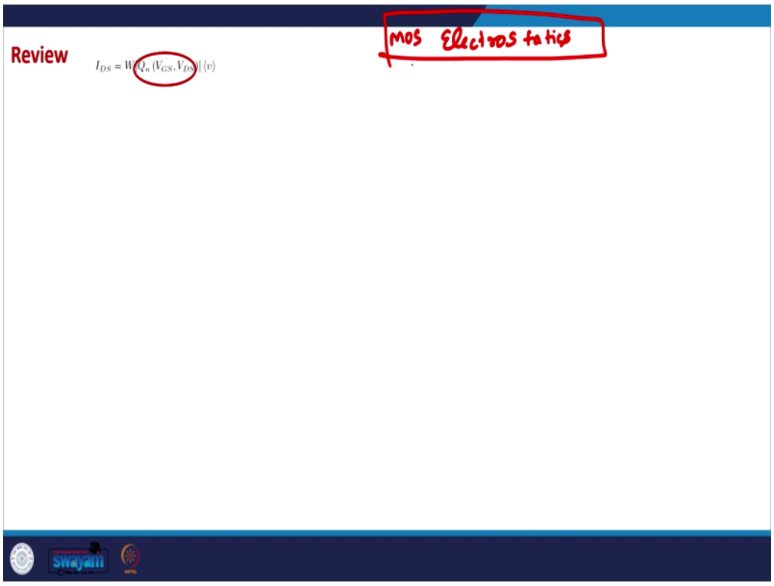
Physics of Nanoscale Devices

Prof. Vishvendra Singh Poonia
Department of Electronics and Communication Engineering
Indian Institute of Technology, Roorkee

Lecture - 51
MOS Electrostatics

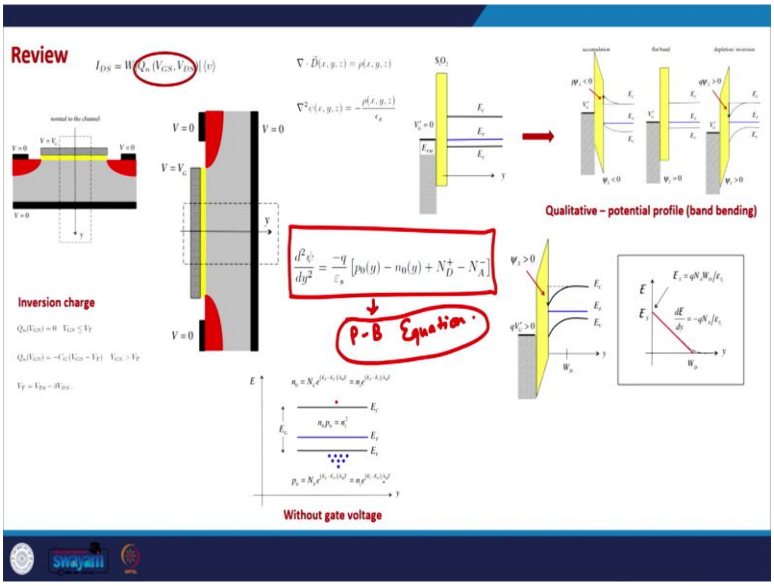
Hello everyone, as all of you know that we have been discussing about the MOSFET Electrostatics and in this video as well in this class as well we will continue that discussion. And the title of the class is MOS Electrostatics, which means Metal Oxide Semiconductor Electrostatics. So essentially, we are right now concerned more about just the gate oxide and semiconductor part, we are not dealing or we are not going into the details of the effect of the voltage applied on the source and the drain.

(Refer Slide Time: 01:01)



So, that is why we can write it as MOS electrostatics as well. And once we understand the MOS electrostatics, then this can be easily extendable to extend it to the MOSFETs as well.

(Refer Slide Time: 01:25)



So, let us quickly review what we have seen so far and we know that the charge in the channel which is this Q_n charge, it depends on V_{GS} and V_{DS} and that is why we try to understand the effect of both V_{GS} and V_{DS} on Q_n. And that is why essentially that is the central topic of the electrostatics and this is how our device looks like and in most of the conventional devices generally 1D electrostatics is considered.

1D means only in y direction, because the electric field in y direction is extremely strong as compared to the electric field in x direction, that is the case in conventional MOSFETs not in the nano MOSFETs.

And in the conventional MOSFETs what happens is that the voltage drops or voltage changes very abruptly in this part of the channel, in this part of the semiconductor which is the channel part of the semiconductor.

And that is why the electric field is extremely strong and in order to visualize this 1D electrostatics, we sort of tilt the device in this way. So, that this y direction is becomes, y direction is the horizontal direction and that is how we try to understand or we try to visualize the electrostatics in a in a MOSFET or in a MOS as well, MOS device.

