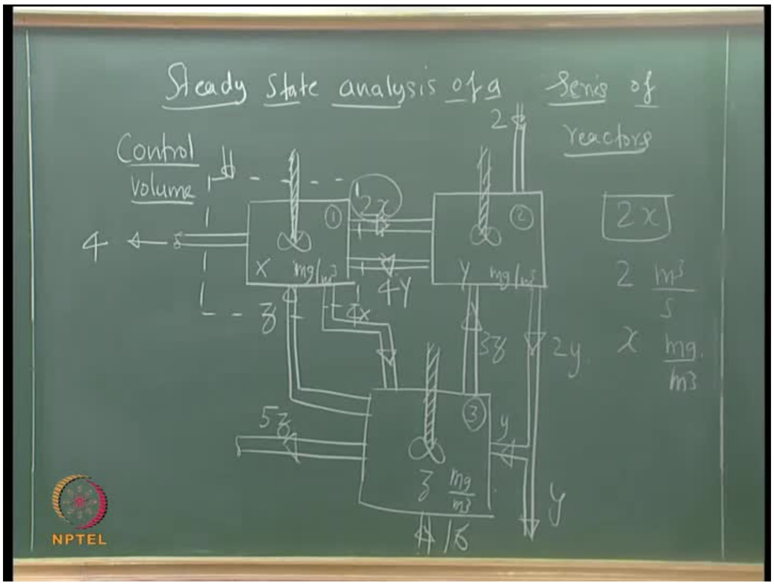
Good afternoon. So, we will continue with our discussion on Gauss-Seidel briefly and then we will go out to curve fitting. So, we will revisit the problem of the chemical reactor.

(Refer Slide Time: 00:26)



Here we solved it. Just to have a small recap, in the Gauss-Seidel method, we first rearrange the equations such that the diagonal dominance is satisfied. Diagonal dominance is a i i is greater than sigma of a i j, i naught equal to j is valid for all i, and then you write out x 1 equal to b 1 minus a 1, a 1 to x 2, and so on, and then you dynamically update the latest values of the iterates. If you do not dynamically update, it is called the Jacobi iteration. So, the Gauss-Seidel method is extremely potent, is extremely powerful, and is widely used by all these EFD software. For example, the fluent which works on the finite volume method, actually, finally you get thse variables like u velocity, v velocity, w velocity, temperatures, and so on at every grid point. Let us connect it to the neighboring grid point using this from finite volume or finite difference equations. These are simultaneously solved using the Gauss-Seidel method. Now, we will turn around and just see how one can get these equations.

(Refer Slide Time: 02:17)



So, let us say, steady state analysis of a series of reactors: Consider three reactors all of which are well stirred. Consider three reactors which are well stirred and we call these reactors as 1, 2, 3. Now, the concentration of the species in each of these reactor is x, let us say, milligram per meter cube. These reactors are interconnected by pipes and we are looking at steady state steady state analysis of a series of reactors. They are well stirred; so, there is a uniform concentration.

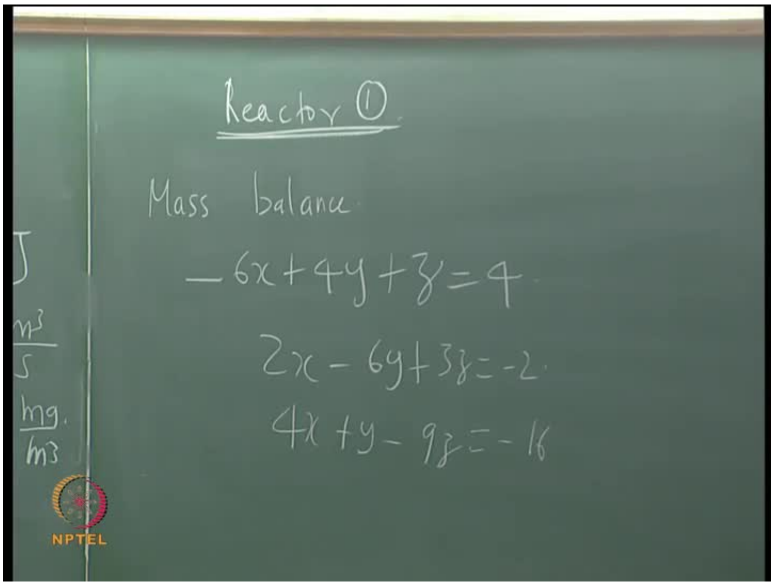
Now, I have a pipe and I pull out four units. So, when it is going, it is x; when it is coming, it is y; when it is going from here, concentration will be x; when it is coming, it will be y; I am saying that 2x is going from here, 4y. So, two units are coming from there; so, I will just hash this so that you do not get confused.

That is 16.

So, for example, this one, what is 2 meter cube per second? X will be milligram per meter cube. So, what are 2? The 2 denotes what? This 2x, it denotes some flow rate; x is basically the concentration, and concentration multiplied with the flow rate will give you the mass flow rate.

