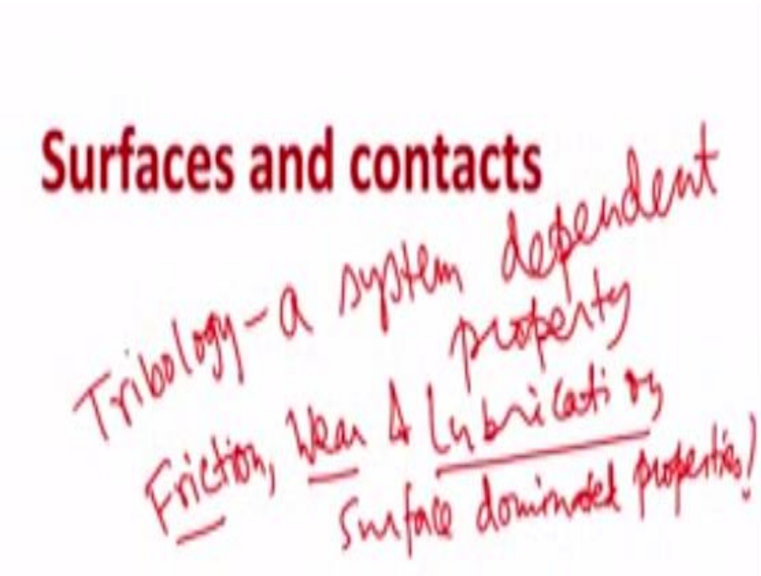
Friction and Wear of Materials: Principles and Case Studies
Prof. Bikramjit Basu
Materials Research Centre
Indian Institute of Science - Bangalore

Lecture – 2
Surfaces and Contacts

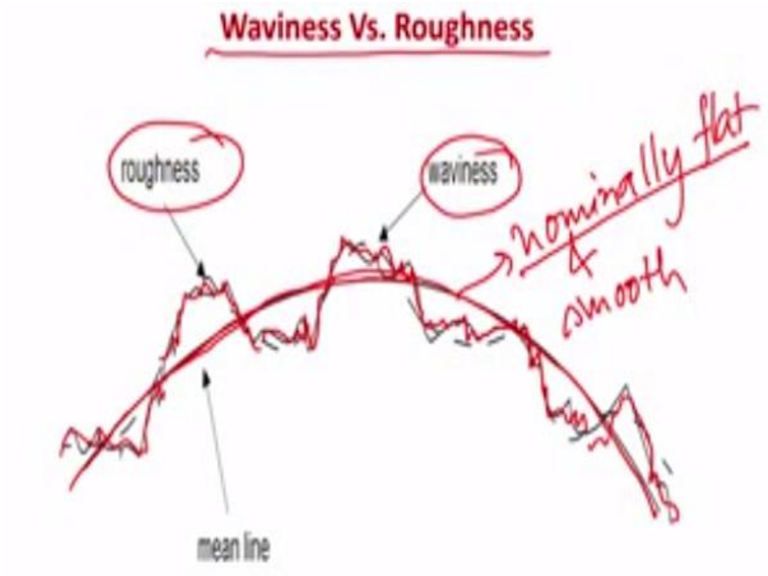
This is the second lecture in this NPTEL lecture series on tribology. In the first lecture, we have reviewed some of the fundamentals of that friction and wear. So, one of the things I would like all of you to carry forward the thought that tribology is a system dependent property okay. **(Refer Slide Time: 00:51)**



So, tribology essentially is a system dependent property and that most of the elements of tribology are like friction, wear and lubrication. These are the 3 pillars of tribology. Friction is 1 pillar, wear is another pillar, and third pillar is lubrication. These are like surface dominated properties. What it means that friction between two mating solids? This is essentially governed by physicochemical interactions between the asperities of the two mating solids.

The result of this interaction leads to wear of materials. How to prevent friction and wear? you need to use the lubricants and essentially the major role of lubricants is to physically separate the asperities from the 2 interacting surfaces. So, this a nut shell that why in the field of tribology one has to understand the characteristics of the surfaces and contacts.

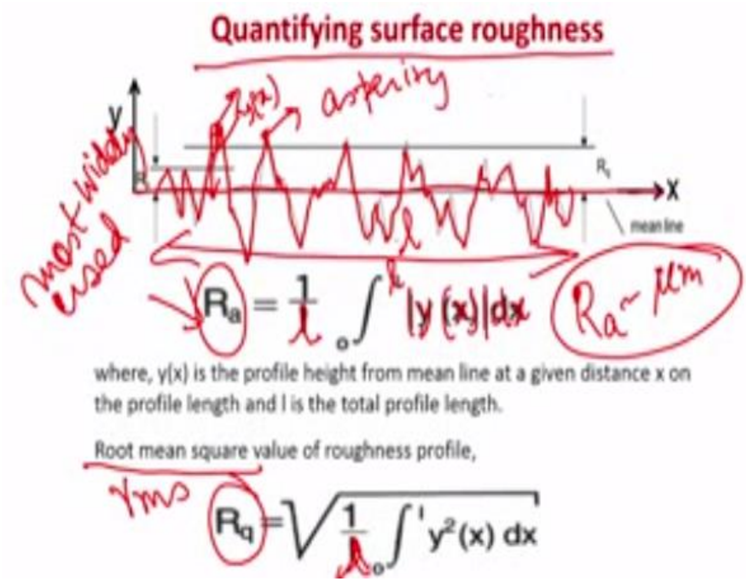
(Refer Slide Time: 02:24)