So, as I told you that that one of the most important visualization tool in electrostatics is the semiconductor bands. And so in order to understand the electrostatics or in order to understand how the gate voltage is impacting the charge in the channel, we see how the gate voltage is changing the bands in the channel or how the gate voltage is changing the bands throughout the semiconductor in the y direction.

And that is what this kind of band diagrams become extremely important in electrostatics discussion. And these are the general assumptions we have seen in our previous classes that generally it is assumed that that the inversion charge is 0, when V_{GS} is less than V_T, inversion charge is given as C_G times (V_{GS} - V_T) and negative sign is there. Because this is the negative charge, we mostly talk about N-MOSFET and that is why this charge is mostly negative charge, when V_{GS} is greater than V_T.

And in order to account for the effect of the drain voltage, this is how the threshold voltage is in many cases written as. It is written as V_T is equal to V_{T0} which is the threshold voltage when there is no V_{DS}, - δV_{DS}. And this delta is the DIBL coefficient, in that is how it is defined.

And in order to understand the electrostatics in MOSFET, this is the equation that we start with this is known as the Poisson equation. And according to this equation the displacement electric field is related to the distribution of charge in the space. And that is how we can relate the potential electrical potential to the charge distribution in a certain region of the space.

And from starting from here, we derive what is known as the Poisson-Boltzmann equation for the MOSFETs and that is essentially that is that basically starts with this kind of formulation and by putting various values of various these parameters this $p_0(y)$, $n_0(y)$ we come to the what is known as the poisson-Boltzmann equation.

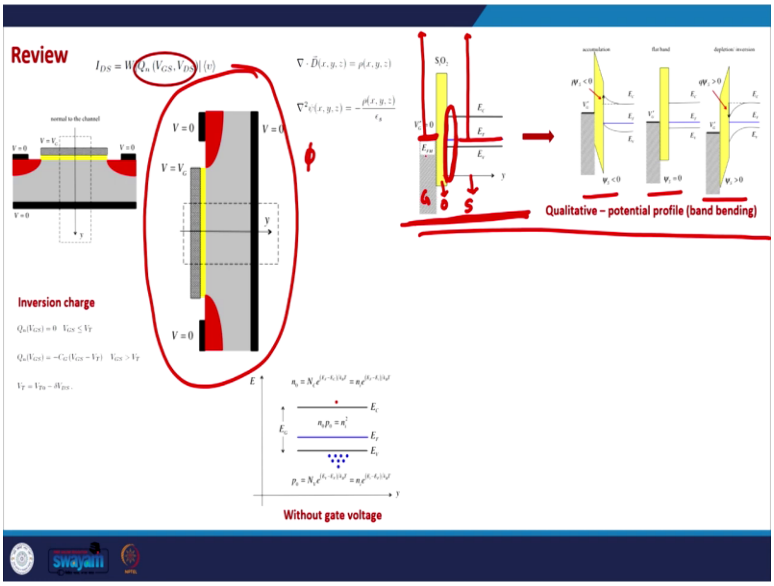
And that is one of the most important mathematical tool with us to solve the. So, this is not the Poisson-Boltzmann equation, from here we derive the Poisson-Boltzmann equation and this becomes one of the most important mathematical tool to find out the charge and the potential in the device.

And generally, the Poisson-Boltzmann equation cannot be solved analytically so that is why the numerical solutions are generally used. So, that is about the mathematical approach, but in most of the cases or large part of discussion on electrostatics is about understanding how things are working. And for that we need a visualization of how things are working in the channel when a certain gate voltage is applied.

And this is as I told you that understanding the bands in the semiconductor, this become an extremely important part. In order to understand the electrostatics we also apart from the mathematical approach we need a visualization of how electrostatics is working in these devices. And that is why we used we use the bands to visualize the impact of the gate voltage on the channel essentially.

And these three plots are extremely essential plots, extremely important plots I would say if we try to understand the electrostatics in a MOSFET. This is the bands, various bands when there is no gate voltage ok.

(Refer Slide Time: 07:10)



And here as you can see, this is the gate this is the oxide and this is the semiconductor and these bands are for this device. So, this area, this region that is very close to the oxide, this is the channel region, this is where the channel will be formed. And here as you can see that there is an implicit assumption, that this the Fermi level of the gate, the Fermi level of the gate metal is aligned with the Fermi level of the semiconductor ok.

So, which means that the distance or the energy difference between the Fermi level of the metal and the vacuum level and the energy difference between the Fermi level of the semiconductor and the vacuum level is exactly the same. So, as some of you might recall that we define the work function of a material as the difference between the vacuum level and the Fermi level. So, in this band diagram, it is assumed that the work function of the metal of the gate metal is same as the work function of the semiconductor.