Now, if you do a steady state analysis, if you do a steady state mass balance for the three reactors 1, 2, 3, you will get three equations. There are only three variables x, y, z. Please draw a control volume around each of these; do the mass balance and let us see whether you are able to get something. Let us start with reactor 1. So, you can just draw.

(Refer Slide Time: 07:58)



So, reactor 1, what do you get? So, 2x is going out from here; 4x is going out from here; so, minus 6x; 4y is coming inside plus z minus 4 is equal to 0. Is this the first equation which we got in the last class?

So, after working out this problem, I went back to my office and I did the and I worked out the inverse problem. How should, if you suppose, the problem is, how should the three reactors be arranged in such a way that you get these three equations, alternatively, I could have just said steady state analysis of series of reactors shown in the figure. There are three reactors well stirred, the concentration of x, y, z, set up the governing equations for this problem and using the Gauss-Seidel method, determine x, y, z; is that clear?

So, otherwise, it will be just like a mathematics class, I mean 4x plus; now, it is a big puzzle from the equations getting back, I struggled for two hours in the last class, but I was trying till I came to the class, it was not working out; so, I gave it up. I do not want to give up; that is why today 15 minutes we will lose from this, but it is worth it. Please check whether you get the other two equations. What happened? You do not want to use these benches or what? If there is any small mistake let me know. Now, you know the goal; we should get the same three equations. Varun are you getting? Got it right? So, what is the second equation?

Student: 2x minus x y.

Where is minus x y? 2 and 4? x equal to minus 2.

So, that one?

Student: 4x plus y minus 9z equal to minus 16.

4x plus y minus 9z equal to…But this you are not reading out. This y is equal to?

Student: Minus 16

See, the beauty is, it should satisfy mass balance, but also be diagonally dominant. Of course, you can realize that this is not directly lifted from some text book. This is my own problem. Therefore, I struggle to recreate this. So, from the next year onwards, I will first give this and then ask them to set up the Gauss-Seidel system, and then solve it. Now, you got an idea of why Gauss-Seidel method will be useful or rather I will put it more straight away.

Where can you expect linear system or equations in engineering? If you are doing reactor engineering and chemical reaction engineering, all this including (()) and this thing fractionating column, petroleum engineering, and all these things, we will get this. In fluid flow take it is also you will get. You designed the apartment complex, the fan and that is what. I have already graded your assignment. For example, there are several branches; instead of an apartment complex, you are working out the distribution for a city itself, then there will be places where it is pumped again and so on.

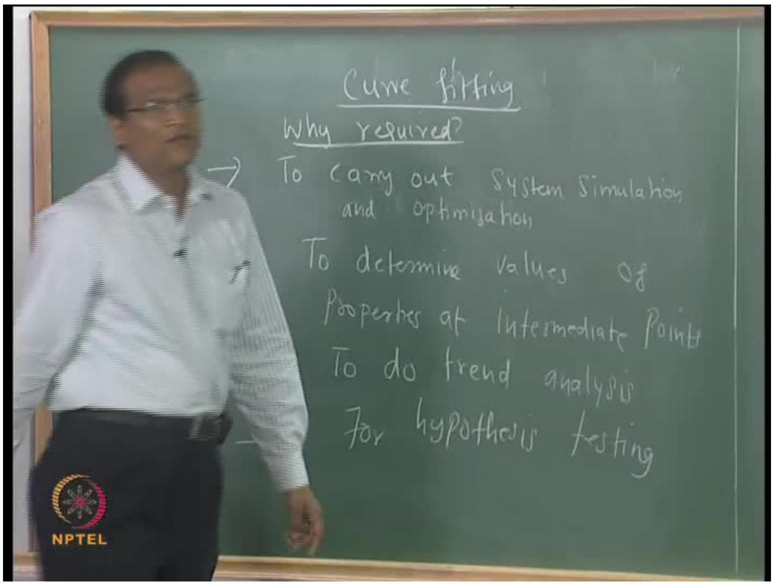
For example, do you know that in Germany nobody has a overhead tank? The water reaches the 5th floor of apartment complexes directly. If open that tap it comes; so, they worked backwards and found out how much pressure is required so that the water will automatically flow. So, here, it is just above the atmospheric pressure. Therefore, we have a motor and pump and we build a overhead tank and everything is collected at the overhead tank; the water comes out in such a high pressure. In the developed countries, if you open the tap, it will automatically flow; they are unheard of UPS, overhead tank, all these are unheard of. They would not know what; a German will not understand what a UPS is because I mean he has never seen a power cut.

Is it uninterrupted or uninterruptible?

Student: Uninterrupted.

So, with this, we close the discussion on Gauss-Seidel method. Now, I will go to equally interesting topic - curve fitting.

(Refer Slide Time: 13:40)



We are talking about authentic and genuine curve fitting; we are not talking about conducting experiments, where you first draw the curve and then get to draw the point; that is the special category of...; so, this is curve fitting 1 over 1; for example, 1 over 1 basically universally 1 over 1 course is the first course; so, this genuine curve fitting 1 over 1. So, why curve fitting at all? Why study curve fitting? Why curve fit at all? Why? Why do you want (( ))? Why cannot it keep quiet?