Now, let us start with some of the basics of how to characterize the surfaces? What you will see here? This is the part of a sphere. Now this sphere it can appear to you it is a nominally flat and smooth sphere. When I say that nominally flat and smooth, essentially what it means? It means that this word nominally flat means what you see with the naked eyes. But if you put the same sphere surfaces under the profilometer or under the microscope, what you see is as follows:

What you see? These essentially normally flat surfaces have lot of asperities which contributes to the roughness. Also, there is some other terms called waviness. Roughness and waviness these 2 terms essentially constitute the surface characteristics of this nominally flat sphere. I repeat you have the mean line. This mean line is what you see with the naked eyes. But if in a much lower length scale if you go, then this nominally flat surfaces will consist of roughness and waviness.

(Refer Slide Time: 03:58)



Now, the next question that should come to your mind is that if you say surface is rough, how to quantify the surface roughness? For the quantification of the surface roughness, there are a host number of parameters which are widely used in the field of tribology. Among them the most commonly used parameter is R_a. R_a is the average surface roughness. Now let me define what is meant by R_a? What you see, again the way I am pressing this line, this line is essentially nominally flat surface.

In this nominally flat surface, you will see there are lots of peaks and valleys. These are the peaks and these are the valleys. Sometimes, the peaks are very sharp or sometimes the valleys can be sharp or they can be bimodal peak and there is lot of undulations right. So, these will constitute the asperity. Essentially these are called asperity.

$$R_a = \frac{1}{l}\int_0^l |y(x)|dx$$

What is l? l is the total sampling line okay.

What is y(x)? y(x) is the profile height. If you take this particular point at this position x, this will be defined as y(x). You take various sampling line, profile line, at many distances x within this span, within this sampling length l. You integrate it, you divide it, the result of integration by sampling length l, so that will define you the average of surface roughness. So, the average surface roughness as I said by far the most widely used parameters.

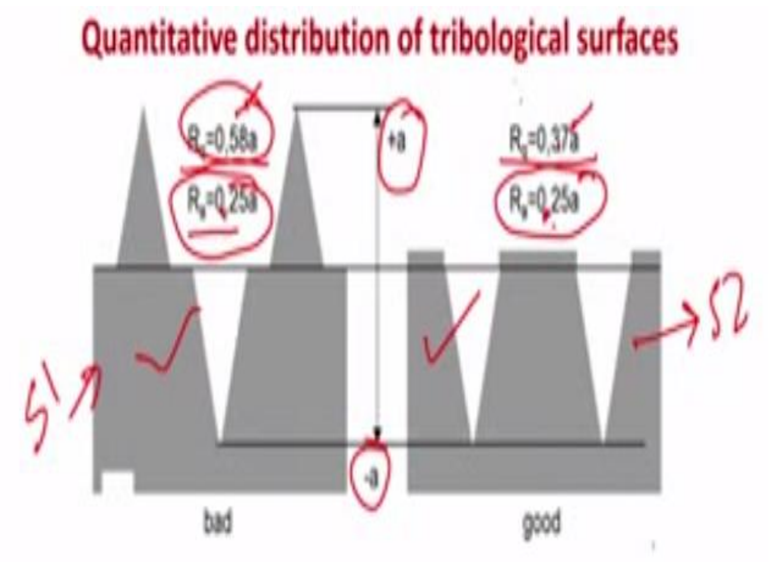
Anytime if you get a machine surface, the first thing people ask that what is the R_a value of the machine surface. Typically, R_a values of the various surfaces the length scale which we prefer is on micrometer. If the R_a value says goes to millimeter, that is extremely rough surfaces. It is not acceptable in most of the industrial applications. R_a value should must be in the order of micrometer scale. The second one is the root mean square, that is called rms roughness, and that is R_q.

$$R_q = \sqrt{\frac{1}{l}\int_0^l y^2(x)dx}$$

What you see from this expression that fundamental difference between R_a and R_q is that in case of R_q you are essentially taking the square. Once you take the square and then you take a square root. Then what would happen? That you know that the deviation of the asperity from the mean line that is being squared. That means it is being magnified or amplified.

