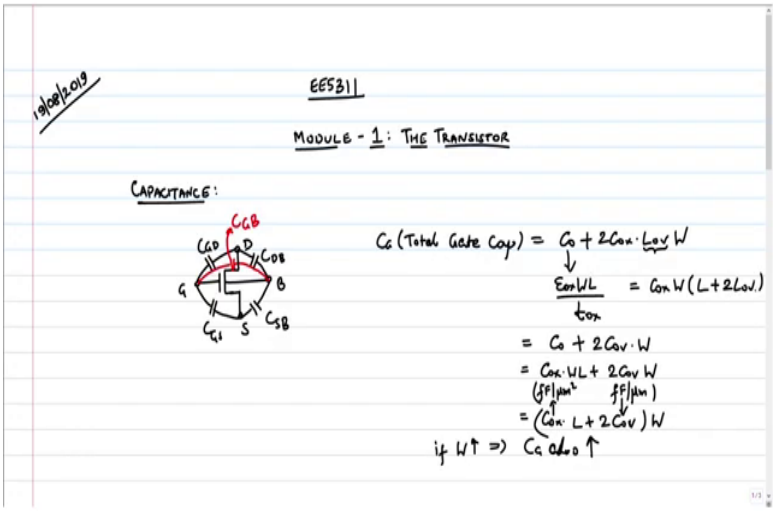
**Digital IC Design**
**Prof. Janakiraman Viraraghavan**
**Department of Electrical Engineering**
**Indian Institute of Technology, Madras**

**Lecture - 13**
**Transistor Capacitance**

(Refer Slide Time: 00:14)



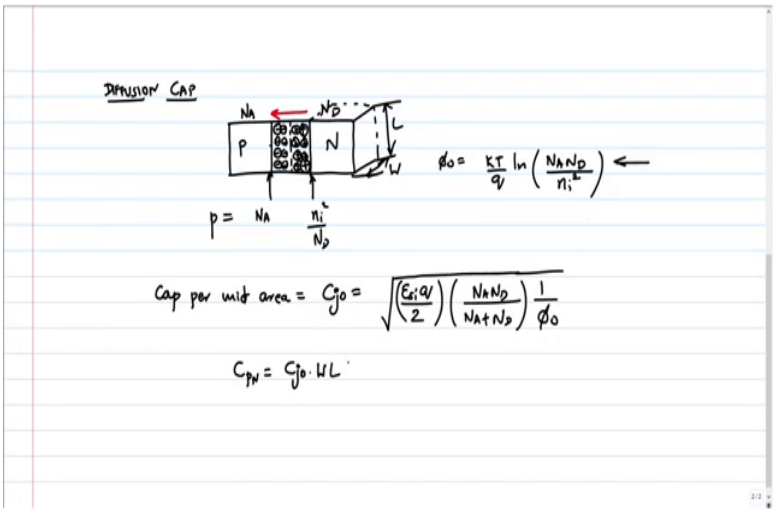

So, the discussion we had last time was the transistor is; obviously, a four terminal device, the gate, drain, source and body and you have capacitances almost across all the terminals right. This is C GD; C GS, then you have the C DB; C SB and you also have one capacitance going across like this and this is C GB ok. So, we said that the C G; total gate capacitance right, gate Cap is basically what? Just C naught right which was epsilon ox W L by t ox plus what is it? C naught into C ox into L overlap into W; 2 times perfect ok.

So, now I just left the expression here, but remember that this L overlap is basically a constant for a given technology right. Once you do the annealing and all that; this L overlap actually is reasonably fixed; it goes in by the same amount and therefore, it can be absorbed as a technology constant right. So, therefore what you do is; you basically say that this is C naught plus 2 C overlap into W ok. And this can be further written as C ox into W L plus 2 C overlap into W ok. Now, it does not matter whether you are going to read this 2 C overlap as a technology constant or C overlap as a constant ok; it is just that maybe a convention ok. So, the point is I can write this as C ox into W L; C ox into L plus 2 C overlap into W.