You take data points from the values; you take the experimental values.

Extrapolate or interpolate?

Student: Interpolate.

Why do you want to interpolate?

Student: Experiment.

What is the goal of the interpolation?

And do what? Write a book?

You are all close, but not close enough.

You want to get the properties of function at unknown points. For example, if these functions are available only at the known number of points, but you want this function that some other points, that is interpolation. Sometimes the extrapolation is also required. For example, you want to do forecasting; forcasting of a demand of a product; forecasting of the monsoon; forecasting of the rainfall; whatever, whether forecasting is the economics or atmospheric science, whatever, now, you need to establish a relationship between the independent variable and the various dependent variables, and you have to prove your hypothesis. That means, out of this, suppose your data is collected over 40 years; you will probably develop a model for 35 or 36 years; then you will start testing on 37, 38, 39, and 40; if it is works, then you say that I have built a model taking data up to 35 for first 0 to 35 years; 36 to 40 it is working reasonably well; therefore, there is no reason why it should misbehave on the 41st year; therefore, I will start predicting for the 41st, 42nd…, that is basically extrapolation of forecasting.

Interpolation, often times, for example, thermodynamic properties and so on, they are measured only at some finite values of pressure and temperature, but you want enthalpy, entropy, and other values of pressure and temperature. So, you will need a method by which you will get these properties. Therefore, that is only one part of the story. In relation to this course, why do you want curve fitting? You cannot, you do not know the exact optimum point. So, you find somewhere two points before the optimum and after the optimum and somewhere

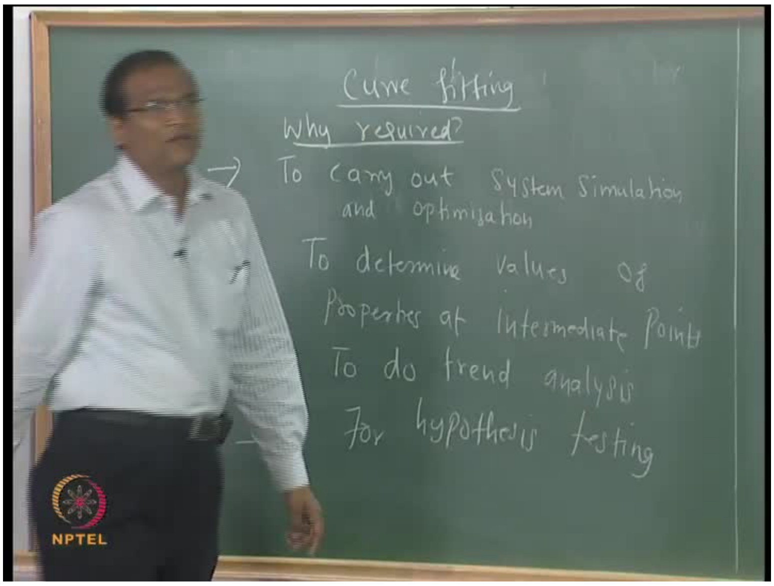
No. Even now fundamental than that, for example, when you wanted to do system, when you wanted to do getting systems, when you want to do get the operating point, what did you start off with? You had of a torque speed characteristics of the load, torque speed characteristics of the engine and all that. Who will give you all that? Who will draw those curves for you? Who will give those equations? Nobody will give those equations. People will give some points, but with these points we cannot hopefully, I mean some optimization techniques will work; discrete optimization and all that, but generally it is more convenient for us to work with functions. Therefore, the first step in technology is to look at these points and first draw an appropriate function; that is the best possible function.

So, already when you are doing curve fitting, some optimization is already inherent in it because you are looking at the best. Now, I will tell you the difference between the best and some other curve. For example, if you got 1000, if you got some 500 data points, we got some 500 data points, it will be absolutely meaningless for you to have a curve which possess through all the 500 points because you know that this data, error is inherent in this data; when error is inherent in this data, then you want to pass a line which gives the minimum deviation from all the points in a least square sense; all that will explain little later. Therefore, now, I have already introduced some more additional concepts to you. Curve fitting does not mean that the curve has to pass through all the points, but there can also be cases when the curve passes through all the points.

So, in the light what I have told you just a minute ago, when will you have an exact fit? When will you have an approximate fit? When the number of parameters is less and you have got absolute faith and confidence in a measurements then go ahead and have an exact fit, but once you have more number of data and if you are using a polynomial, then the order of the polynomial keeps increasing. What is the basic problem with higher order polynomials? Equation

That is now granted that we can handle complicated equation. The polynomial tends to get very oscillatory and if you want to work with the additional things like derivatives and so on, it will become very very messy. Therefore, if we are talking about high accuracy data, limited number of data, it is possible for you to go ahead and do exact fit, but if you are dealing with the large number of data points where it is error proven, it is okay to have the best fit. So, already curve fitting can be of two types: best fit and the exact fit.

(Refer Slide Time: 20:10)



Now, why is curve fitting required at all? So, why required? To carry out system simulation and optimization; to determine values of properties, for example; you also want to do what is called a trend analysis. You want to make sense of the data; what trend is it following? Is it following an exponential trend, a linear trend? So, it is to do trend analysis.