So, you would be able to get even the very finer deviation that becomes magnified. So essentially R_q value in reality captures very close to that real roughness values of any surface. If you are given the choice between R_a and R_q, I would choose R_q value because R_q value captures much better the surface deviation, asperity deviation from the mean line.

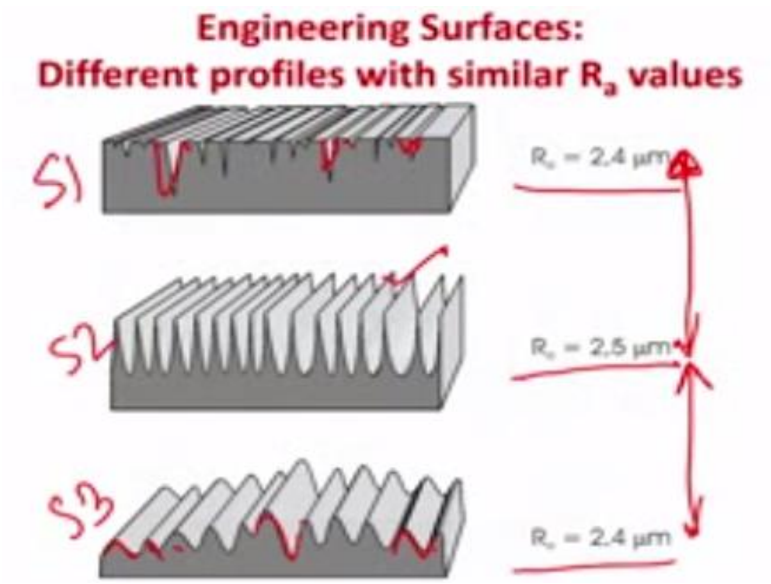
(Refer Slide Time: 08:02)



Now, this particular slide is important. What it does? It says that the 2 surfaces can have same R_a values. This is your solid surface 1, this is your solid surface 2. So, S1 and S2. As you can see that S1 R_a value is 0.25 times a; S2 R_a values is 0.25 times a. So when you say a, a can be any values and the typical a value is mentioned here. Now although their surface roughness is equal, but their R_q values are quite different.

In this particular case R_q value is 0.58 times a and this time the R_q values are 0.37a. 0.37a and 0.58a, certainly this surface is much rougher. So, if you follow strictly the R_a values, then these 2 surfaces S1 and S2 may be equivalent. But in reality, they are not equivalent surfaces simply because they have a 2 different R_q values.

(Refer Slide Time: 09:18)

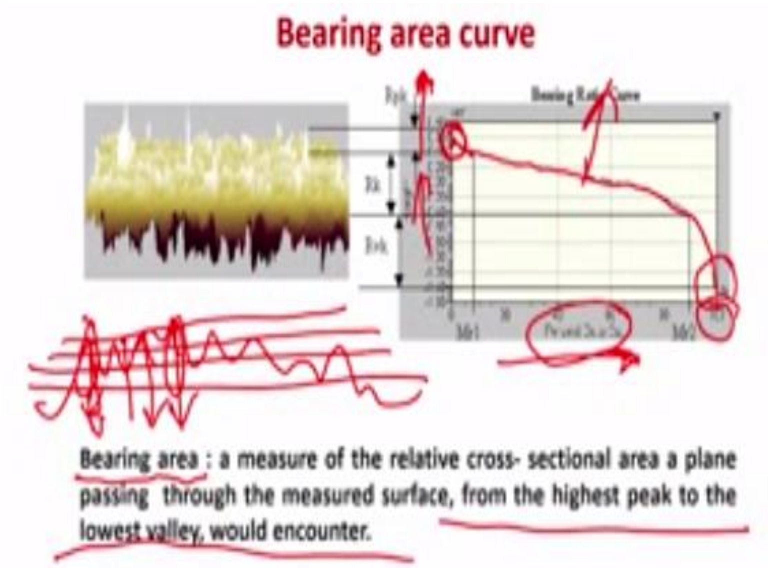


This is another example that these 3 surfaces again, S1, S2 and S3. These 3 engineering surfaces. To any layman, these 3 engineering surfaces must be different. Because if you see the nature of the asperities are different. These surfaces, it is much more regular asperities. These surfaces, the valleys are much more, much deeper. These surfaces, there. are irregular asperities. You can see these asperities is much higher than these asperities and so on.

So, by far these surfaces may appear to be much more uniform. But if you quantify the surfaces on the terms of R_a values, then what you see, they are like equivalent, 2.4 micron is almost closer to 2.5 micron right. So, what I am trying to point out through these 2 slides is that that R_a value cannot be used in absolute scale to distinguish and differentiate between different engineering surfaces. There is a need to consider other surface roughness parameters like R_q values and another which I introduce into later.