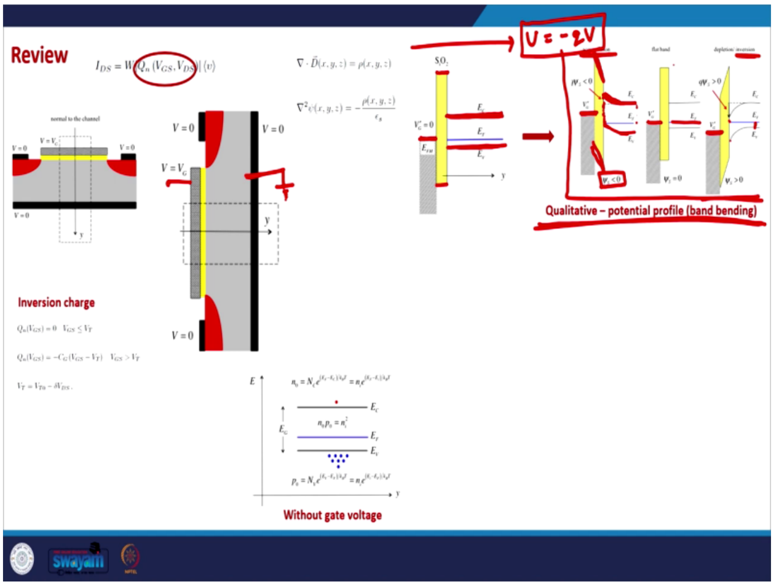
So, that is why there is this what is known as the flat band condition in which the bands are flat when the gate voltage is 0 and the Fermi level of the gate is aligned to the Fermi level of the semiconductor, then we have the conduction band of the semiconductor and the valence band extrema of the semiconductor. So, this is the bottom of the conduction band this is the top of the valence band.

As you can also see that the oxide also has the conduction band minima and valence band maxima, but these this band gap, so called band gap in the oxide is extremely large as compared to the semiconductor. And in this case, if we apply a negative voltage on the gate material what it does is that it will attract the holes in the semiconductor to the interface and the bands will be bending upwards.

Because the Fermi level will always remain uniform because there is no current in the y direction in this direction there is no current. So, the Fermi level in various materials will always be straight lines.

So, and that is why the conduction band minima in the semiconductor will move up, because now the whole concentration is increasing in the semiconductor because of an applied negative voltage in the metal.

(Refer Slide Time: 09:58)



Or in other words what we can say is that because of the negative voltage on the metal, the potential energy of electrons in the metal is increased. Because the potential energy of electrons is -q times applied voltage and that will increase the Fermi level as well. So, this Fermi level of the metal goes up as compared to the Fermi level of the semiconductor and along with these Fermi levels the band edges also go up in the semiconductor.

So, the valence band maxima and conduction band minima, they also go up along with the this gate metal Fermi level. And the reason that they go up is that in order to in, so if a negative voltage is applied on the gate material positive holes will accumulate at the interface and that will essentially increase the distance between the E_C and E_F.

And that will decrease the distance between E_V and E_F. So, this is how the bands bend and this is the situation in accumulation, accumulation is when the majority carriers in the semiconductor accumulate in response to a gate voltage. And in this case this quantity Ψ_S which is essentially the surface potential in the semiconductor or which is essentially the electrical potential at the interface this is a negative quantity.

Because the bands are bending upwards so the potential needs to be negative there and in other words, we can also understand it because the V_G is negative, so this Ψ_S will also be negative because the voltage is 0 deep inside the semiconductor. Because the other terminal of the semiconductor is assumed to be grounded.

So, whatever voltage is applied on the gate side, it drops from this V_G value to 0 voltage deep in the semiconductor ok. So, this is accumulation when we apply a negative voltage in an NMOS, in case of 0 voltage this is known as the flat band situation.

So, this E_F of the gate and the E_F of the semiconductor are aligned E_C and E_V are like a p-type material and this oxide also has the sort of the flat band edges condition. In this in the accumulation case, as you can see that along with the band bending in the semiconductor there is a kind of bending in the oxide as well.

Because now there is an electric field across the oxide and that will essentially bend the bands in the oxide as well. So, now, this third figure is the figure when the applied voltage is a positive voltage and in response to a positive voltage what happens is that this the gate Fermi level will go down.

Because according to this relationship if v is positive u is negative. So, the potential energy of the electrons is decreased. So, this E_F of the gate will go down and along with this the bands of the semiconductor will also be bent in the downward direction, that is one way of understanding this. And the second way of understanding this is that when a negative voltage is applied on the gate terminal it will attract or it will repel holes in the semiconductor or it will attract electrons to the interface.