Sometimes, logical extension is to also do hypothesis testing; for example, it follows some basic model of economics; it follows some basic model of micro biology; it follows some basic model of fluid mechanics - heat transfer; whether I have carried out experiments, whether the measurements, if I re-plot it or draw a curve, whether it is useful for testing a hypothesis; so, for hypothesis testing.

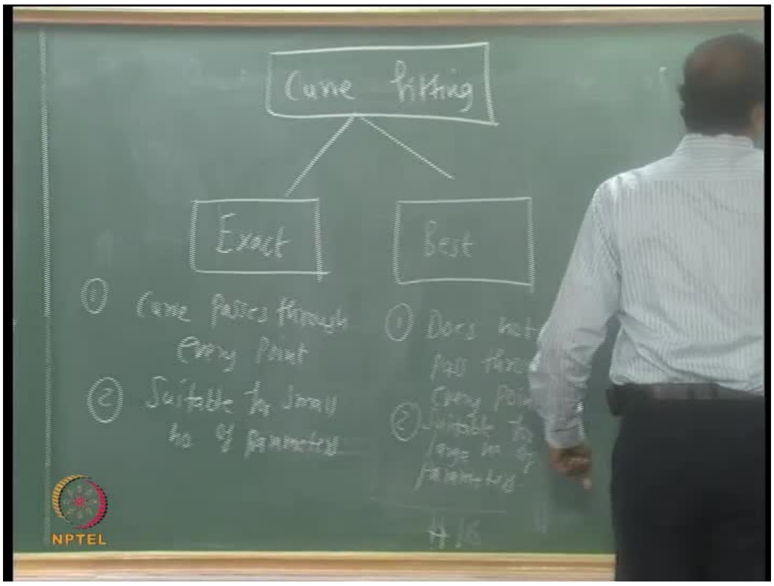
Whether if you take a course like ME 110, if you plot the marks of all the 800 people, the hypothesis is it will follow Gaussian distribution. Whether that hypothesis, whether the test is okay? If you do, you can actually plot a curve. What happens? When instead of doing for the whole class, you do it separately for A batch, B batch, C batch, D batch, and then you can split again dual degree students. So, if you keep on, then you can, then you can eventually, you can either write a paper or you can go mad to…

So, in this case, I want to test A is because Central Limit Theorem says that all data, when the number of points become too large, it follows the Gaussian, normal distributed. So you can use it for…, you can actually plot, draw curve, and check whether it is this thing; find out the two important parameters of the distribution; what are the two important parameter of the Gaussian distribution?

Student: Mean and variance.

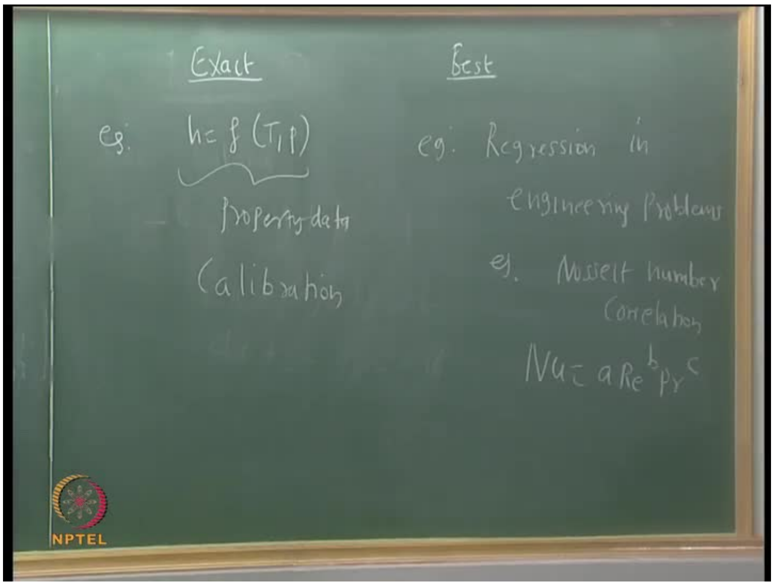
So, the mean and variance; so, these are the two parameters, two important parameters of the Gaussian distribution. So, in order to able to do all this, you need to know curve fitting.

(Refer Slide Time: 22:55)



Curve fitting: Exact and in exact or approximate or best. Exact: curve passes through every point; suitable for small number of parameters, suitable for large number of parameters.

