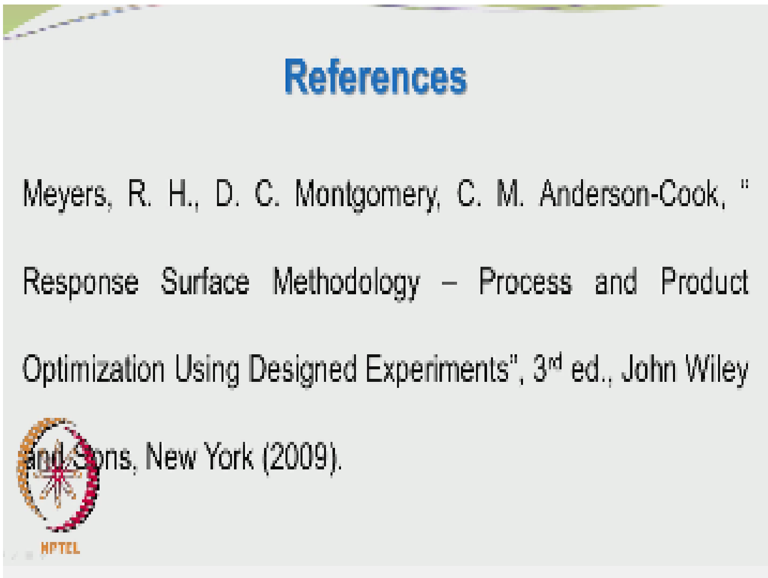
Statistics for Experimentalists
Prof. Kannan. A
Department of Chemical Engineering
Indian Institute of Technology – Madras

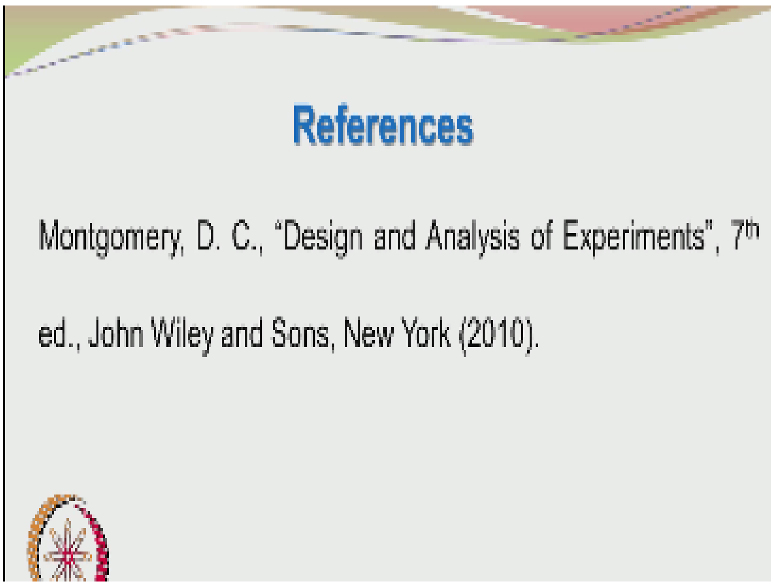
Lecture – 46
Experimental Design Strategies - A

(Refer Slide Time: 00:24)



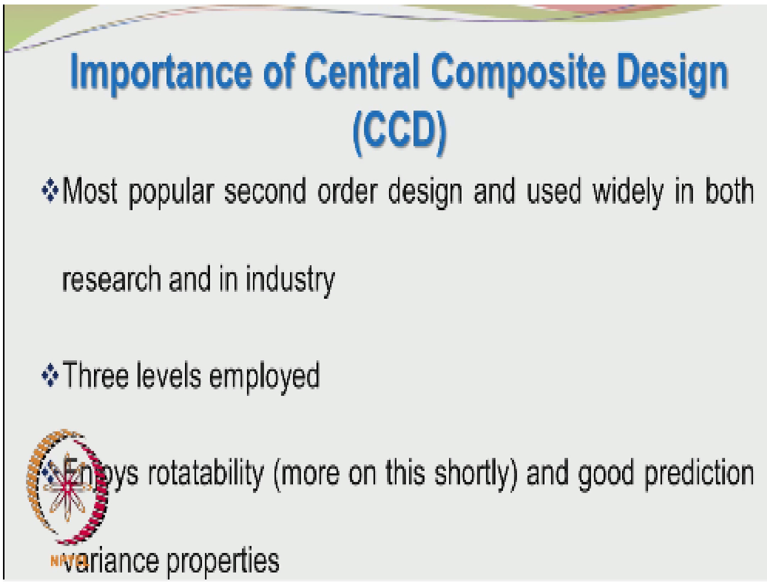
Welcome back in today's lecture, we will be looking at some experimental design strategies. The references for this lecture are the book written by Meyers Montgomery Anderson Cook, Response surface methodology, process and product optimization using designed experiments, 3rd edition, John Wiley and Sons, New York, 2009.

(Refer Slide Time: 00:39)



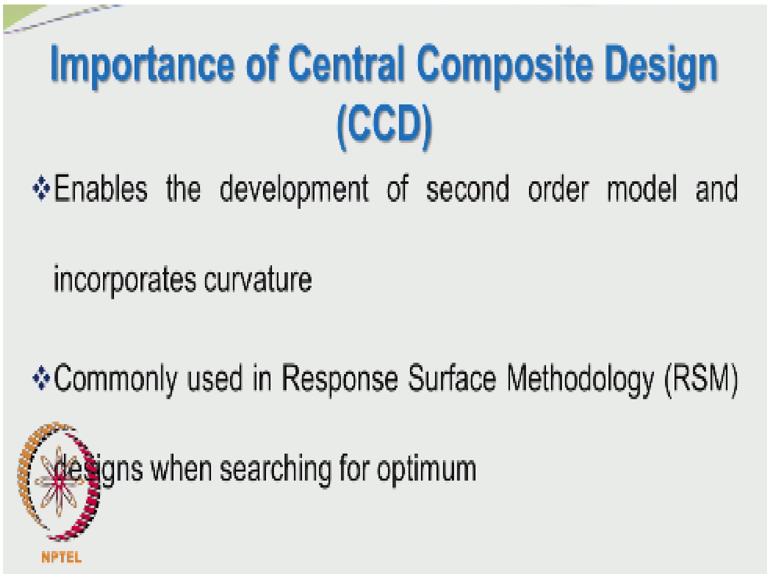
You may also want to refer to Montgomery design and analysis of experiments, 7th edition, John Wiley and Sons, New York.

(Refer Slide Time: 00:54)



The importance of central composite design will be stressed upon in this lecture, it is a very popular second order design used widely in both research and in industry, 3 levels are employed and there are some features like rotatability and good prediction variance properties.

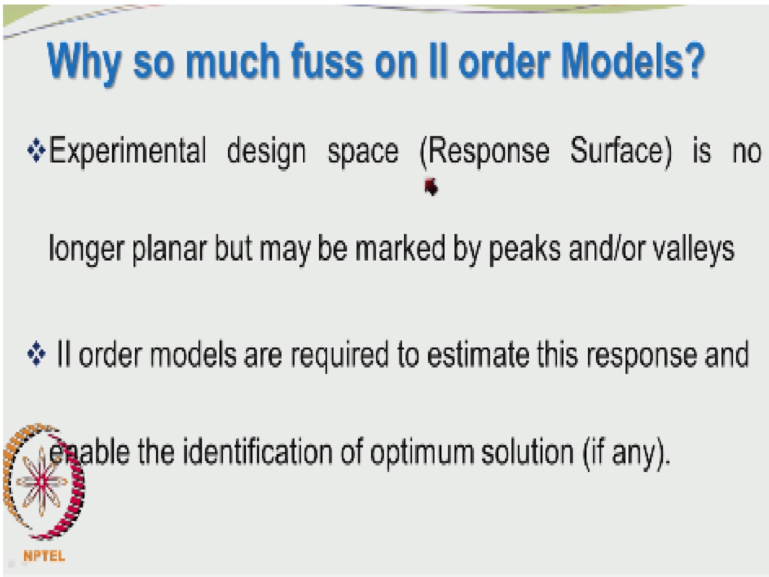
(Refer Slide Time: 01:28)



The central composite design enables the development of second order model and incorporates curvature. What is meant by a second order model? So far, we have been looking at the main factors and the interaction between the factors, when you want to expand in the model space, the response may show curvature and in the multi-dimensional coordinate system, you will have the response surface in the form of a 3 dimensional surface.

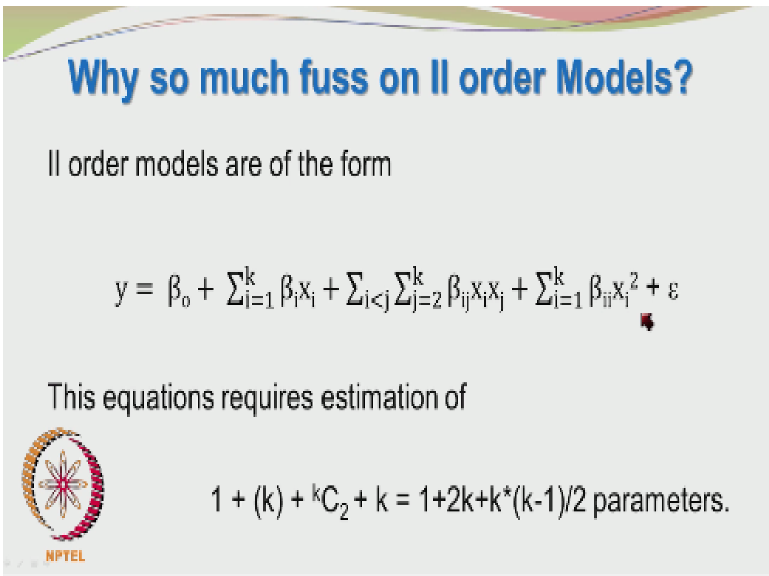
And to describe such kind of response surfaces, we need higher order terms in the model equation; second order terms like x1 squared, x2 squared. Usually, we do not go for models higher than second order unless, it is absolutely essential. So, let us see how we may develop the second order model using the central composite design approach. This is also used in Response surface methodology designs, when searching for the optimum.