So, basically I am bringing out the design parameter and the technology constant separately right. And both these capacitances scale with W that is the main conclusion that we will come to right. If W increases implies C G also increases; as a designer this is what you should really know ok. Just to be careful, remember that the units of this C ox will be in femto Farad per micrometer square. It is a capacitance per unit area whereas, this guy has to be femto Farad per micrometer; it is a capacitance per unit length ok; so per micro vector; just keep this in mind ok.

Dimensionally, if you know what; if you understand the dimensions of these constants, you can easily actually predict what the actual capacitance must be; you must just multiply by either area or by some length that is all ok.

(Refer Slide Time: 04:20)




So, we were discussing the diffusion caps. So, we said that if you have a P N junction; P then you have some depletion region right. The concentration is N A here of the acceptor impurities, donors concentration is N D right? Then, just to refresh your memory what is the hole concentration at this edge here? Ah? N; NA right. So, this is N A; what is the hole concentration here?

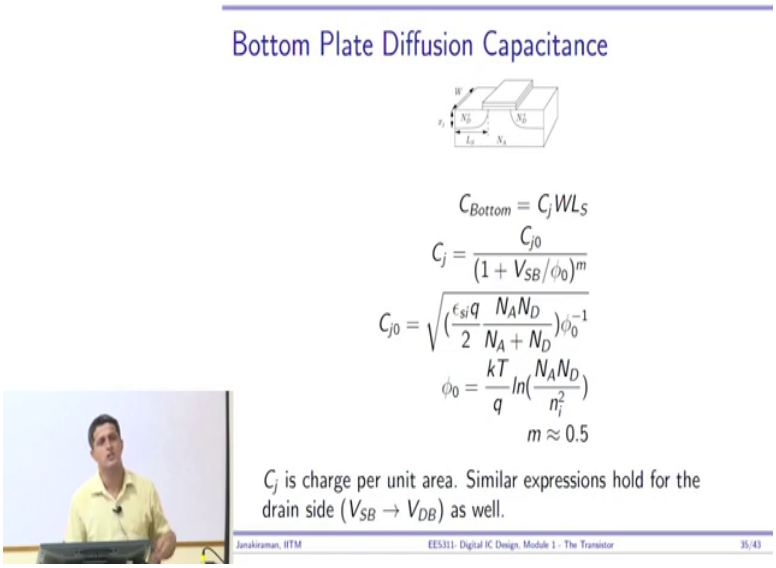
Student: (Refer Time: 05:15).

n i squared by N D right ok; n i squared by N D which is what this is whole concentration ok. So, therefore, there must be a potential drop across these two right. So, you can find out what that drop is as phi naught is equal to k T by q; lon of n i; N A, N D by n i square. Just like we did for the C MOS transistor; you can calculate what this difference is ok.

Now, this because there is a depletion capacitance here if you look at this; there is a depletion capacitance here; when the hole goes you get negative in mobile carriers here right and this is plus plus; you have an electric field from the N to the P region and that itself is like a capacitance now ok. And if you increase the reverse bias then this depletion region goes up and therefore, the capacitance will drop ok.

So, you do not have to remember any of these expressions that I am going to you know show you on the slides; now you will be given those expressions or they will all be evaluated as technology constants, you should just know where these numbers are coming from ok.

(Refer Slide Time: 06:50)




So, it turns out that the capacitance per unit area is as shown here. Let me just; maybe I can just write it down capacitance per unit area; k is equal to C j naught is what? Root of epsilon

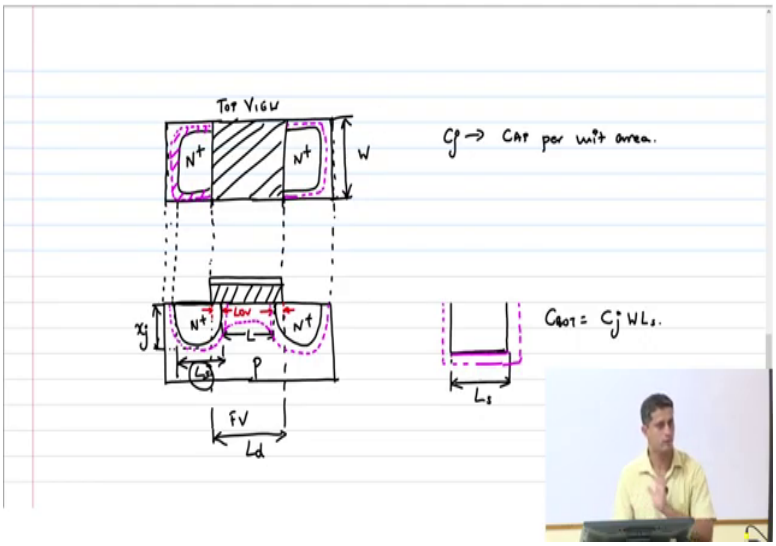
silicon q by 2 into what? Into N A; N D by NA plus N D into 1 over phi naught right and phi naught of the equation is given here right.