(Refer Slide Time: 24:26)

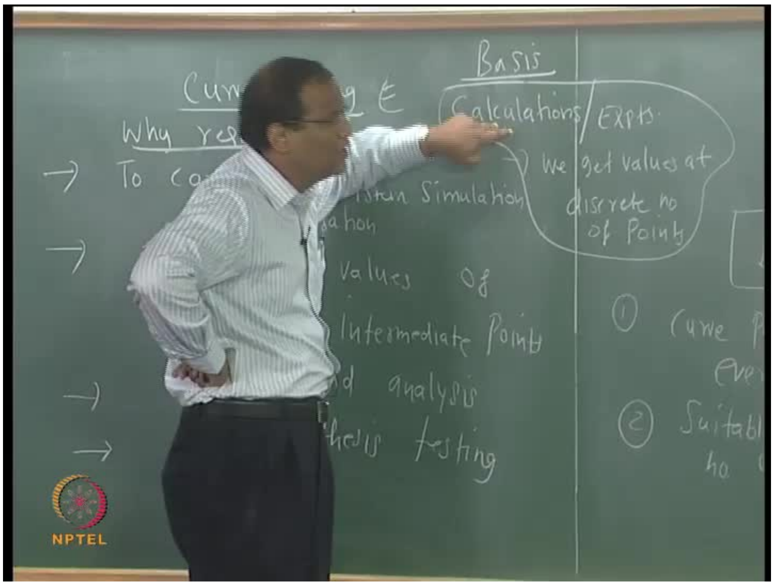


Example, so h is the function of temperature and pressure; of course, we did not use regression in ME 110 that is Thermo Dynamics; ME 110 is now it is called ME 1100, the first course in Thermo Dynamics. There we gave you tables, but it is very messy; if you want to write the program, you cannot use tables. So, it is better to convert all these tables into equations form. For example, now software, various kinds of software is available; I mean, you can you give just give the temperature and pressure and automatically give the enthalpy and all those days of tables and all that gone, but tables are still useful for the class room environment. That is why we still continue to use that for the first year course at least.

Now, if you want something, enthalpy is equal to function of temperature and pressure, so, property data like this; you can go for an exact fit. Calibration data: what is that calibration data? EMF thermo couple; EMF for system temperature where everything is highly accurate, then you can have an exact fit.

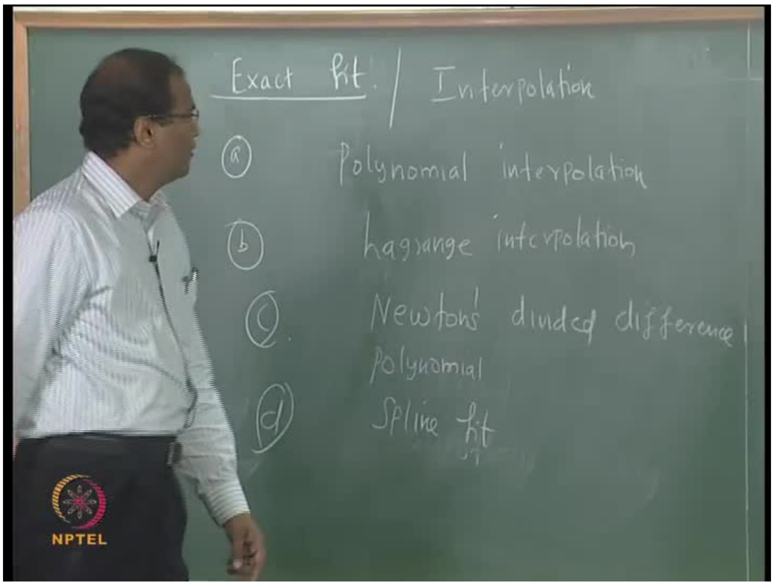
Any regression? Example, Nusselt number correlation; you consider flow over a cylinder; you heat the cylinder; so, in cold air; you do experiment under steady state; measure the total amount of heat which is dissipated in order to maintain the cylinder temperature constant; increase the velocity; for every Reynolds number, find out the heat transfer rate, heat transfer rate using Newton's law of cooling; convert the heat transfer coefficient; from heat transfer coefficient, get the dimension less heat transfer coefficient called the Nusselt number; plot the Nusselt number versus the Reynolds number; if you so like, change the fluid from air to water, from water to liquid metals and so on. So, now, if you know the correlating variables, for example, Nusselt number goes as a Reynolds to the power of b Prandtl to the power of c. If you have sufficient number of data, it is possible for you to get a, b, c from experiments.

(Refer Slide Time: 27:36)