(Refer Slide Time: 02:52)



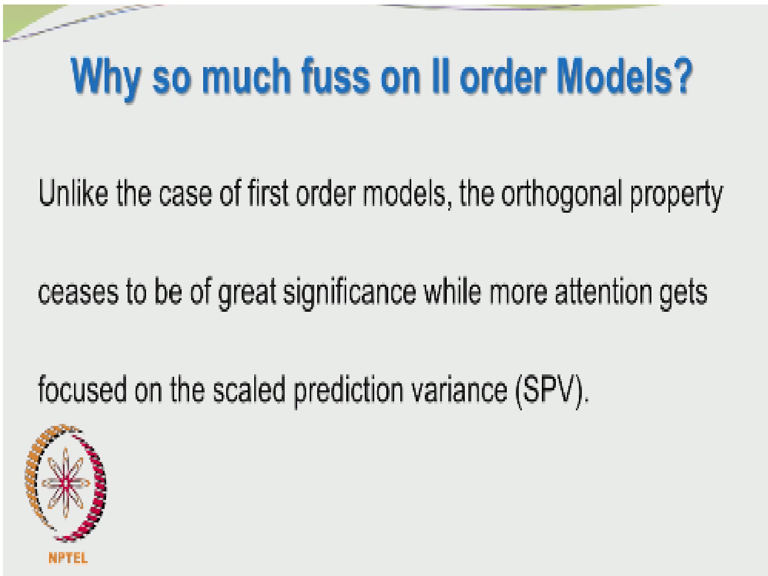
So, why do we emphasize so much on second order models? The experimental design space, the response surface is no longer planar but maybe marked by peaks and or valleys. Second order models are required to estimate this response and enable the identification of optimum solution if any.

(Refer Slide Time: 03:15)



Second order models are of the form y is beta0 + i = 1 to k beta I xi sigma i sigma j beta ij xi xj + sigma = 1 to k beta ii xi squared and so on, so this is the main factors, this is the interaction; binary interaction between two factors taken at a time and then you have the second order terms x1 square, x2 square and so on. This is of course, the error term. So, when you want a fit of model, you fit one intercept, k main factors, kC2 binary interactions and k second order terms.

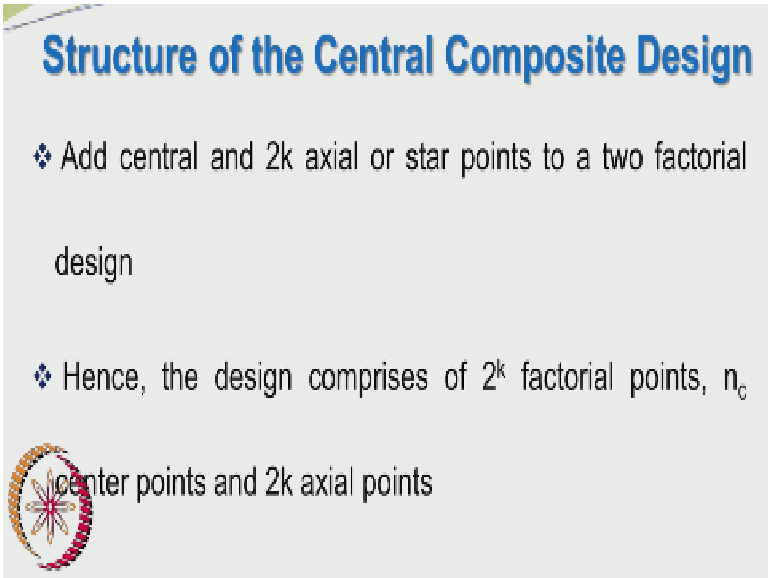
(Refer Slide Time: 04:53)



So, that would be 1 + 2k + k * k – 1/ 2 parameters, if k is = 2, you will have 1 + 45 and then 2 * 2/2 is 1 that would be 6 parameters totally you have to estimate. So, those would be beta0, beta 1, beta 2 that makes it 3 and then one interaction term; x1, x2 that makes it 4 and then x1 squared and x2 squared coefficients that makes it 6. So, when we go for central composite designs, we are no longer able to retain the orthogonal property.

And we shift our attention from the orthogonal property advantages to the suitable low values of the scaled prediction variance. So, when we develop a second order model, we are very worried about its prediction capability and we want to make the variance in the predicted response as low as possible. So, what would be the suitable design strategy, which will bring down the scaled prediction variance, is our goal.

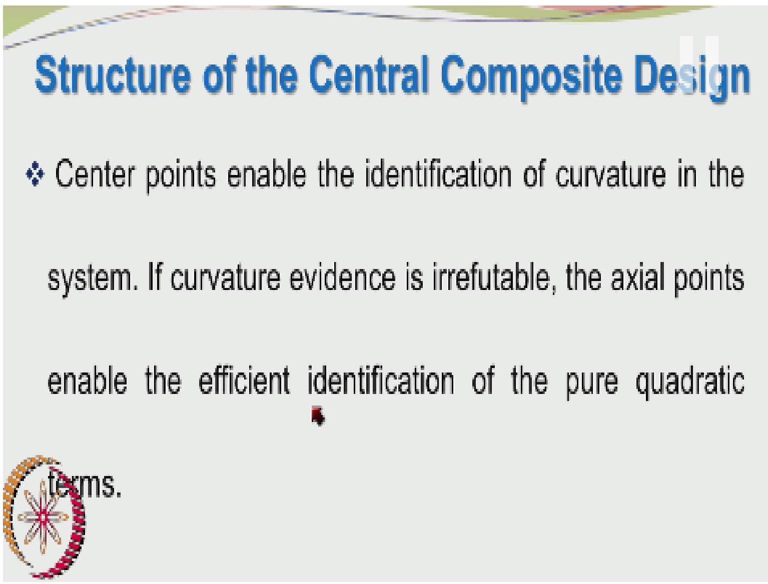
(Refer Slide Time: 05:37)



What is the structure of the central composite design? So, we add a central and 2k axial or star points to a 2 factorial design, so let us take a simple case of a factorial design. First, we add center points, we have already seen the center points at the geometric center of the design space, they were used to not only get an idea about the experimental error but also regarding the significance of the curvature in the response.

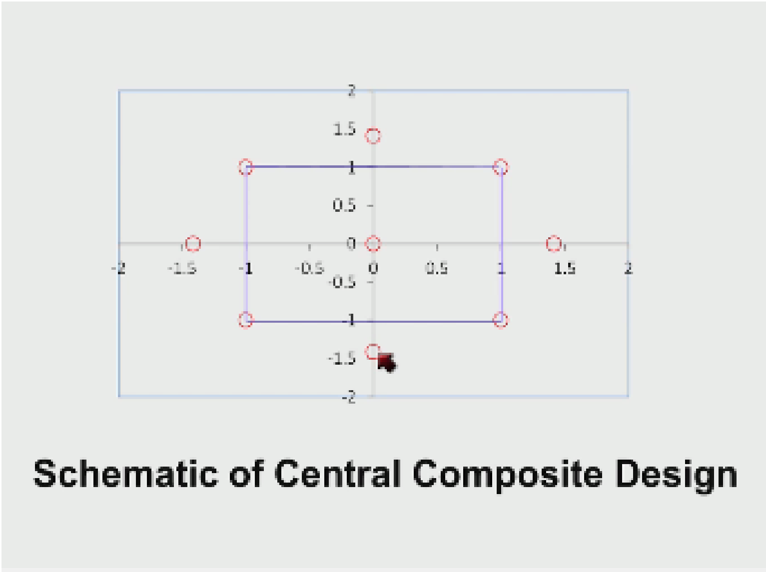
Then on top of the center points, we also add points along the axis. For a 2 factorial design, we have to 2 access; the x and y axis or x1 and x2 axis and you put certain points at select locations on the axis. On each axis, you put one pair of points symmetrically, so when each axis contains one pair of points for a 2 factor design involving 2 axis; x1 and x2 will have for 4 axial points totally or 2 pairs of axial points.