So, that will initially create a depletion in the semiconductor because the majority carriers are getting repelled from the interface and this depletion essentially means that the number of holes are going down and number of electrons are increasing.

So, this conduction band will come close to the Fermi level and the valence band maxima will go away from the Fermi level. If we apply a large enough negative voltage what happens is that this Fermi level is quite close to the conduction band right at the interface. Far away in into the bulk semiconductor, the Fermi level is close to the valence band of the material.

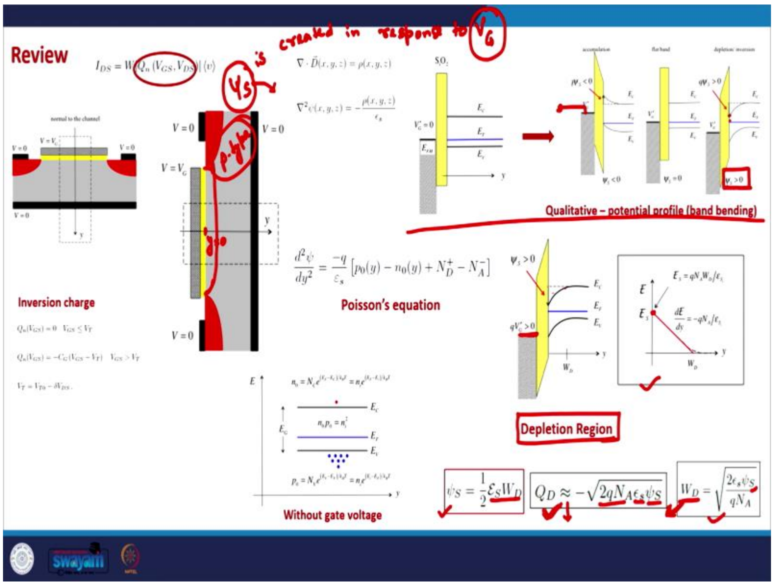
So, this is a peculiar situation, because at the interface the Fermi level is extremely close to the conduction band minima and deep inside the semiconductor the Fermi level is quite close to the valence band maxima. So, what it means is that, deep inside the semiconductor the material is p-type because the E_F and E_V are very close to each other which means that number of holes are large.

And right at the interface the E_C and E_F are closed, which means that the semiconductor becomes n-type at the interface or there is a large number of electrons accumulated at the interface of the oxide and the semiconductor and that is known as the inversion region.

So, this is the case of inversion and that is also what happens when the channel or negative channel is formed inside a p-type bulk material ok. So, this is a qualitative picture of the potential profile and along with this band bending also happens in the semiconductor in y direction.

So, if we properly understand this band bending in semiconductors on application of various gate voltages, we can easily find out or we can intuitively understand what happens when a certain gate voltage is applied or we intuitively can understand the electrostatics in a much better way.

(Refer Slide Time: 15:59)



So, what is also clear from here is that, that the surface potential or the electrical potential at the interface Ψ_S is related to the band bending. So, in this case when the bands are bending downwards the Ψ_S is positive ok. So, in a nutshell what we can say is we are applying a voltage V_G, a certain voltage Ψ_S is getting created in the semiconductor at the interface. So, on application of V_G we are producing Ψ_S in the semiconductor at the interface or at y equal to 0, which is essentially this point Ψ_S is created in response to V_G.

So, that is an important point because the Ψ_S governs most of the things in the semiconductors, this Ψ_S governs most of the electrostatics of the semiconductor. So, this Ψ_S is an also very important quantity and its relationship with V_G is also very important to understand the electrostatics.

So, this was the qualitative picture of the potential profile in the semiconductor when we apply various gate voltages. Then in the previous class we also saw what happens in the depletion region from the Poisson equation itself, that if we assume the depletion approximation which means that we assume that on application of a positive gate voltage the holes in this part of the semiconductor, the holes very close to the interface have depleted.

And moreover, the negative elect charged electrons are also not there in significant number, which means that the depletion has happened, but the inversion is yet to happen.