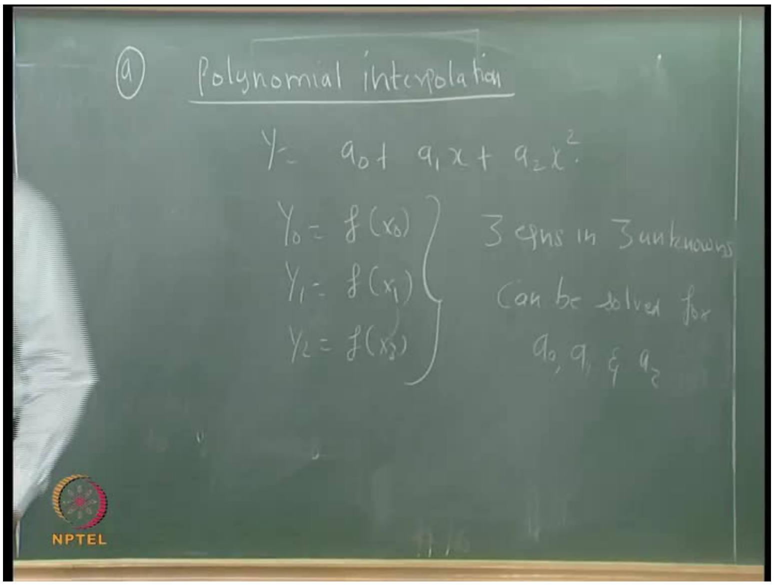
Now, why required to carry out and all these? The basis for all this is calculations; regardless of whether do calculations or experiments, you get values only at discrete number of points. Correct? However, my idea of doing experiment and calculations is not just to get values at discrete number of experiments. I know that as a scientist, if we do the experiment carefully, I will get some five values for Nusselt number for five values of Reynolds number, but I am not going to stop there. I am going to have a predictive relationship like this which I can use for optimizing the whole systems and so on. I want to do systems simulation and then I want do optimization. I am not just satisfied with trying to find how the heat transfer varies with Reynolds number. That is that is one over one kind of thing, but I want to do more. I want to do system simulation optimization, but this is the limitation I have; therefore, there is an intermediate step which is in is involved, which is which requires you to convert all this into equation form which can be played around with and you can use your simulation or optimization. Is it… So, this gives a broad overview of the two widely used classifications namely exact fit and best fit, the rest of today's class we will look at some possibilities of exact fit and will also solve problems.

(Refer Slide Time: 29:46)



Polynomial interpolation, Lagrange interpolation, Newton's divided difference polynomial, and your spline fit. The exact fit can also be called as an interpolation.

(Refer Slide Time: 31:22)

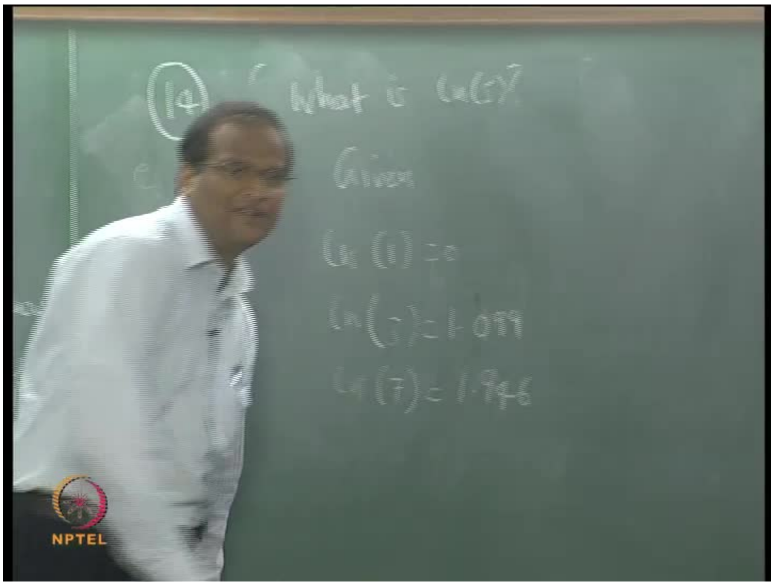


Let us start with a; so, I propose that y varies quadratically with x. So, I write the relationship as y equal to a naught plus a 1 x plus a 2 x square; a naught, a 1, a 2 are the three constants; so, in order to get three constants I need the values of y at…, how many values of x? Three values so, it will exactly satisfy. There will be no error when you applied this formula at x naught, x 1, and x 2. There will be some error when you apply it in between x naught, x 1, x 2 because we are approximating the whole physics by a, we are approximating the physics by a quadratic. You can only force three points: you can draw a parabola; I mean you can draw quadratic; you can make a quadratic pass through; you can only do that much; every point in between you cannot force because after all you are representing some physics; let us start off. Let us have a warm up.

So, problem number 18 is it?

Student: 14.

(Refer Slide Time: 33:48)



Evaluate the natural logarithm of 5, problem number 14. We will start with some simple mathematical problem; then we will go to a thermal system problem. Evaluate the natural logarithm of 5 using quadratic interpolation. Evaluate the natural logarithm of 5 using quadratic interpolation. Evaluate the natural logarithm of 5 using quadratic interpolation given: (Refer Slide Time: 34:45). Please report the error. The exact value you can get from the calculator.

Vinay, what is the problem?

Is it 2 point? What is it?

No, there is mistake here what is (( ))