(Refer Slide Time: 07:03)



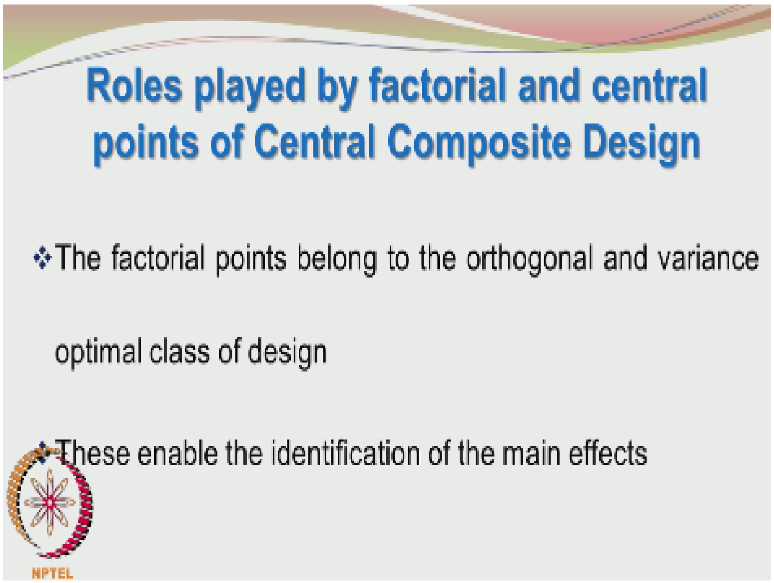
So, the design compresses of 2 power k factorial points, the ones which are located at -1, +1 and so on, nc center points and 2k axial points. So, the center points enable the identification of curvature in the system, if curvature evidence is irrefutable from a T test or a suitable test. The axial points enable the efficient identification of the pure quadratic terms, so each point in the central composite design has its own significance.

(Refer Slide Time: 07:28)



So, what is the central composite design, when looking at it pictorially? You have a central composite design shown here for 2 factors, these are the points in the experimental space as usual the 2 power 2 factorial design has 4 corner points each located at -1 and +1, so this would be -1, 1, 1, 1, 1, -1, -1, -1, the usual factorial design and this is the center points, you can have more than one center point.

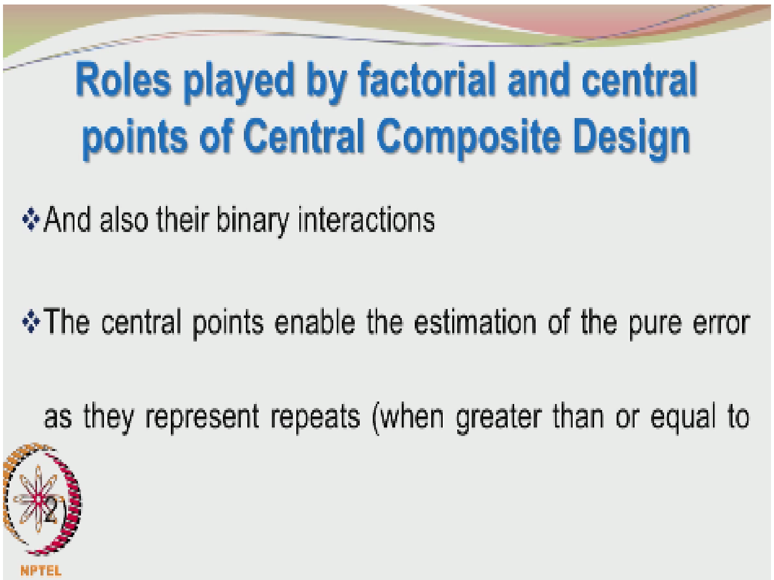
(Refer Slide Time: 08:39)



And then what is unique about the central composite design from the regular factorial design is the presence of the axis or star points, you can see that each axis, this is the x1 axis is having 2 points located at 1.414 and -1.414. Similarly, you have 2 points located on the y axis or x2 axis and they are also located at 0, 1.414 and 0, -1.414. So, the factorial points belong to the orthogonal and variance optimal class of designs.

And these enable the identification of the main effects. The factorial design we saw comprised of points, which were located on the extremes of the design space for that particular design -1 and +1 and since the points were located at very far off positions, you can visualize that the x prime x inverse matrix would be x prime x inverse matrix would be pretty small and that would reduce the variance of the predictions.

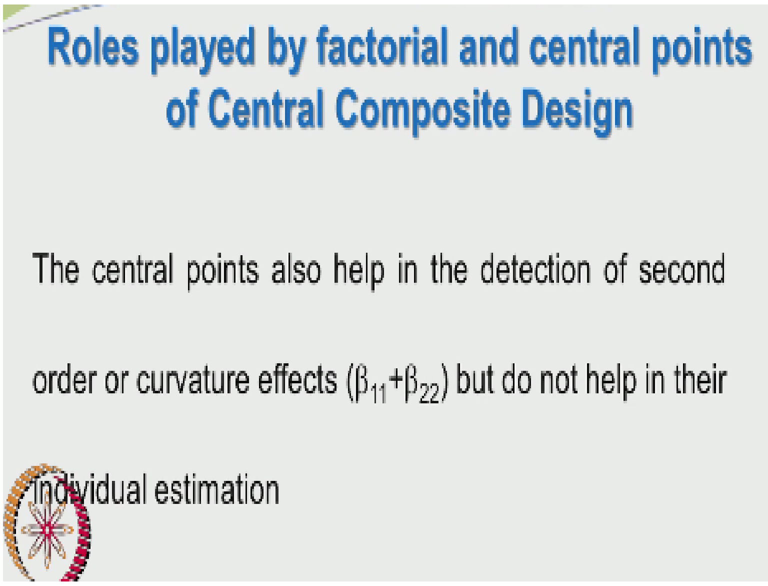
(Refer Slide Time: 09:34)



And hence, it was termed as a variance optimal design, the factorial points are used to find the main effects and the interactions, you find the main effects and interactions in exactly the same way as you did for the regular factorial design. The center points also enable the estimation of the pure error as they represent repeats. So, you need at least 2 or more repeat points and rather than repeating the experiments at all the factorial points, you may want to do the repeats at the geometric center.

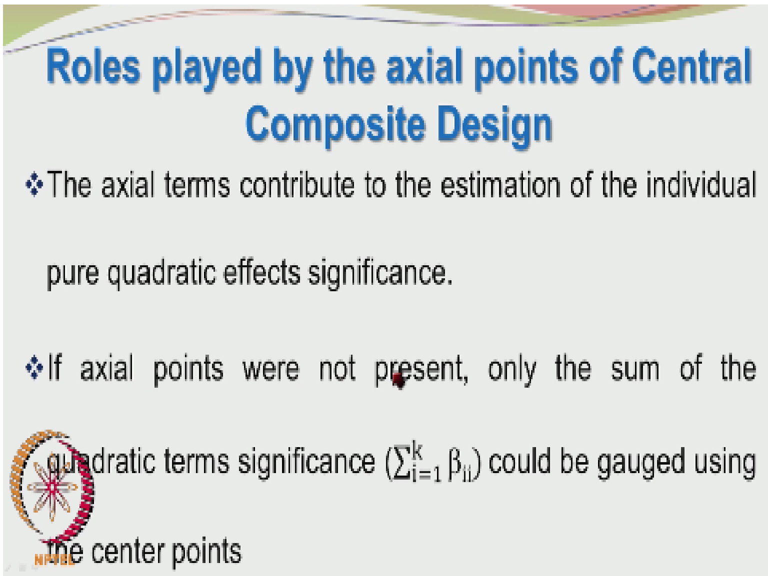
By this way, you can get an idea about the experimental error and also you can save time on doing the experiments at the corners of the factorial design of course, that would lead to more number of runs but in some cases that may be inevitable for the simple reason that certain research requirements require the reporting of the experimental measurements averaged over triplicates.

(Refer Slide Time: 10:36)



So, the center points in other cases are helpful to find the experimental error but in addition to this, they also have another utility. The central points also help in the detection of the second order or curvature effects but do not help in their explicit individual estimation, center points also give us a hint on whether curvature effects are important or not and they tell that whether the curvature is significant, okay.

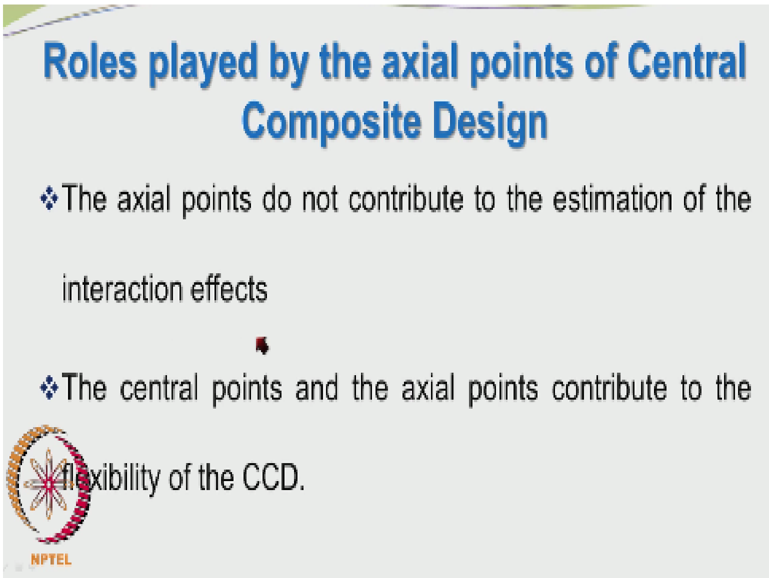
(Refer Slide Time: 11:24)



But it does not help us to explicitly quantify the curvature, it only indicates whether curvature should be considered in the model or not. So, in order to identify the curvature effects explicitly, we require the axial points. Why should the axial points be located at - root 2 or