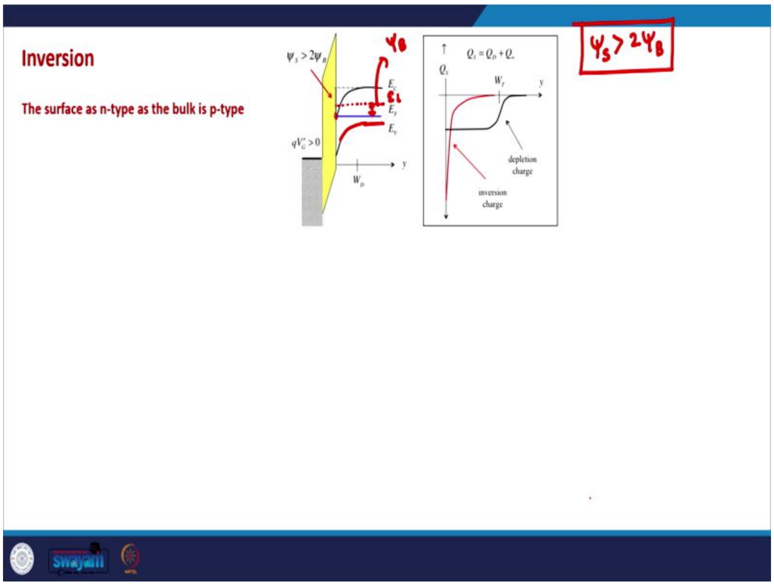
In that case we derived in our previous class that the surface potential is given by $E_S W_D/2$; E_S is the electric field which is which has this kind of profile and E_S is the electric field at the surface of the semiconductor.

And this Q_D, which is the depletion charge is given by $-\sqrt{(2qN_A \epsilon_S \Psi_S)}$ and this depletion width the width of the depletion region is given as $\sqrt{(2\epsilon_S \Psi_S / qN_A)}$. So, as you can see that, once we know this E_S we can find out Ψ_S and from here from Ψ_S we can find out the depletion charge which is the charge because of the exposed acceptor atoms, because this is a p-type material and the doping is p-type here.

So, because of the whole depletion, now these accepted atoms are left behind and that is that charge is known as the depletion charge and this is given by this expression and similarly the depletion width is given by this expression.

So, as you can see that the depletion charge depends, depletion charge depends on Ψ_S, also the depletion width depends on Ψ_S. So, Ψ_S as I told you earlier as well it is an extremely important quantity and that is why we one of the main focus of our discussion in the electrostatics part is to understand the Ψ_S when a certain V_G is applied.

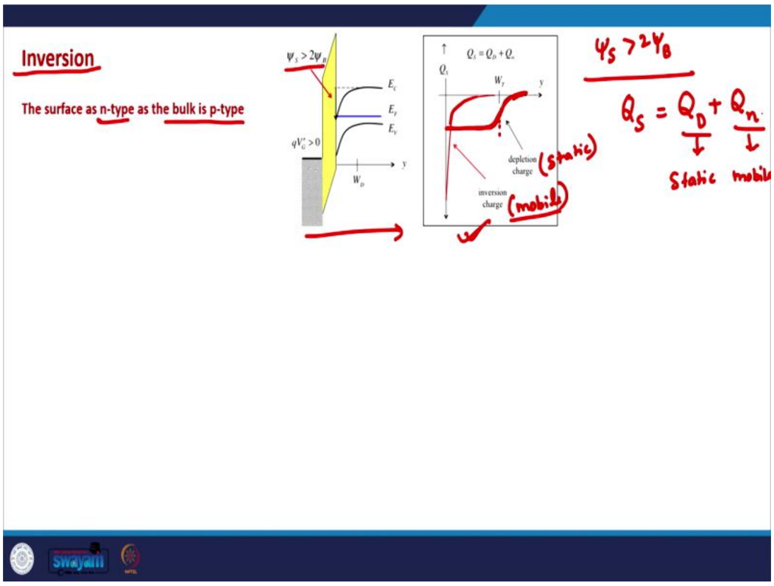
(Refer Slide Time: 19:56)



So, with this background let me quickly go to the inversion region and the inversion region is the region when the gate voltage is extremely large. Which means that the gate voltage is such that the surface of the semiconductor or the interface between the oxide and the semiconductor has become so negatively charged as the bulk is positively charged in the bulk. Or deep inside as the semiconductor is positively charged the surface is as or more negatively charged.

So, what it means is that this Ψ_S is now greater than what is known as $2\Psi_B$ ok, what is Ψ_B essentially? Ψ_B is actually this quantity which is; so this is the intrinsic region Ψ_B is the difference between the intrinsic region and the E_F which is the Fermi level of the semiconductor ok.