While R_a value which is widely used in the community. I have already shown several, at least 5 examples, where the 2 surfaces can have identical R_a values but they are essentially different surfaces as far as the physical features of the asperities are concerned.

(Refer Slide Time: 11:02)



So, in this perspective one of the things that may be very relevant is that bearing area curve. Now, what is the bearing area curve? Bearing area curve essentially means that if you plot that deviation of any point on the surface from the mean line and how much deviation that particular height or that particular deviation is there on this particular surface along X-axis. So, what I say along Y-axis that relative cross-sectional area of a plane passing through the measured surface and then from the highest peak to the lowest valley would encounter.

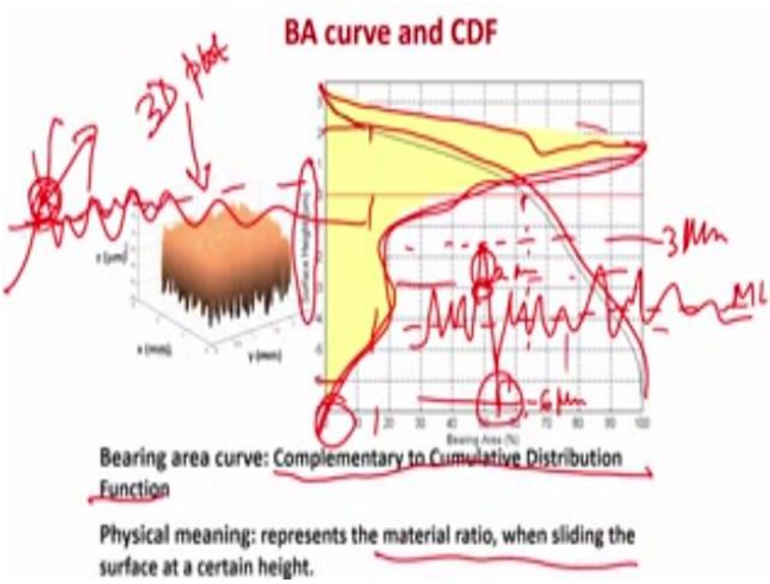
So, what is the bearing area curve? Bearing area curve the way I am expressing that is a typical nature of the bearing area curve. So, what means is that, that here along the Y axis, it is the deviation of the profile on the surface on the mean line and that is how much percentage of the plane that passes through the deviation that is plotted along the x axis. So, if you see that higher bearing area higher the surface deviation that is covered by a very small fraction of the area whereas very low amount of this deviation that will constitute more or less like 100% surfaces.

This is a very important concept which is used in many cases to understand the difference between different types of commercial lubricants. Traditionally bearing area is defined as a measure of the relative cross-sectional area of a plane. So, relative cross-sectional area means suppose this is a plane. This is theoretically rough surface. Now you take a plane, you take a different plane. So, these different planes are cutting, not all the asperities, but some of the asperities are at a particular height. So, this is another asperity this is in particular heights.

So you have to only consider what is the amount of material that is contained by this particular plane which is cutting through the different asperities and these percentage of the area of the plane which contains the material that is plotted along the X axis and the height of the asperities that is plotted along the Y axis. If you plot it theoretically, it will give rise to this kind of characteristic curve. This is called bearing ratio curve and through the measured area from the highest peak to the lowest valley would encounter.

So, highest peak is this one and lowest valley is this one. This is a very important concept which is used in lubrications.

(Refer Slide Time: 14:13)



This is another representative surface. This is the 3-D plot of the engineering surface and **(Video Starts: 14:27)** as I said in the last slide that is in the surface height. If you go back to this plot, I am saying that this is the surface height for any plane. Let us say if you take this plane, the surface height is this one right. So, if you put the surface height 3 or 2 microns, then you go particular here. Then you say at least 15% the materials that will be under the plane passing through 2 micron height from the mean line.

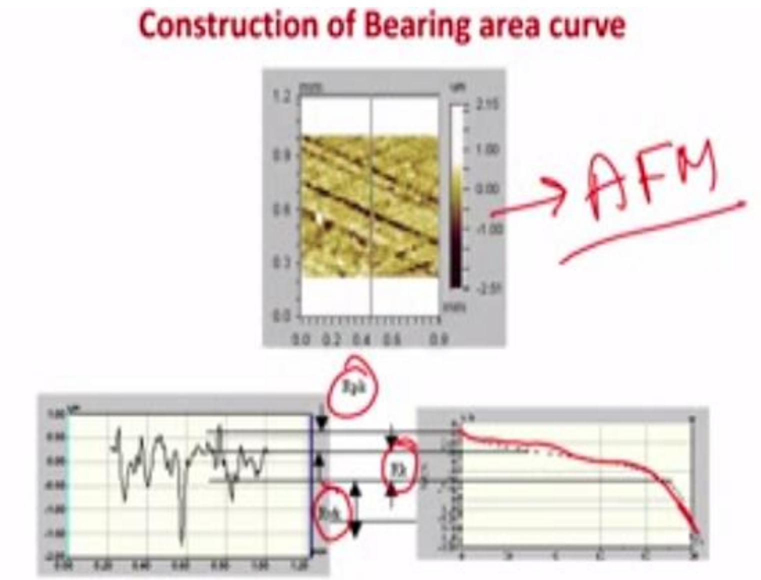
Now if the surface deviation is -7 micron. Let me just clarify these things in this hypothetical thing. So, if you say this is your mean line, this dotted line is your mean line. Then for example if you take this as your maximum, this is like 3 microns. you see this material that is contained in very minimal like it is close to 0. Now if you go to -6-micron. then you say that most of the materials, all the materials -6 micron is 100%.