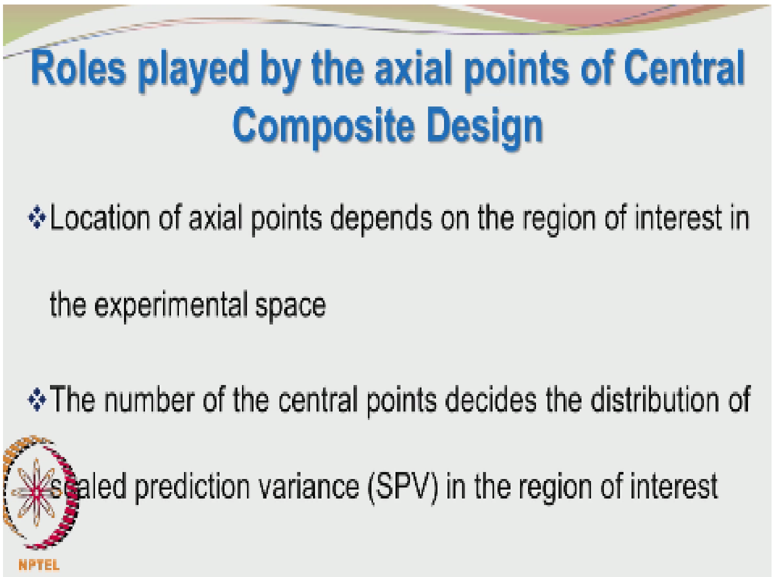
-1.414, and + root 2 or +1.414, the answer to this would be given shortly. So, the axial points contribute to the estimation of the individual pure quadratic effects significance.

(Refer Slide Time: 11:39)



And if the axial points were not present only the sum of the quadratic term significance could be gauged using the center points and the axial points do not contribute to the estimation of the interaction effects, the central points in the axial points contribute to the flexibility of the central composite design. So, by adding the new central points and axial points, which are variance or enhancements to the regular factorial design, we make the experimental design more flexible.

(Refer Slide Time: 12:03)



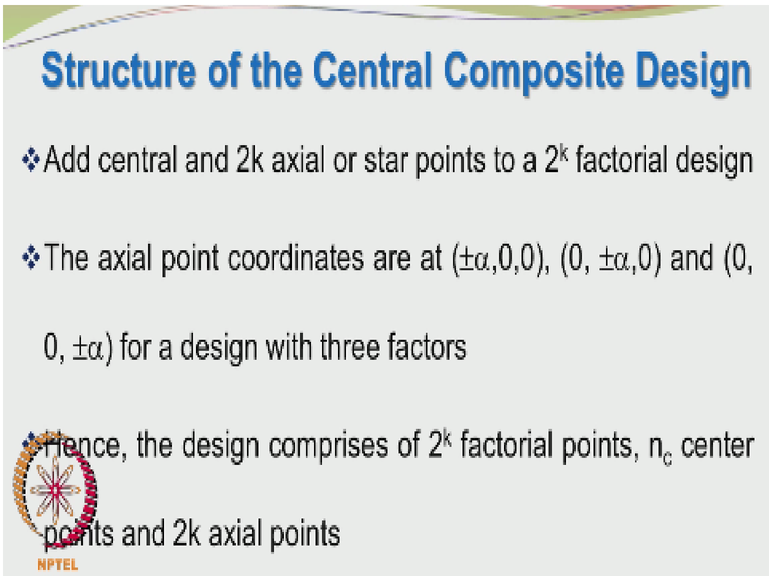
And so, where do we exactly locate the axial points is the next question, it depends upon the region of interest in the experimental space and the number of central points determine the distribution of scale, the prediction variance in the region of interest. So, the location of the axial points depends on the region of interest in the experimental space and the number of central points determines the distribution of a scaled prediction variance in the region of interest.

This is a very important statement because we want to have our model predict uniformly as much as possible in the entire design space. If the variability in the prediction is unmanageably high in our design space, then the models utility is reduced, it is not enough if the model predicts well in the center of the region; the center of the geometric design space, the geometric center of the design space.

But also as we move away from it as we approach the edges of the design space, we want the variability in the predictions to be kept as low as minimum because we normally want to predict the response of the experiment at points further and further away from the geometric center, we may want to even extrapolate sometimes the experimental response beyond the factorial points.

In such cases, if the variances in the predictions keep increasing, then the utility of the model is lost. So, planning for this, we should see what should be the appropriate design strategy and we should also consider parameters like number of center points that would reduce or minimize the scaled prediction variance and an important thing to note here is when you are planning the design strategy, you do not need the experimental data explicitly.

(Refer Slide Time: 14:54)



You can find the scaled prediction variance even before you carry out the experiments and see whether for the experimental strategy, you have adopted the scaled prediction variance is manageable and is acceptable. So, what we do is; we add central and the 2k axial or star points to a 2 power k factorial design. Suppose, you have a central compost design with 3 factors, then you locate the axial points at + or - alpha 0, 0; 0 + or – alpha, 0 and 0, 0, + or – alpha.

(Refer Slide Time: 15:35)

	x1	x2	x3
1	-1.00000	-1.00000	-1.00000
2	1.00000	-1.00000	-1.00000
3	-1.00000	1.00000	-1.00000
4	1.00000	1.00000	-1.00000
5	-1.00000	-1.00000	1.00000
6	1.00000	-1.00000	1.00000
7	-1.00000	1.00000	1.00000
8	1.00000	1.00000	1.00000
9	-1.68179	0.00000	0.00000
10	1.68179	0.00000	0.00000
11	0.00000	-1.68179	0.00000
12	0.00000	1.68179	0.00000
13	0.00000	0.00000	-1.68179
14	0.00000	0.00000	1.68179
15	0.00000	0.00000	0.00000
16	0.00000	0.00000	0.00000
17	0.00000	0.00000	0.00000
18	0.00000	0.00000	0.00000
19	0.00000	0.00000	0.00000
20	0.00000	0.00000	0.00000

CCD for 3 Factors: MINITAB® Data Entry

How to determine the alpha is an important question? we will answer it shortly. So, the design comprises of 2 power k regular factorial points nc center points and 2k axial points. So, let us look at the Minitab output for a central composite design involving 3 factors, you can see the 3 factors are represented by x1, x2, x3 here and then the first 8 experiments are the regular factorial design points.

(Refer Slide Time: 16:34)

Model Coefficients in Coded units

Let us estimate the intercept, main factors, binary interactions and quadratic effects only.

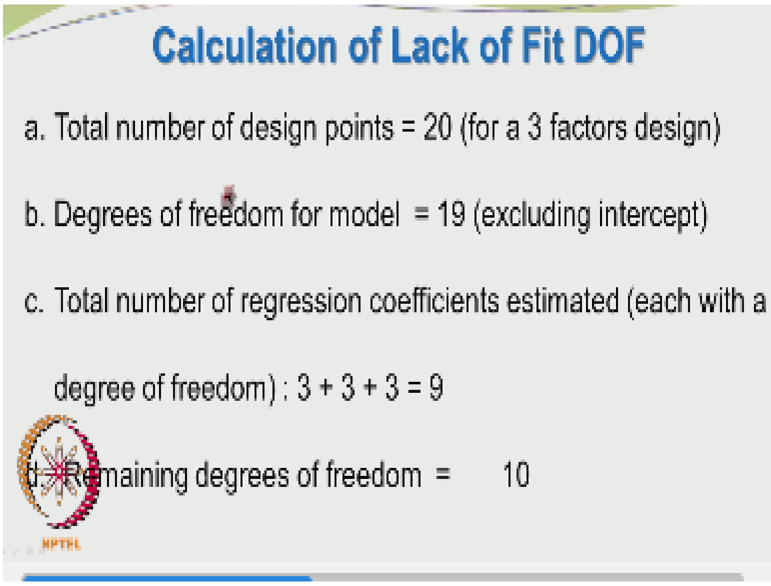
Then we need to estimate the coefficients in addition to β_0 the following

(a) x_1, x_2, x_3 (b) x_1x_2, x_1x_3, x_2x_3 (c) x_1^2, x_2^2, x_3^2

You can see -1, -1, -1, 1, -1, -1 and so on, then the 8th one is 1, 1, 1, then you have these axial points, - alpha 0, 0 + alpha 0, 0, 0, - alpha, 0, 0, + alpha 0, 0, 0, - alpha, 0, 0, + alpha where alpha is 1.68179. What is it special magic number? We have to see shortly and then you have as many as 6 repeated points at the geometric center of the design. So, let us estimate the model coefficients, which are in associated with data that are in coded units.

We have to estimate the model coefficients for the experimental data that are in coded units, so we have to estimate the intercept, main factors, binary interactions and quadratic effects only. So, in addition to the intercept beta 0, we need to estimate the main factors coefficients, the coefficients associated with x1, x2, and x3. Then, we have to identify the binary coefficients associated with the x1 x2, x1 x3, x2 x3.

(Refer Slide Time: 17:36)



And then, we also have to find the coefficients associated with x1 square, x2 square and x3 square. So, the total number of design points is 20, for a 3 factor design with 6 repeats.