So, using this we now have to go back and find out what the diffusion capacitance components are in the transistor ok. So, how this gets affected by the bias; I will tell you later when the when we put down the picture of the transistor. But effectively we have a C j 0; which is a capacitance per unit area and if this has a width like this right and it has an area right.

So, what happens is; there is basically say this is my L and this is my W; then what is the net capacitance out there of P N junction? C j 0 into W into L this is what you must keep in mind ok. So, now, the complication in the transistor is not really the formulae; it is just imagining and visualizing where these diffusions are because we are basically looking at it in a 2 D plane, but the transistor is actually a 3 D picture.

So, to get that I will try to illustrate it with some figures here and with that you should be able to easily calculate what the net capacitance is ok.

(Refer Slide Time: 09:21)




So, I will go back to some engineering drawing things; where I look at the top view and the front view with that we should be able to construct the; this is the front view of the transistor. I have my gate oxide and my diffusion; why does the diffusion go inside? Because of annealing right; that is what gives rise to the overlap N plus N plus. So, let us assume that this width is x j and this spread here is L S ok. So, remember this small delta here is what? L overlap and the channel length effectively has now come down; it is going to be between these two points.

So, in fact if you go back to the expression of the gate capacitance that we wrote; you might now ask why is this not just; you know if I remove this from where from here ok. I can write this expression as C ox into W; L plus 2 L overlap right. So, you might think that; well this is effectively the same channel length right, but there is a difference because this channel length

that you that we are showing here is what is known as the drawn channel length; means this is the intended channel length that you wanted, but because of diffusion and all it reduces.

So, when I say L; it means the same L that appears in your current equation mu and C ox W by L. It is not mu and C ox W by L D; that L that we talk about is always what appears in your current equation ok. And that is why I am not writing this capacitance as just some L D; I am writing it in terms of W L and the overlap links, I am not combining them and writing it as L D yeah. So, now how does the top view of this transistor look like ok? This is my top view I am going to like just extend this ok; this is my top view ok. Obviously, here is where I have my gate ok; this is my oxide first tell me what this dimension is, if this is the top view? Yeah W; great ok.

Now, clearly you can see that these two regions to the left and right of the gate oxide are my N plus regions and there is now going to be a depletion region that forms around all around it right. So, if you look at this; if you look at the diffusion region here let me draw it in to magenta, it will be like this right and now going to map that diffusion region on top ok.

So, what do I do? If you look at this, I will let me just take that up further right and then I have a depletion region around it like this; do you agree? This is the N plus right. Similarly, here also I have something like that; N plus now where is the depletion region under the; on the other side on the fourth side for both drain and source, I have shown the depletion region only on three sides right; this is like a rectangle.

So, where is the depletion region on the fourth side? Yeah, under the channel; so that picture is not so straightforward. Because what happens is now when I apply my gate voltage that depletion region just seamlessly merges with the depletion region under the channel. So, therefore we do not consider that fourth side; when we are taking into account to the calculations for capacitance here that fourth side is basically accounted for under the gate ok.

Because we have already considered the gate capacitance and other things already there so we do not want to double count that here this is just one convention by the way; I am just giving you the rationale behind this kind of a convention clear? Ok. So, now, this is a P N junction

and I have a between the N plus and the P; I have a capacitance per unit area called C j; capacitance per unit area. Now, all we have to do is find out what all the area components are right.