But if you go to like 2 microns for example, then at least some part of the material, that the material is here, that some part of the material and this may be less than 10% or something that is there. So, if you go to mean line here. Essentially it is like 60% or more than 60% because some fraction of the material is in the top part and some fraction of the material is at the bottom part. So, this is the cumulative distribution function of the surface height, you can see this is cumulative.

So, this is not a single mean or mode. This is a bimodal type of distribution. What you see here? This is your bearing area curve. Bearing area curve as I said that is a typical characteristic feature. So, what it mentioned here bearing area curve is complementary to the cumulative distribution function and physical meaning of the bearing area curve is that it represents a material ratio when sliding at the surface at a certain height.

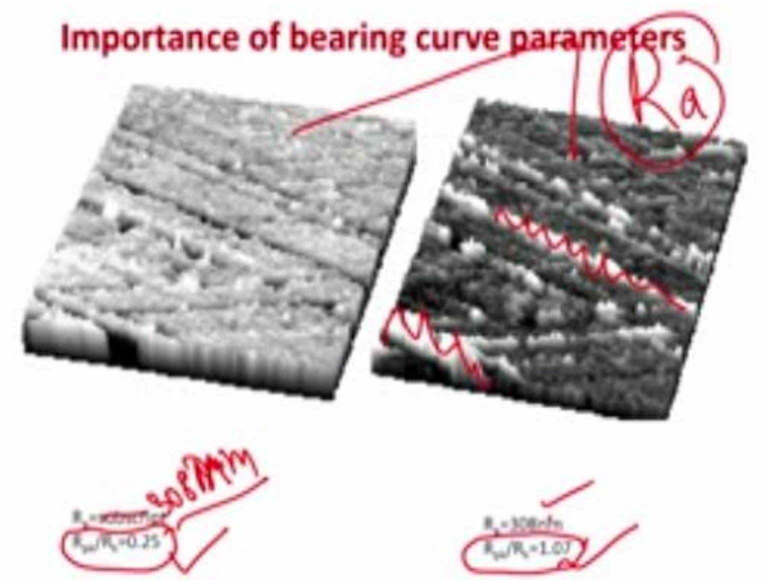
What it means that? Suppose your bearing area curve, if the material is sliding at this particular height, you know this much material will be in contact with this particular mating solids. Now if this particular area is getting chopped off like you know it is getting lifted, then it allows the next level of materials **(Video Ends: 17:25)** to be in contact to the mating solids.

(Refer Slide Time: 17:30)



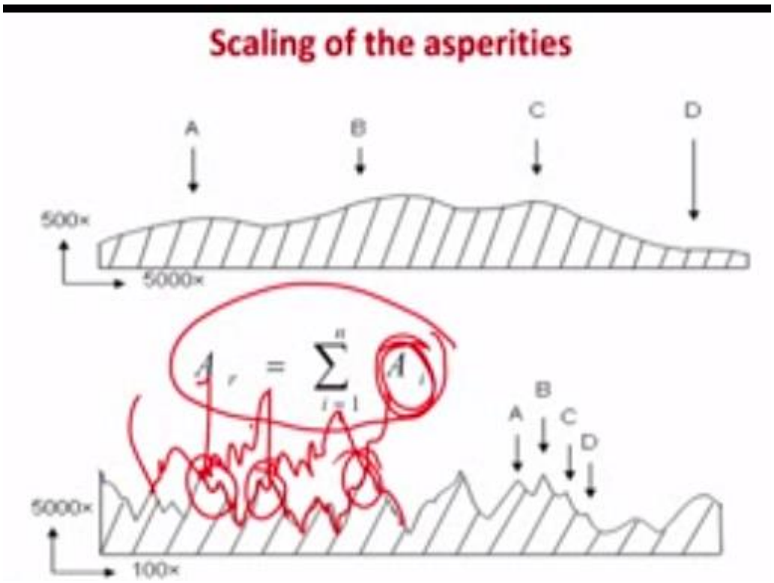
So, this is the construction of the bearing area curve. You know you can get all this 2-D profile. From the 2-D profile, you can get a bearing area curve. This is a typical AFM image, AFM stands for atomic force microscopy. That gives a very high resolution like it goes to micron scale and even lower surface roughness. And from there, you can find out that what is the valley roughness? what is the peak roughness? what is the mean roughness of the surface aspect ratio?