(Refer Slide Time: 17:42)

	x1	x2	x3
1	-1.00000	-1.00000	-1.00000
2	1.00000	-1.00000	-1.00000
3	-1.00000	1.00000	-1.00000
4	1.00000	1.00000	-1.00000
5	-1.00000	-1.00000	1.00000
6	1.00000	-1.00000	1.00000
7	-1.00000	1.00000	1.00000
8	1.00000	1.00000	1.00000
9	-1.68179	0.00000	0.00000
10	1.68179	0.00000	0.00000
11	0.00000	-1.68179	0.00000
12	0.00000	1.68179	0.00000
13	0.00000	0.00000	-1.68179
14	0.00000	0.00000	1.68179
15	0.00000	0.00000	0.00000
16	0.00000	0.00000	0.00000
17	0.00000	0.00000	0.00000
18	0.00000	0.00000	0.00000
19	0.00000	0.00000	0.00000
20	0.00000	0.00000	0.00000

CCD for 3 Factors: MINITAB® Data Entry

So, you can see that there 20 independent experimental settings that is not correct, it is not 20 independent experimental settings, you have 14 independent experimental settings and then even though, you have 6 repeats that will constitute only one independent experimental setting, so that would mean 14 + 1, 15 independent experimental settings are there. So, you have total number of design points as 20.

(Refer Slide Time: 18:44)

Calculation of Lack of Fit DOF

d. Number of center points : 6

e. Number of degrees of freedom for pure error : 6-1 = 5

f. Lack of Fit degrees of freedom : 5

This implies that five additional coefficients may be fitted to the model without the risk of aliasing.

The degrees of freedom for model is 19, excluding the intercept and total number of regression coefficients estimated each with the degree of freedom is 3 + 3 + 3 that is = 8, the remaining degrees of freedom is 10, so the number of center points is 6, lack of degrees of freedom for pure error is 6 - 1 which is = 5. So, lack of fit degrees of freedom is = 5, so this is a very interesting calculation for the degrees of freedom for lack of fit.

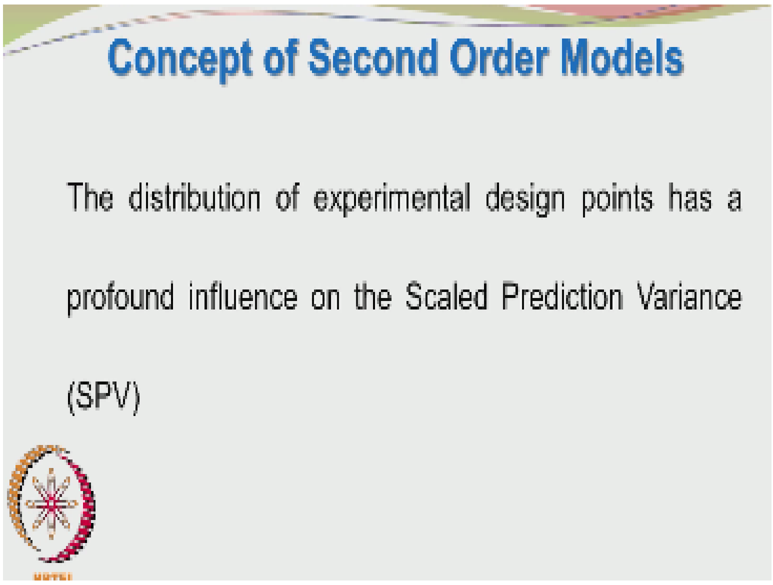
So, even though we have fitted 1, 4, 4 + 3, 7; 7 + 3, 10 parameters, the model possibilities are not exhausted, so there are still some; there is still some scope for expanding the model and adding more coefficients. What can be the number of coefficients that can be further added to the model has to be first estimated. So, if you look at the model, you are having 20 experimental settings but out of that, you are having 14 central composite design points.

The factorial points on the axial points that would be 8 + 6 because you have for 3 factors 2 factorial design, you are having 8 factorial points and for 3 axis, you are having 6 axial points, so that makes it 8 + 6, 14 and then you are having 6 center points but the center points are repeats only that means, that would constitute only 1 independent data setting, so in total, we have something like 14 + 1, which is 15 independent experimental settings.

And if we have already estimated 10 parameters and there are 15 independent experimental settings, we can quickly say that we can additionally estimate 5 more parameters to the model that may not be really necessary but it gives us the option of adding another 5 parameters to the model because of processed knowledge and prior experience, there may be some unusual terms like x1 x2 squared or x2 squared x3.

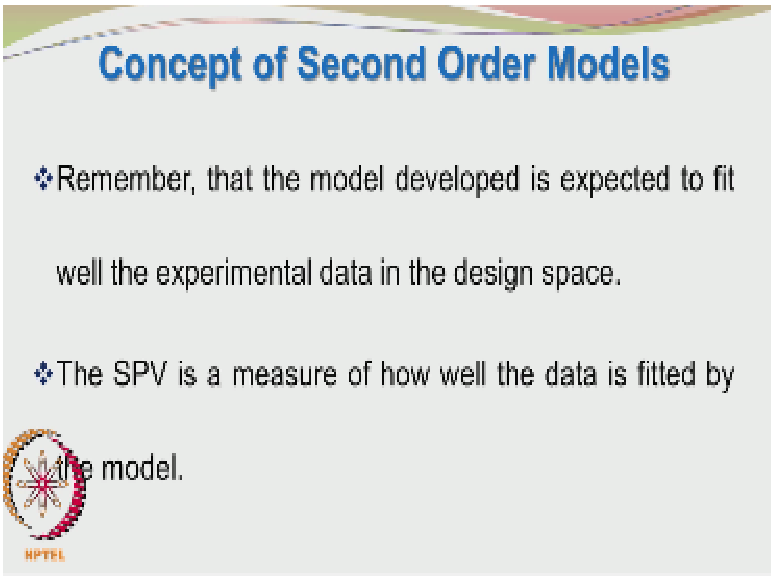
This kind of terms may have to be added to the model because of the peculiarities of the process and then you may need to identify the coefficients associated with those variable combinations. Hence, we have 5 more degrees of freedom for fitting additional model parameters and this is nothing but the lack of fit degrees of freedom. Sometimes, even with 10 parameters, there may be scope for model development.

(Refer Slide Time: 22:24)



And so, the analysis of variance table would indicate that the lack of fit degrees of freedom is significant and hence we may have to consider adding of more terms to the model. So, the lack of fit degrees of freedom is 5, as we discussed just now. So, we can fit additionally 5 more regression coefficients after expanding the model appropriately without the risk of aliasing and now, the distribution of experimental design points has a profound influence on the scaled prediction variance, okay.

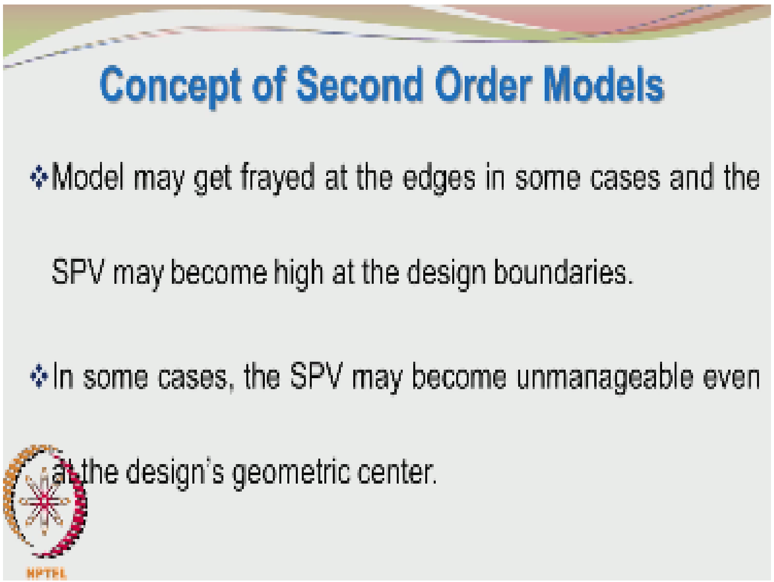
(Refer Slide Time: 22:35)



So, recall that the model developed is expected to fit the experimental data properly in the design space. The SPV is a measure of how well the data is fitted by the model. So, these concepts are very interesting for the simple reason that these are over and above, what we usually are aware of in experimental design. There are numerous instances of sighting of central composite designs in research papers and they give the justification that they are being mainly meant for considering the second order terms in the model but many of these papers do not discuss further as to why the central composite design was chosen among different options available.

And how good is the prediction capability of the model developed using the central composite design, so these are probably beyond the scope of the particular research article but it is very important for us as data analysts and researchers to assess the quality of the developed model, how good the model is and it is also good to be informed about the limitations of the model in the design space.

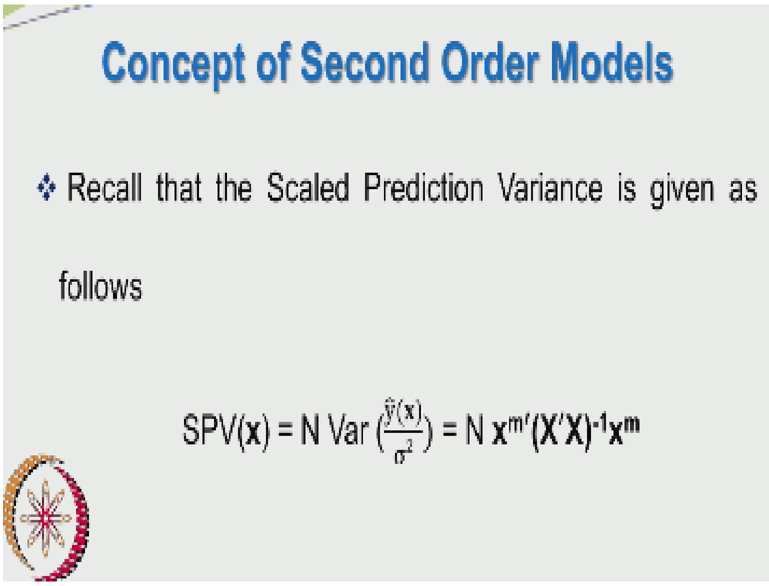
(Refer Slide Time: 24:17)



One important indicator of the limitation of the model in the external design space is this scaled prediction variance and that is the reason why we are harping on it for so many slides. In some cases, the model may get frayed at the edges, so that the scale the prediction variance may be very high at the boundaries, you know the scaled prediction variance may look manageable in the interior portion of the experimental design space.