So, where is the; if you look at this front view you agree that there is going to be one diffusion capacitance. Now, let us assume that this is a rectangle; proper rectangle ok, I am not worried about all the deformations we can correct for that later through some empirical correction. So, this is a rectangle this length is how much? L S and then we have a depletion region like this. So, what is the area under this region? So, if you look at the top view; I am asking there is a bottom plate the bottom, N plus region which is in contact with the P; I want to know what the area of that region is, yeah how much? No, no I cannot hear W?

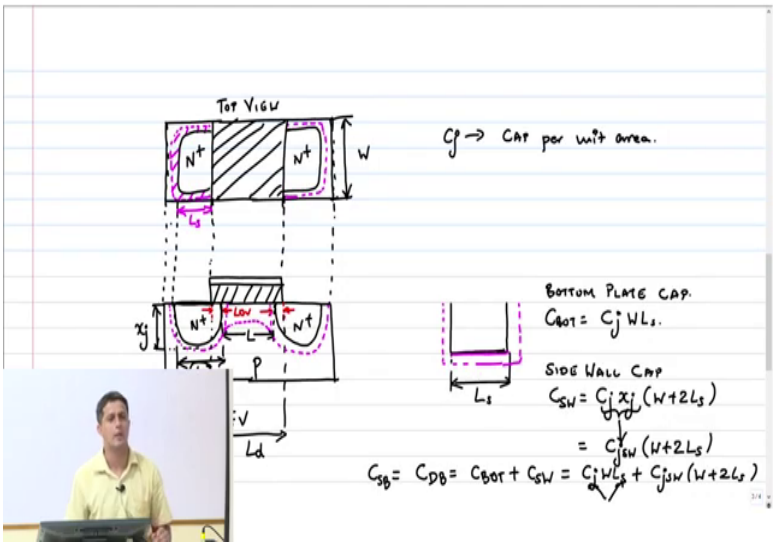
Student: (Refer Time: 16:53).

W into L S that is the area. Now, will you have a top plate component? No, because the implant diffusion implant is happening like this; s the diffusion N plus happens to the top there is no P region after that ok. The N plus cuts off abruptly here therefore, there is no top plate; you have only a bottom plate capacitance ok. So, how much is the bottom wall capacitance? C bottom is C j into how much? W into L S yeah? L S; what is L S?

L S is basically the how much of diffusion I have; here you look at this, the N plus width is what I am calling; not width, since this is the width; I will the other dimension of the N plus region is L S ok. Now, that is just one component; now there is also the perimeter capacitance which I need to consider right.

So, what is; what are the perimeter components I need to consider? I need to consider these components right because I have accounted for the bottom plate already here.

(Refer Slide Time: 18:26)




So, what is my perimeter capacitance or rather it is called side wall capacitance for obvious reasons ok; this is called bottom; bottom plate CAP or bottom wall does not matter ok. So, side wall CAP is again going to be C j ok; side wall C sidewall is huh? Exactly, it is going to be x j into what? W plus 2 L S right because this dimension is L S and this other dimension is W; total perimeter we are talking about is W plus 2 L S.

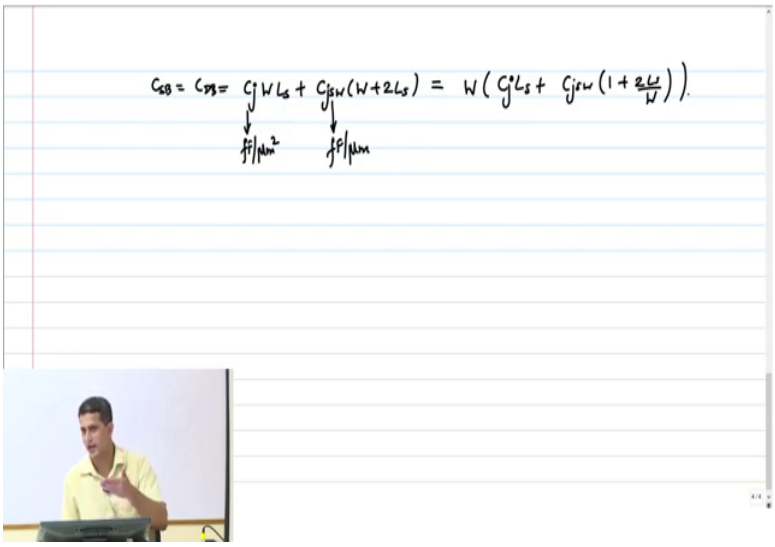
So, it is C j into x j into W plus 2 L S; now x j is also a technology constant right. Therefore, what do I do? I am simply going to write this as C j S W; side wall into W plus 2 L S ok. Both these components are not dependent on the channel length quite evidently because that dimension does not even come into the picture here; channel length is under the channel ok, it only depends on the width ok. So, again we are going to do the same trick; we are going to split this as technology constant multiplied with some design parameter.