(Refer Slide Time: 18:00)



Now, one of the importance of the bearing area ratio curve is suppose if its R_a is 308 microns and this is also 308 microns. So, in both the cases, R_a values are the same. But if you consider that R_{pk} to R_k that is the peak roughness to the mean roughness values from the bearing area curve, this is less rough, but this is more rough. You can see that this is more rough, although that same R_a values, but if you quantify the other roughness values, the R_{pk} a versus R_k, then they are quite different.

(Refer Slide Time: 18:47)



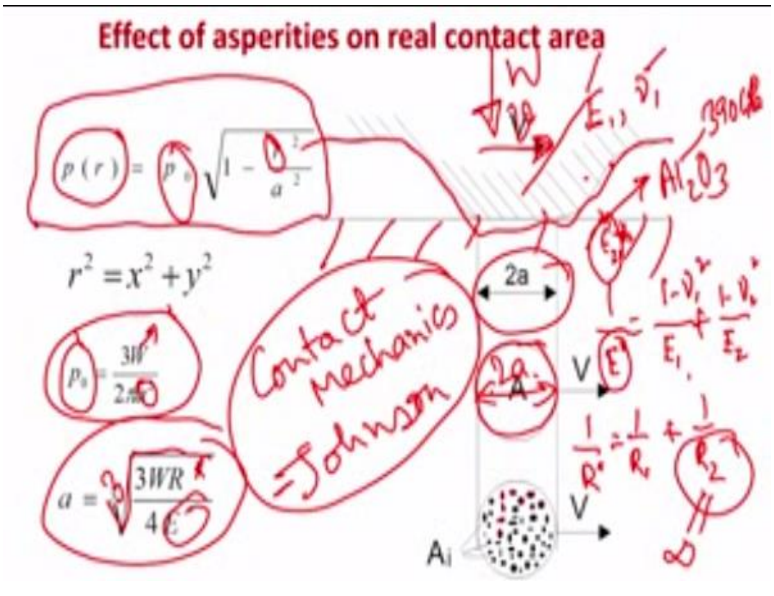
As I said before, the sliding of the asperities what happens? you have that local asperity-asperity junction that forms during friction and these asperity-asperity junctions that will contribute to the wear of the materials and this real area of contact is

$$A_r = \sum_{i=1}^{n} A_i$$

A_i is your instantaneous cross-section area of the asperity.

So, all these asperity-asperity contacts will contribute to a real area of contact. But nominal area of contact is the A_i at individual asperities.

(Refer Slide Time: 19:34)



Now, comes to is contact mechanics. For contact mechanics that I said in the first lecture that Kenneth Johnson's book is the bible of contact mechanics. What is being shown here? This is the nominally flat surface and this is one of the rough surface. These goes in the relative motion V and this is pressed against P as a load or W as a load. So, this you can see this is replaced by W.

This is your total asperity, this is your nominal asperity area A, and this is your individual spot. This spot is essentially asperity-asperity contact, this is the real area of contact. Now summation of all these spots like features that will constitute your real area of contact. Now how to find out that what is the contact radius a? Suppose if this is 2a,

$$a = \sqrt[3]{\frac{3WR^*}{4E^*}}$$

$$\frac{1}{E^*} = \frac{1-\vartheta_1^2}{E_1} + \frac{1-\vartheta_2^2}{E_2}$$

So, this is your E_1, ϑ_1 and this is your E_2, ϑ_2. So E is your elastic modulus and ϑ is your Poisson's ratio. Poisson's ratio in most of the solids is close to 0.3. Now if you consider for example E_1 is your steel ball where elastic modulus is 210 GPa and E_2 is your ceramic for example alumina where elastic modulus is 390 GPa or vice versa like you know this is your alumina ball and this is your steel flat. So, in both cases you can find out that what is the effective modulus E* and these E* values you have to apply here.

Similarly,

$$\frac{1}{R^*} = \frac{1}{R_1} + \frac{1}{R_2}$$

What is the R_1? R_1 is your radius of curvature of this particular asperity. For flat surfaces, R_2 is nothing but infinity. So, if you consider these flat surfaces, the bottom surfaces, the R_2 value is infinity. So now, I have defined that a value that is the contact radius, so 2a is contact diameter. So, 2a is nothing but this right contact diameter. From there, you can find out the contact pressure.