Student: Lon 3 is 1.09

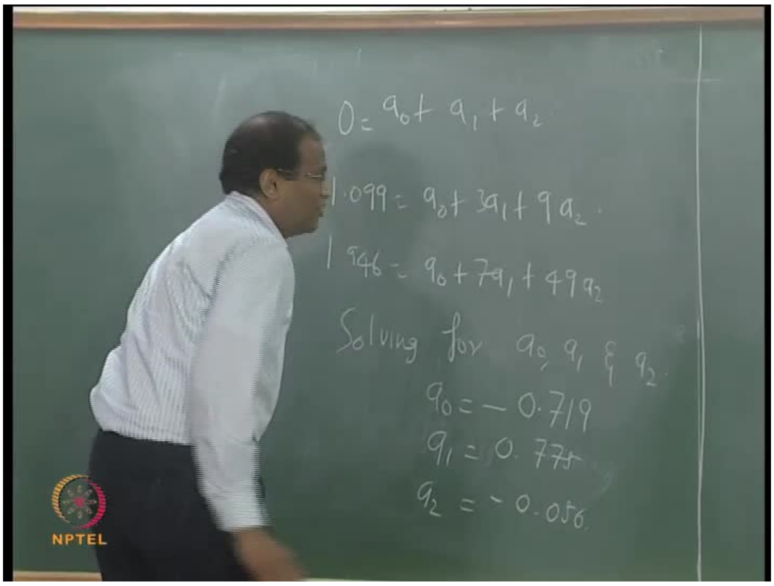
I do not want lon 3, what is lon 7?

Lon 7 is correct; lon 3 is wrong

This is the mistake. Anyway, what is lon 5?

What do you get? You will not get the same; so, what you get? Minus?

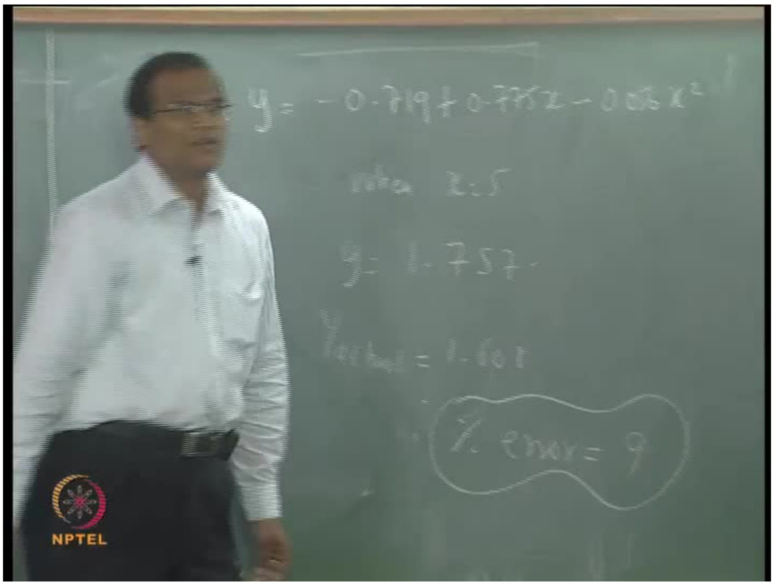
(Refer Slide Time: 36:13)



1.099; x is 1; it is correct, right?

So, a naught is minus point correct? a 1…, 0 5; that is it.

(Refer Slide Time: 37:55)



So, linear interpolation you have done in thermo dynamics; polynomial interpolation may be is new to you; It is not very difficult, but now, if you do, therefore y is equal to minus 7.

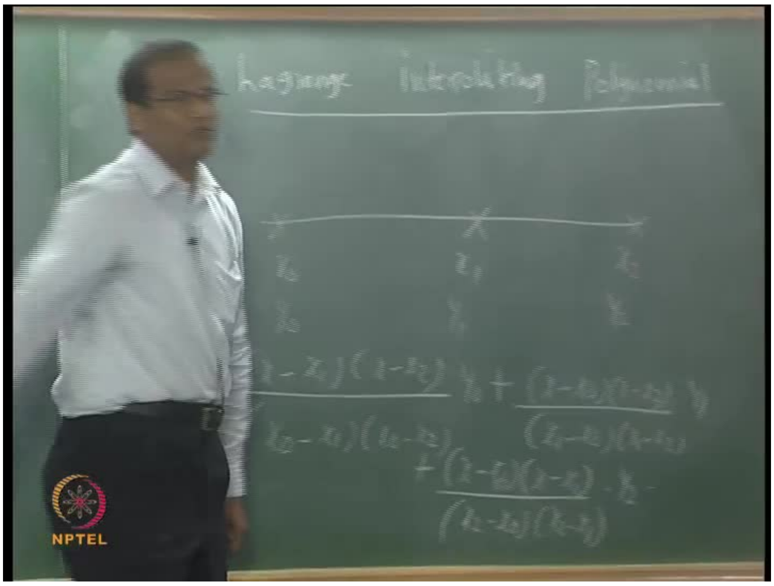
9? It depends how you…

Percentage error 9.

A simple example, it gives, demonstrates, it will let us understand how this works. Of course, we can complicate it by having cubic; if you have 8 points, you can have a 7th degree polynomial and all that, but what happens is beyond a certain, beyond cubic in the polynomial it tends to get highly oscillatory. It is not advisable. You may rather go in for what is called cubic's plane or a quadratic plane, where locally three points you will fit with the lower order curve. So, that generally gets a smooth curve.

So, if I want to…, if you have a highly accurate calibration data, thermo couple data, EMF versus temperature and so on, then you can use this. Then, the next one I am going to teaching you is most powerful Lagrange interpolating polynomial.