(Refer Slide Time: 21:23)



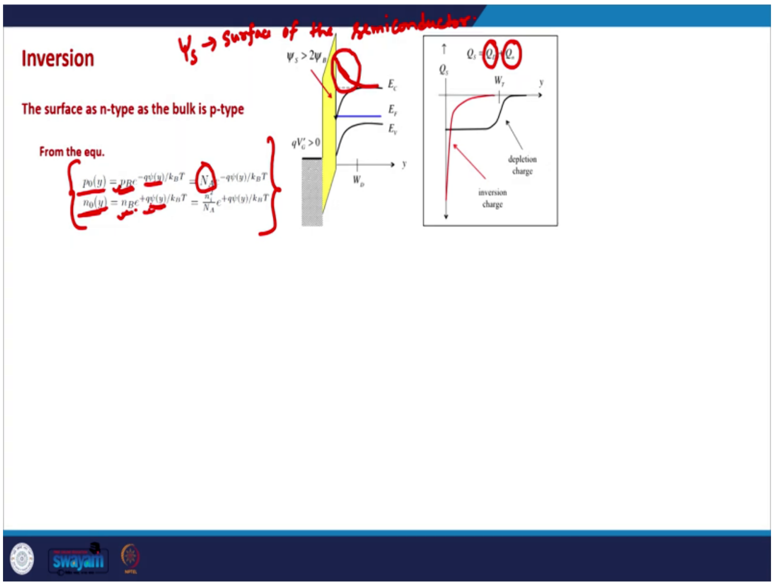
And that essentially quantifies the onset of the inversion. So, the condition for the strong inversion is, that the surface of the semiconductor becomes as n-type as the bulk is p-type. So, that is one of the conditions of the inversion. And if we plot so to say, the charge in the semiconductor as a function of the distance in y direction in this direction, this is how the charge profile looks like and here we have made a distinction between the depletion charge and the inversion charge.

So, the depletion charge is because of the exposed acceptor atoms and the inversion charge is because of the because of the inversion layer of electrons at the interface. So, that is why the depletion charge is the static charge and the inversion charge is the mobile charge so to say because this is because of the electrons in NMOS and in PMOS it would be holes.

Also, as we have seen previously that thus this depletion charge is directly proportional to $\sqrt{\Psi_S}$ and generally in the depletion region this is almost uniform and this decays like this. The inversion charge as soon as this V_G is large enough suddenly there is or not suddenly actually it is a gradual process, but gradually a lot of electrons start accumulating.

And after this condition once this Ψ_S is greater than $2\Psi_B$, there is an exponential increase in the inversion charge and that is what we will try to understand. So, this is the intuitive picture this is the, I would say the qualitative picture of the charges in the semiconductor and so that is why in semiconductor generally we write that the total charge in the semiconductor is thus sum of the depletion charge and the inversion charge. And this depletion charge is the static charge and the inversion charge is the mobile charge.

(Refer Slide Time: 23:46)



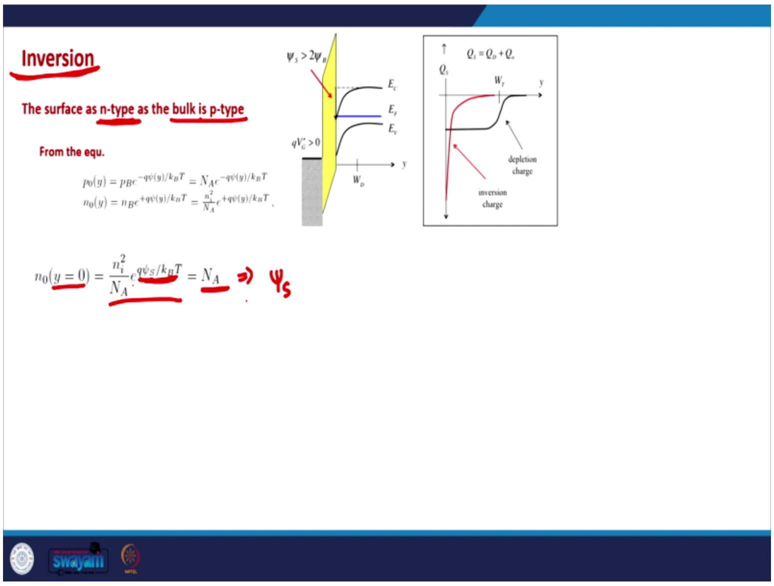
So, we will try to understand this mathematically, at any point in the semiconductor the number of holes is given has a relationship with the amount of band bending at that point. So, this $\Psi(y)$ accounts for the band banding at a point y and Ψ_S is actually the amount of band bending at the interface at the interface between oxide and semiconductor or at the surface of the semiconductor.

So, $p_0(y)$ which means the number of holes at any point is given by p_B times p_B is the number of holes in the bulk times $\exp(-q\Psi(y)/kT)$ and generally p_B which is the number of falls in the bulk is equal to the number of acceptor atoms. Because we assume that almost all the accepted atoms are ionized. Similarly, the number of electrons at any point is n_B times exp(+qΨ(y)/kT).