Design parameter here is what? Only W; L is not available here even for an log race right. So, the C DB is basically C bottom plus C sidewall and of course, this is also equal to C S B; if the biases are the same right; so C SB equal to this. So, how much is that? It is C j into W L S plus C j S W into W plus 2 L S. Now, why did not they combine C j into L S? L S also is fixed right? Why did not they combine C j into L S as a technology constant?

See, what is the idea of making it a technology concert? If you I give you that one number, you do not need any more information right. Now, if I give you C j S W; I do not need to give you x j, for whatever reason you do not need x j right. In this model, I just need to give you that one number and I am done; I multiply it by W plus 2 L S, I will get the answer.

Now, if I give you C j into L S as one number do; I need to give you L S? Yes, why because I need it for the perimeter capacitance. Therefore, this you just leave it as a capacitance per unit area and multiply it by L S ok.

(Refer Slide Time: 22:22)




So, the total capacitance C SB equal to C DB is equal to C j into W L S plus C j S W into W plus 2 L S. Yeah, no my question I ask you is like I combined this into C j S W; why cannot I combine C j into L S and give you one number; say that that also I will make it as capacitance. Because now look at what is happening this guy is femto Farad per micrometer square capacitance per unit area.

This guy is femto Farad per micrometer; what I am asking you is why do I have to worry about that; why did I make this also because L S is also fixed for a technology, does not change with W and nothing ok. So, why cannot I give you that as one number; the point is you need L S in order to calculate the other capacitance; perimeter capacitance, side wall capacitance and therefore, I have to give you that number. So, it does not make sense to combine those two as one constant that is all ok.

So, this is the total capacitance and the origin of all the capacitance components that we have right. Again, the most important thing is to see is whether it is C SB, C DB or the gate capacitance; every capacitance scales with W ok. This also I can write as W into C j L S plus C J S W into 1 plus 2 L S by W ok. If you leave out that 2 L S by W; if W's are large right, then typically into a first order everything just scales with W or that is effectively like a bias term right; it is y equal to m x plus C; that is the C of the equation. W is basically the slope of the equation that we are talking about ok.

So, as a designer this is what you need to know; if I scale W all the capacitances will scale linearly (Refer Time: 24:59) right. So, if I if you tell me that I will just make the size of the transistor much bigger; then be aware that you are going to actually load the previous guy significantly more. On the other hand, if you increase W like we discussed in last class, the resistance drops; our equivalent is proportional to 1 over W; the resistance drops, there is no doubt about it.

So, this fight between resistance and capacitance is what is going to dominate your delay and there is therefore, going to be a sweet spot somewhere; you can optimize your circuit to get minimum delay. Because if I increase W; the resistance is coming down, but capacitance is going up. Of course, if I multiply these two components; it seems to cancel out which is true; the what is known as parasitic delay of a transistor is to first order is independent of W. Parasitic delay is due to its own capacitances; whatever I showed you now are the transistors own capacitance; it is not any load capacitance; so they are called parasitic capacitances.

So, parasitic delay is to a first order independent of W right, but of course, because I have a load guy coming in; it brings in some very interesting optimizations that I can do we will look at all of that later on in this course ok. Any questions here ok? Just the final thing, how does this C j scale with the applied potential ok? So, if I apply V SB right; then this basically becomes C j into 1 by 1; V SB by phi naught, where phi naught is given here power m and m is typically 0.5 ok. This again is an empirical fit and you do not need to remember this equation; just remember that if I reverse bias fit for with a higher voltage, the capacitance drops.