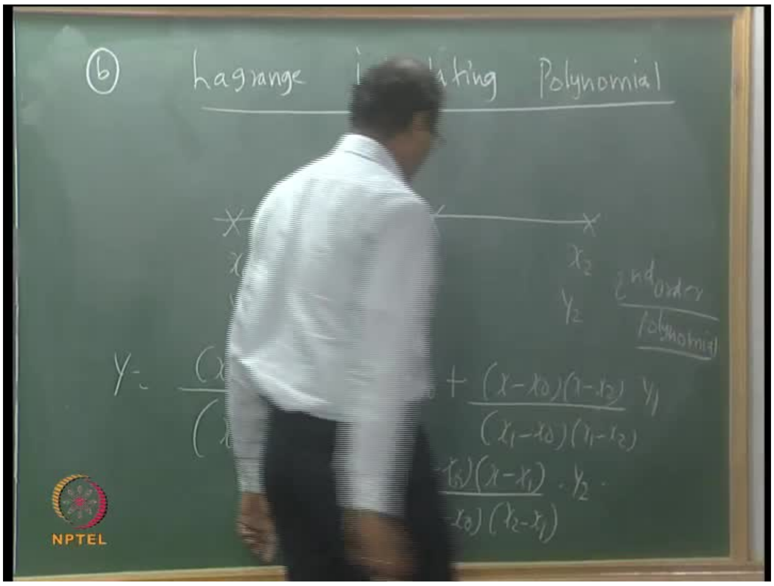
(Refer Slide Time: 39:52)



So, I will start with a second order Lagrange interpolating polynomial. For example, if you have, you know that you can use the interpolating polynomial which you just studied now; the polynomial interpolation for this but the Lagrange interpolating polynomial is much more, what shall I say, it is much more potent in sense that it can be easily extended to any number of points; it is easier for it take the first derivative, second derivative, and so on, and a much of the I mean, for example, the post processing of CFD work and so on. After we get the velocity and temperature, in fact left right and center, we all use the Lagrange interpolating polynomial to get the gradient of temperature and the gradient of velocity. Once you get the gradient of temperature, we can multiply with thermal conductivity to get the q from q. Using Newton's law, you can get h from h; we can get Nusselt number; then we correlate from the Nusselt number, and so on; so, it is idea of post processing. So, what is this Lagrange interpolating formula.

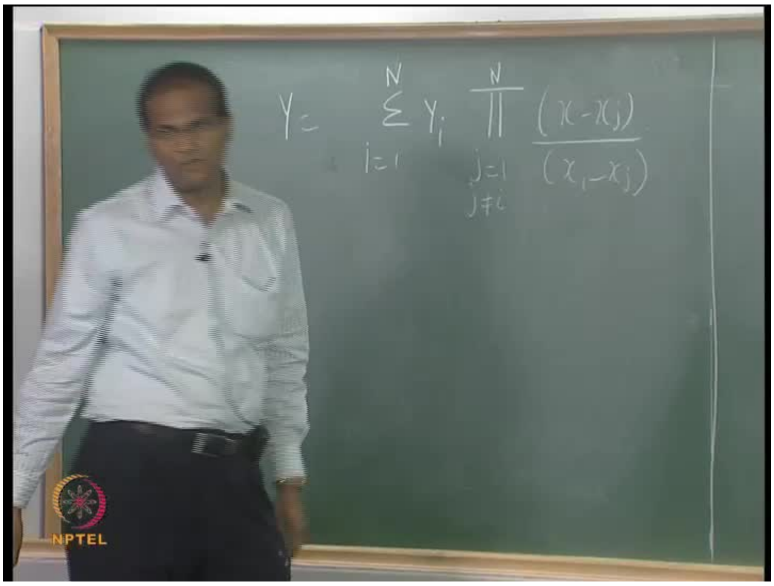
So, given x naught, x 1, x 2, they can be unequally spaced; there is no need that they have to be equally spaced and corresponding values of the dependent variable are y naught, y 1, and y 2. The Lagrange polynomial is like this y (Refer Slide Time: 41:43)

(Refer Slide Time: 42:35)



Lagrange interpolating formula: This is called the Lagrange interpolating polynomial order 2. So, it starts with this. Suppose, you are looking at this, so you do x minus x 1, x minus x 2 divided by x naught minus x 1 x naught minus x 2 into y naught plus x minus x naught and x minus x 2 divided by x 1 minus x naught x 1 minus x 2 into y 1 and so on; that is the way you do.

(Refer Slide Time: 43:25)



Can we write this compact mathematical form? Can you try? Y phi of x minus x j divided by…, here?

j is equal to

That phi is for multiplication for people who are unfamiliar, that phi is mathematical this thing for multiplication. So, this can be written in a sink or compact from like this. So, it is possible for you to get first derivative, second derivative, possible for you; the first derivative will still be a function x; the second derivative will be a constant. The second derivative will not be a function of x and so on. This is for a second order polynomial. You can go up to third order polynomial where you can take y naught y 1, y 2 and y 3. The next class, we will solve a heat transfer problem with the Lagrange interpolation formula and then we will go on to Newton’s, I mean divided difference polynomial and then we will round it off with cubic planes.