As we go further towards the extremes or the boundaries of the experimental design space, the scaled prediction variance may shoot up very alarmingly and then the model is not very good at the edges of the design space. In certain cases, there may be problems even at the center of the experimental design space; the scaled prediction variance may be high at the center of the design space as well.

(Refer Slide Time: 25:05)

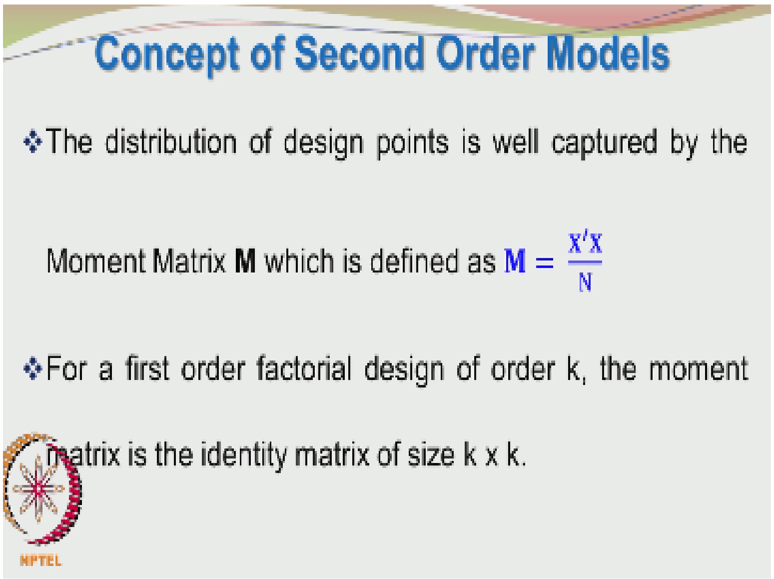


And hence to keep it down or control it, we need to increase the number of center points in certain designs. So, the scaled prediction variance, if you recollect is given by SPV of x is = N, which is the size of the experimental run, the total number of runs in the experiment is N, variance of Y hat X / sigma square and by doing so, we are making the prediction variance independent of sigma squared, which we do not know anyway.

So, we are also getting rid of sigma squared and we are also scaling the design for the size. Certain designs, which are having large number of observations, may artificially bring down the prediction variance because of the large size of the runs. To account for that or to normalized for this effect, we are multiplying by the term n. As an example, if an experiment is performed with large number of repeats, let us say 20 experiments have been or let us say 25 experiments have been performed with large number of repeats.

The prediction variance in such a case would be lower than another experiment, where the number of runs was only restricted to 20, so to compensate or account for the size of the run, we multiplied by n and so the prediction variance, which is multiplied by n and then divided by sigma squared is termed as the scaled prediction variance and we have already seen how to determine the scaled prediction variance, we use x prime x inverse matrix.

(Refer Slide Time: 27:04)



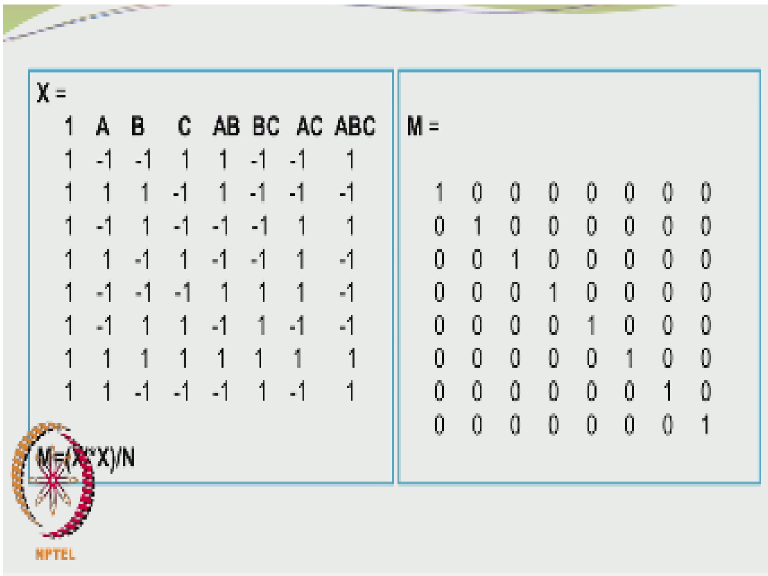
And we also take the coordinate at which we want to estimate the scaled prediction variance and expanded to model space, as was discussed in one of the previous lectures. We introduced at this point the moment matrix M, which is defined as M is = X prime X/N, we saw that the variance covariance matrix is given by X prime X inverse sigma square, so the X prime X inverse or the X prime X matrix is a very, very important term.

Because it captures the essence of your experimental design, whatever design strategy you are implementing is present in the X prime X matrix and the inversion of that matrix help us not only to determine the coefficients of the model proposed but also the variability in the model coefficients and also the variability in the process response. So, these are very significant in experimental design analysis, experimental data analysis and linear regression.

And in such a context, the X prime X matrix assumes the center stage. So, what is the moment matrix, it is X prime X/ N, for a first order factorial design of order k with; that means k parameters, the moment matrix is identified with an identity matrix of size k/k. Suppose, you are having the order as, k, the identity matrix would be having order of k/k. So, let us now look at the second order models more closely.

We define the moment matrix M as X prime X/N, for a first order factorial design of order p, the moment matrix is the identity matrix of size or dimensions p/p, we recollect that p is = k + 1, where k is the number of regression coefficients; beta hat 1, beta hat 2 and so on to beta hat k, in addition to the intercept beta hat 0, so we are having p is = k + 1 regression coefficients. So, the X prime X matrix will also have dimensions of p/p.

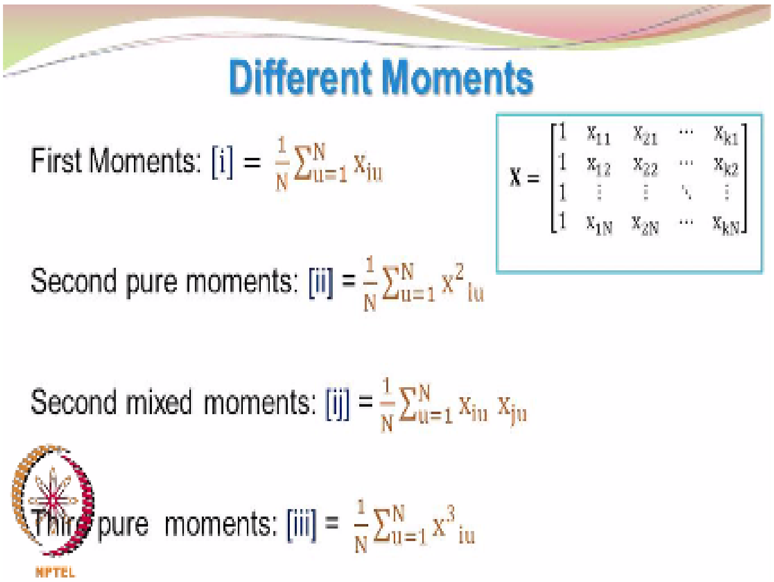
(Refer Slide Time: 30:11)



And the moment matrix X prime X/N would be an identity matrix, the X prime X matrix for a first order factorial design would be a diagonal matrix and when you scale this diagonal matrix by the total number of runs, we get an identity matrix of dimension p/p. Let us demonstrated here and we are having the X matrix, which is given by 1, ABC, AB, BC, AC, the 3 binary interactions and then you have the ternary interaction ABC.

So, this is the X matrix and this is the column of one's and this is the column containing -1, 1, -1, 1, -1, -1, 1, 1 and so on, so we have the entire X matrix. To generate AB, we just simply multiply the elements of the A and B column vectors, similarly for BC and AC and so on and then you also have ABC, which is 1 because it is -1 * -1, which is +1; 1 * 1 is 1, and so when we do M is = X prime X/N, we take the transpose of the X matrix.

(Refer Slide Time: 31:42)



And we then multiply with the X matrix again, then divided by the number of settings N, in this case is = 8; 1, 2, 3, 4, 5, 6, 7, 8 and when we do that the X prime X matrix would be diagonal matrix having 8, 8, 8 and all that but when you divide it by 8, then it becomes an identity matrix of dimension 8/8, so very interesting. So, now let us define the different moments, you have the first moments represented by i and that is given by 1/N sigma u is = 1 up to N Xiu, okay.