Contact pressure according to the basic contact mechanics theory.

$$P_o = \frac{3W}{2\pi a^2}$$

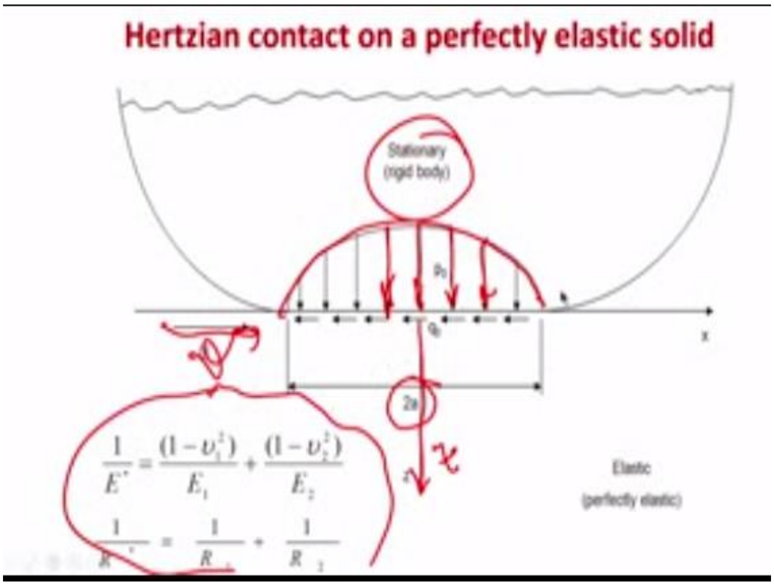
where W is your load and a is your contact radius.

How this pressure is distributed in the contact region? Pressure distribution

$$P(r) = P_o\sqrt{1 - \frac{r^2}{a^2}}$$

where P_o is your nominal contact pressure, r is the distance from the center of the contact zone and a is your contact radius.

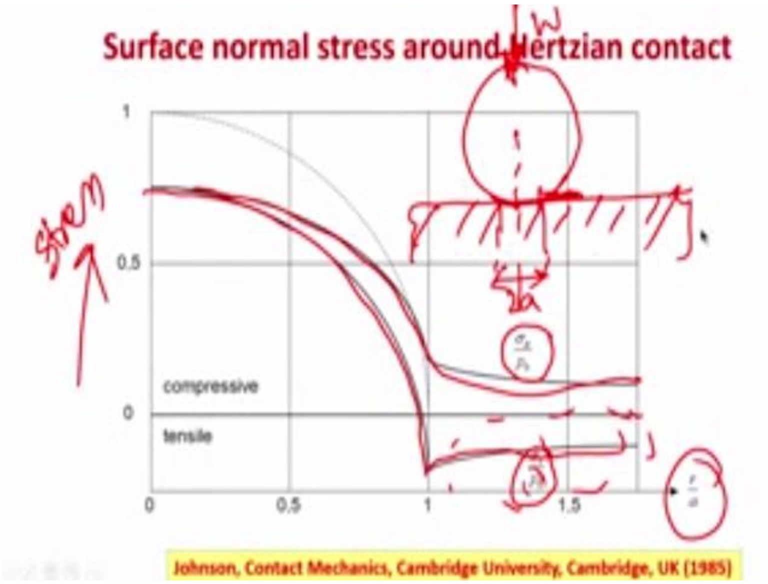
(Refer Slide Time: 23:23)



So, these are more clearly defined what I mentioned in the last slide. That how to find out an effective elastic modulus and how to find out a radius and effective radius of the surfaces. This is the case for the Hertzian contact on a perfectly elastic solid and you can see that you know how Hertzian contact that pressure is kind of varying. And you can see this is the 2a values here and this is your z and this is your traction.

So, this is your perfectly elastic solid and then top one is a stationary rigid body. Essentially the motion is here, v sliding velocity and that leads to the relative motion between the 2 solids.

(Refer Slide Time: 24:19)



Let me spent some time here just to show you that what is the surface normal stress around the Hertzian contact. Why it is called Hertzian contact? **(Video Starts: 24:31)** All these theories whatever I have mentioned in this slide as well as this slide, a scientist called Hertz first proposed these equations and on his name that we all know that this is called Hertzian contact. Let me schematically explain this again.