So, this Ψ(y) is positive which means that the bands are bending downwards, which means the conduction band is getting close to the Fermi level, in that case the number of electrons will increase and number of holes will decrease in the semiconductor. Similarly, if the Ψ(y) is negative, which means the bands are bending upwards in that case the conduction band minima is going away from the Fermi level. And that sort of says that the number of holes are increasing in that region where the bending is happening and number of electrons are declining there.

So, these quantities actually account for the mobile charge in the semiconductor and the static charge in the semiconductor is because of the depletion region because of the dopants.

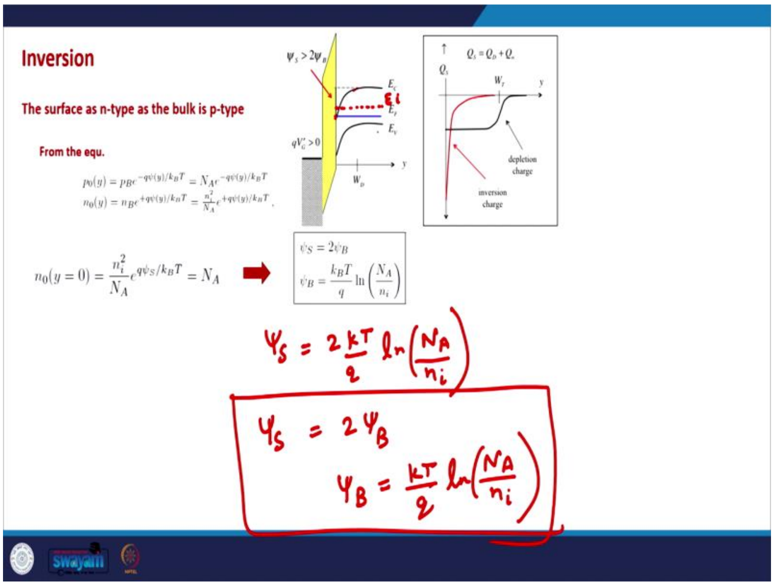
(Refer Slide Time: 25:53)



So, at the interface or at the surface of the semiconductor; so the interface between semiconductor and oxide and the surface of the semiconductor are essentially the same thing actually. So, they are being used interchangeably which is the y equal to 0 point, the number of electrons is n_i^2/N_A times exp($q\Psi_S$/kT) ok and what we assume is that and in the inversion region by definition as I told you that the surface becomes as n-type as the bulk is p type.

So, the bulk is p-type which means in the p-type bulk the number of holes is equal to the number of electrons at the surface, when the inversion has happened. So, this number is equal to the number of holes in the bulk which is N_A. So, $(n_i^2/N_A)\exp(q\Psi_S/kT)$ is equal to N_A, and from this we can easily find out the value of the psi S essentially.

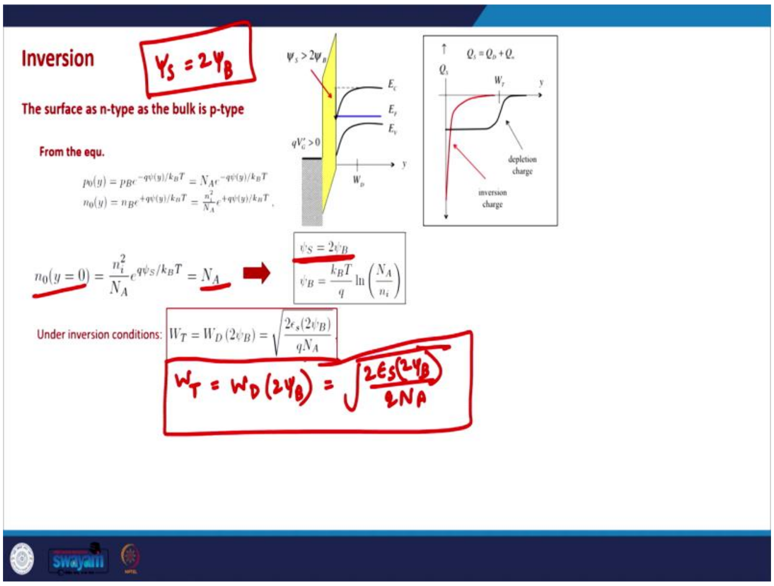
(Refer Slide Time: 26:57)



So, let me not write it is given by this quantity, Ψ_S is actually $(2kT/q)\ln(N_A/n_i)$ or this can also be written as Ψ_S is 2 Ψ_B, where Ψ_B is $(kT/q)\ln(N_A/n_i)$. And just to sort of clarify this Ψ_B is essentially the distance between or the energy difference between the intrinsic which is actually at the midpoint of the this band gap, intrinsic level when there is no doping and the Fermi level.