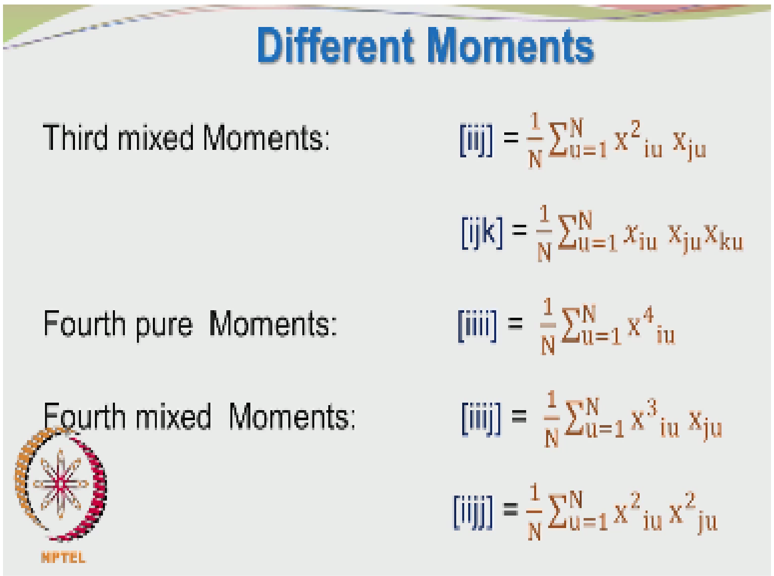
Use the index for incrementing from 1 to N and i refers to the ith column or the ith model parameter for example, if you look at this particular column, if we are talking about X1u, then we take this column corresponding to all ones in the first index and then u is running from 1 to N, so we go from X11, X12 so on to X1n, so the simple thing to note here is we are referring to the ith column and summing over the elements present in the ith column.

So, and that summation is carried over all the experimental settings in the data set and when you look at the second pure moments, we have no adulteration of i with j and vice versa, i is present with i and since its present as a couple, it is a second pure moment and how do you find that? We take the square of the column elements we are choosing, suppose we have chosen bracket; i close bracket corresponding to the 3rd column, then it would be will go to the 3rd column in this X matrix.

And then, we will do X31 square, X32 square so on to X3N squared in this matrix, if i were to be 3. When you have second mixed moments, the column vectors we are considering are different from each other, we are conduct; we are considering 2 column vectors and in these 2 column vectors, i and j are different and so we multiply the individual corresponding elements in each column vector.

So, that we get second mixed moment, this i and j, this i is not = j, i and j are different and hence it is called as mixed moment and that we do 1/N sigma u is = 1 to N Xiu * Xju, so we also have the third pure moment i, i and i, which is since it is pure, there is no additional or a different term in the moment consideration, it is i, i and i that means 1/N sigma u is = 1 to N X cube iu.

(Refer Slide Time: 35:19)

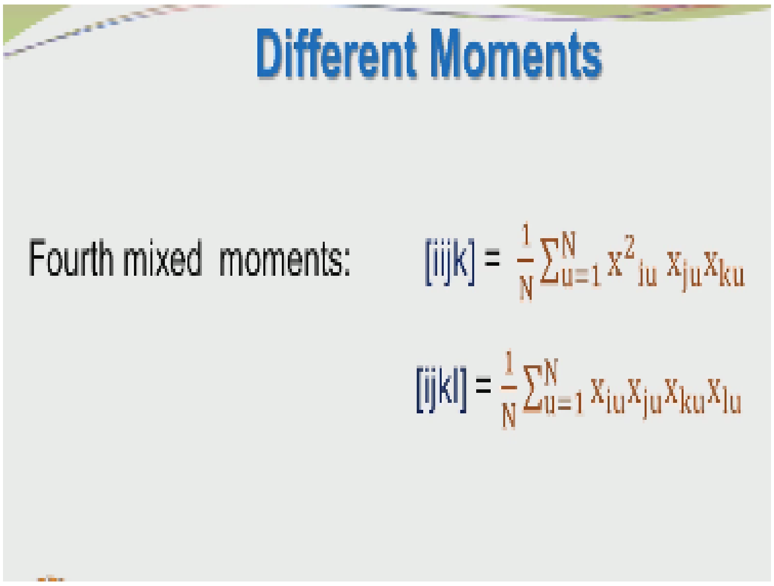


So, for the ith column vector, we just take the cube of each element in that particular column vector and then sum it up; sum it up over all the experimental settings and similarly, we can have all these other moments also, 3rd mixed moment; the total order of the moment is 1 + 1 + 1, which is 3 and mixed moments means, there can be elements, which are different from one another.

You can have two same elements and then you can have a different element j, so this i's and j's obviously refer to the ith column and the jth column in the moment matrix, so we have 1/N sigma u is = 1 to N x squared iu and Xju okay. A correction at this point, these not refer to elements in the moment matrix, they are referring to the elements and column vectors in the X matrix okay.

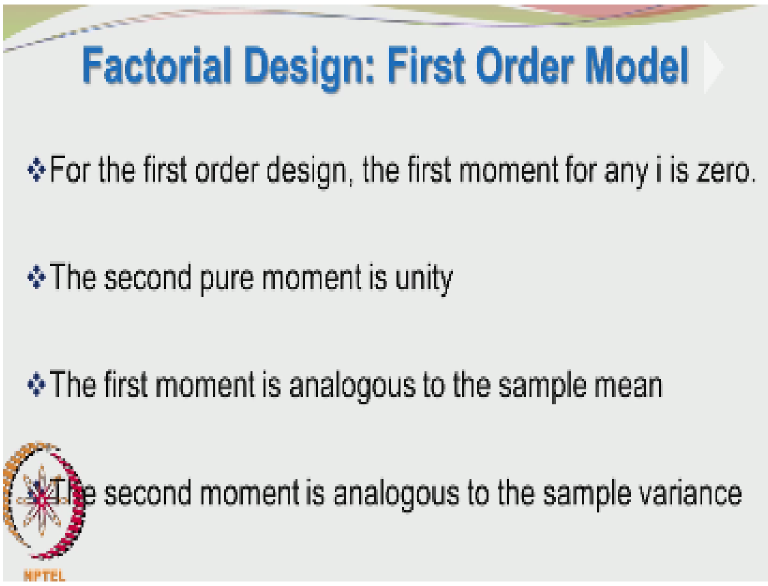
We use the elements in the X matrix, we use the column vectors in the X matrix to find the different moments and if you look at the 3rd mixed moment ijk, it is 1/N sigma u is = 1 to N, Xiu, Xju, Xku. Fourth pure moments are also possible, where we take the 4th power of the elements in the ith column vector and then sum it over the N experimental settings, fourth mixed moments would be the presence of different elements.

(Refer Slide Time: 37:39)



Even i square j squared is considered as the 4th mixed moment, even though you are having a 2 of a species or 2 of certain type together and that is; let us say, ii is present together and jj is present together but since i and j are different, we term it as a 4th order mixed moment and that would be given by 1/N sigma u is = 1 to N x squared iu x squared ju. So, you can also have ijkl, where all the elements within the brackets are different.

(Refer Slide Time: 38:18)



So, they obviously refer to different columns, the ith column, jth column, kth column and kth column in the X matrix and all the i's, j, k and l are different from one another. Here, we have the case of i squared jk, the elements again are being different from each other and hence it is called as a mixed moment. So, for a first order design; a factorial design, the first moment for any i is 0.

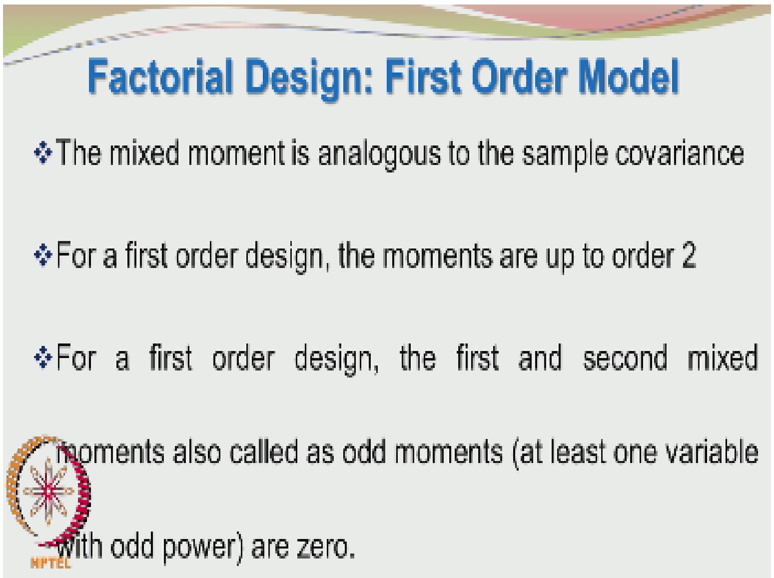
So, when you are looking at the X matrix, if you look at any column, we are not having columns with the contributions from X i squared like X1 squared or X2 squared, so all the elements in the X matrix for this case, would be comprising of -1, +1 and so on, except for the vector of one's, all other columns would be having -1 and +1 and when you total it up for each column, it will become 0.

For example, if you look at the main effects; X1 or X2 or X3, each column would be having -1 and +1 in an equal number and so when you take the sum, it will go to 0 and that is what is meant by the first moment for any i is = 0 for the first order design and the second pure moment is unity, you may ask how it is possible. The second pure moment is either i squared or j squared and so each of the -1 or +1 will uniformly become +1 only after squaring.