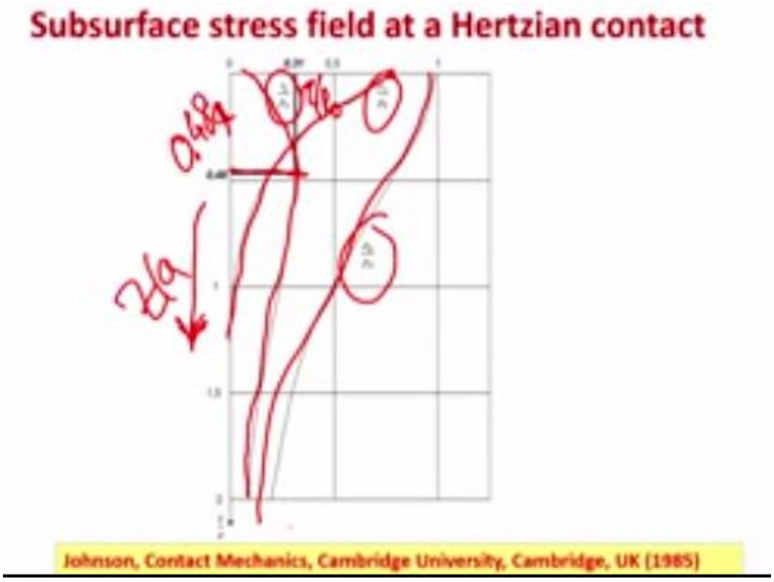
This is your nominally flat solid and this is the load W that is a kind of press against the spherical ball. So, this is your 2a, that I have defined before and this is your contact zone. If you plot it, this is your center of the contact or half way between the contact. And then what I am showing here that under this kind of configuration how this compressive stress and surface normal stress will vary spatially? Along X axis it is r/a is there and again along Y axis it is the stress is plotted. So, there are 2 kinds of stresses I have plotted here.

So, if you consider not the Cartesian coordinates, one is the cylindrical coordinates and one is the spherical coordinate. So, in case of spherical coordinates, you have rθ, so accordingly you can define σ_r and σ_θ. So, σ_r is the radial component of the stress which is normalized with respect to P_o and what is P_o? P_o I have already mentioned here, that is the normal contact stress.

If you plot it, what you see here that σ_θ/P_o, it will be going like this and σ_r it goes from compression to tension and as you grow along this region outside the contact zone, the stress is particularly tensile in nature. So, this is tensile stresses and this is the compressive stresses.

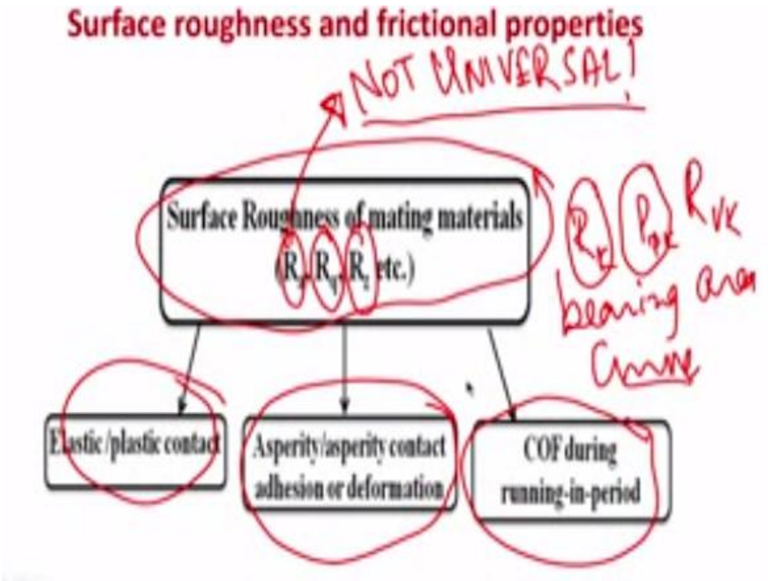
Therefore, most of the contact zone is essentially dominated by compressive stresses **(Video Ends: 27:04)**.

(Refer Slide Time: 27:05)



The second one is the subsurface stress field. Now subsurface stresses are essentially the shear stress that is the τ value, τ / P_o. Now in the shear stress, it goes way maxima at around 0.48 times a. So, along these directions, it is z/a is plotted. So around 0.48 times a, that means almost half of the contact diameter and at that depth the stress goes through maxima and σ_r value it goes like this and σ_z value. So, in these cylindrical coordinates, you have r θz. Then you have σ_r/P_o and σ_z/P_o and they will vary in the subsurface field in this fashion.

(Refer Slide Time: 27:59)



To summarize, in last half an hour or so, I have mentioned that how surface roughness of the mating materials can be effectively quantified by R_a value, R_q value, R_z values and then other values are R_k value, R_{pk} values or R_{vk} values surface roughness. Here also, I have put the bearing area curve which is very important and also, I have mentioned that although the 2 surfaces R_a is certainly should not be used as a universal surface roughness parameter.

Simply because R_a values can be same for two engineering surfaces, but the R_q values can be quite different or the ratio of one of these two parameters can be very different for two surfaces having the same R_a values. I have also mentioned that how to quantify stress distribution in elastic/plastic contact as well as asperity-asperity contact. Whereas adhesion or deformation and coefficient of friction during running in period, that I will explain some of the things in the next lecture. Thank you.