So, when you are having let us say, 8 runs, you are going to have the sum as 8 but please remember according to the definition of the second moment, we are; or for that matter any moment, we are dividing by N, so that 8 will get cancelled with the 8 and hence you will get 1. So, you can see the second of pure moment is having 1/N here sigma u is = 1 to n Xi squared X squared iu and so all these things would be 1's.

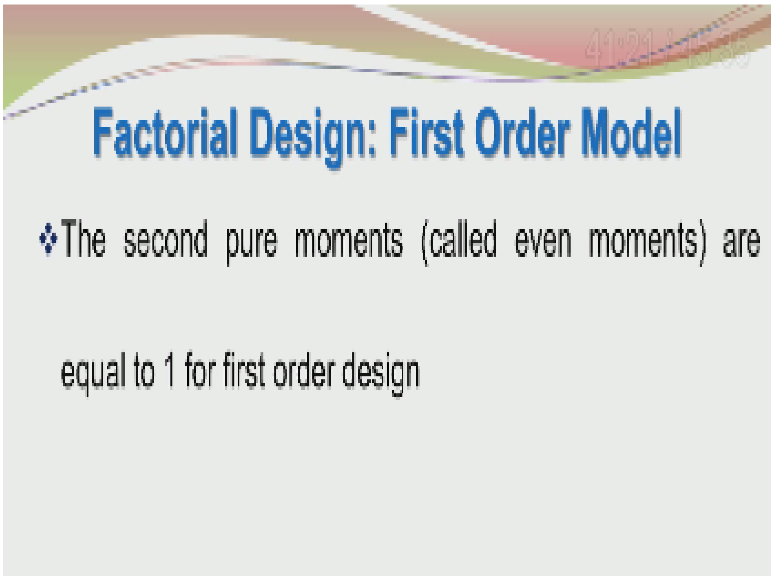
And you are adding it up to N times means, you will get N and N/N would be = 1, the second pure moment is unity that is what we saw just now because we are dividing all the squared elements with the size of the run and so they cancel out and the resulting answer is just 1. The first moment is analogous to the sample mean; the second moment is analogous to the sample variance.

(Refer Slide Time: 40:49)



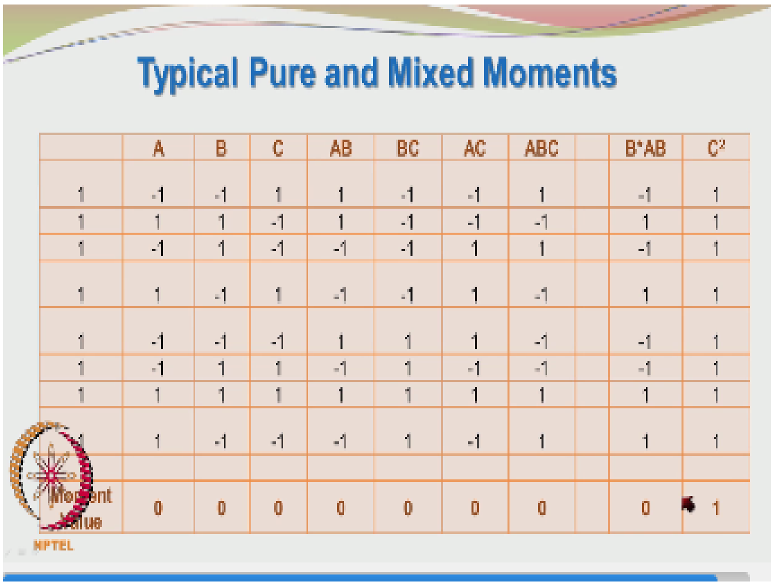
And this mixed moment is analogous to the sample covariance, for a first order design, the moments are up to order 2 and for a first order design, the first and second mixed moments also called as odd moments, at least one variable with the odd power are 0. So, the odd moments are 0, for a first order design.

(Refer Slide Time: 41:15)



The second pure moment called as even moments or = 1, for the first order design. So, now let us look at a saturated 2 power 3 factorial design X matrix.

(Refer Slide Time: 41:28)

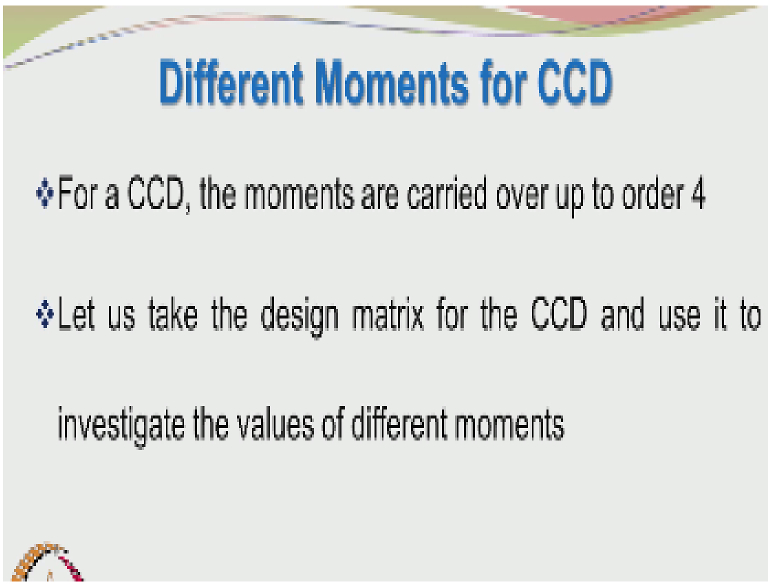


	A	B	C	AB	BC	AC	ABC		B*AB	C²
1	-1	-1	1	1	-1	-1	1		-1	1
1	1	1	-1	1	-1	-1	-1		1	1
1	-1	1	-1	-1	-1	1	1		-1	1
1	1	-1	1	-1	-1	1	-1		1	1
1	-1	-1	-1	1	1	1	-1		-1	1
1	-1	1	1	-1	1	-1	-1		-1	1
1	1	1	1	1	1	1	1		1	1
1	1	-1	-1	-1	1	-1	1		1	1
Moment value	0	0	0	0	0	0	0		0	1

So, you are having typical pure and mixed moments for 2 power 3 design X matrix and so you see that the X matrix is having the usual column of 1's, it is having the column of A, B, C, the main factors; AB, BC, AC, interactions, ABC, which is the ternary interaction and then you also have B into AB, C squared, so all these things are created very easily. For example, the column B AB or B squared A is created by squaring B squared sorry; by squaring B.

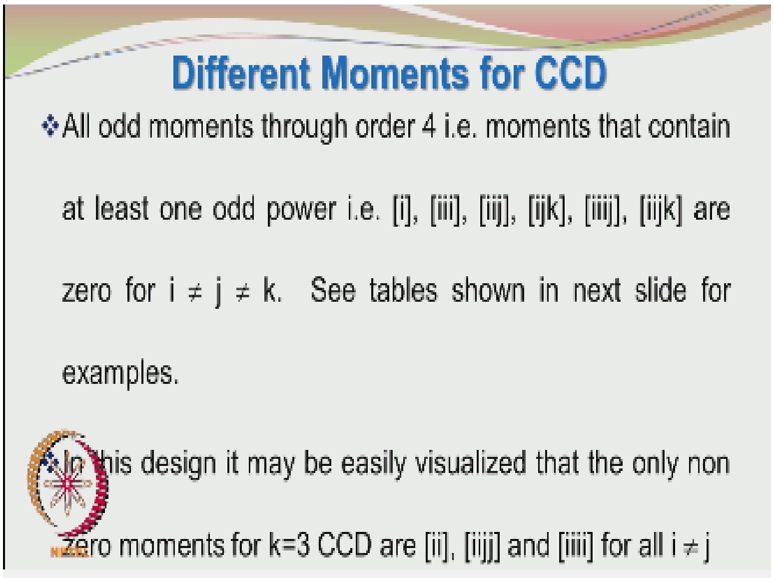
So, these values will all become 1 and then multiplying with A, so when you get B square A; 1 * -1 will be -1 and similarly, B Square will be 1 and A would be 1and so you are having +1. Similarly, you can find out C squared and B squared A would be a 3rd order moment; 3rd order mixed moment because A is also present here and when you look at the 1st order moments corresponding to the main factors, when I am totalling all these values it becomes 0.

(Refer Slide Time: 43:36)



And if I am even looking at B squared AB, it is also = 0 because it is having equal number of minuses and plus here but when I do C squared, it becomes 1 throughout the column and when I add it up; 1, 2, 3, 4, 5, 6, 7, 8; it becomes 8, the size of the run is also = 8; 8/8 will be = 1 and that is why you have 1 here. For a central composite design, it can be shown that the moments are carried over up to order 4 and let us take the design matrix for the central composite design.

(Refer Slide Time: 43:50)



And use it to investigate the values of the different moments. So, for the central composite design, we can look at the values taken by different moments, all odd moments through order 4 that means orders 1, 2, and 3 are also included, it says that is why through order 4 that is the moments that contain at least 1 odd power like i or i cube or i square j, so there is at least 1 order power corresponding to the power of j and then ijk.

This is all completely odd moments because i is different from j and j is different from k and i cube j and i square the jk are 0, for i is not = j not = k, so tables are shown in the next slide as examples. So, in this design it may be usually visualized that only nonzero moments for k is = 3 are i squared i square j square or i to the power of 4, for all i not = j. So, we will continue on this topic after taking a break.