So, this is just a way to define the inversion because the bands have bends such that now this material becomes as p-type as the. Or so, this material at the surface becomes as n-type as it is p-type in the bulk ok.

(Refer Slide Time: 28:14)

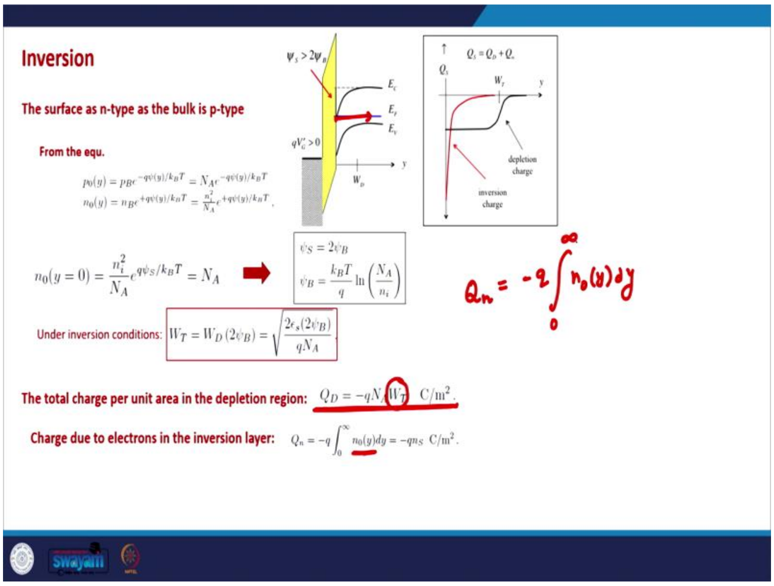


So, under just at the onset of the inversion, which means when the Ψ_S is exactly equal to 2 Ψ_B or Ψ_S or which also means that Ψ_S is such that the number of electrons at the surface is equal to the number of holes in the semiconductor.

So, even at this point, the we generally assume or generally at this point the depletion width is denoted by what is known as the W_T, which is this T comes from the threshold voltage because the gate voltage at which this happens at which Ψ_S becomes equal to $2\Psi_B$ is known as the threshold voltage.

And at the threshold voltage the depletion width is defined by W_T and it becomes W_D at $2\Psi_B$. And if we put instead of Ψ_S in W_D expression if we put $2\Psi_B$, it is essentially $\sqrt{[2\epsilon_S(2\Psi_B)/qN_A]}$. So, this is the thickness of the depletion region at the onset of the inversion ok.

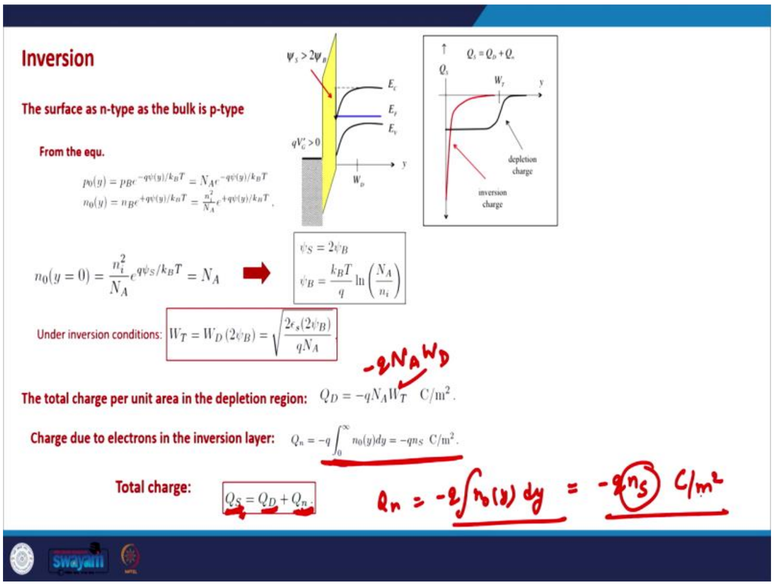
(Refer Slide Time: 29:35)



So, the total charge per unit area in the depletion region is this, this is at the onset of the inversion -qN_A times W_T, minus qN_A times the thickness of the depletion region. And the total charge due to electrons in the inversion layer how do how would we calculate that? That is essentially calculated by taking or taking an integral of n_0(y) in the y direction.

So, the charge per unit area is this or the inversion charge per unit area will be the -q integration of the electron density in the material and the integral is taking in the y direction. The limits are from 0 to ∞ which means that we are taking limit from this point to right up to deep inside in the semiconductor. But generally, we do not need to integrate to up to ∞ and with proper approximations this limit can be changed as we will see shortly.

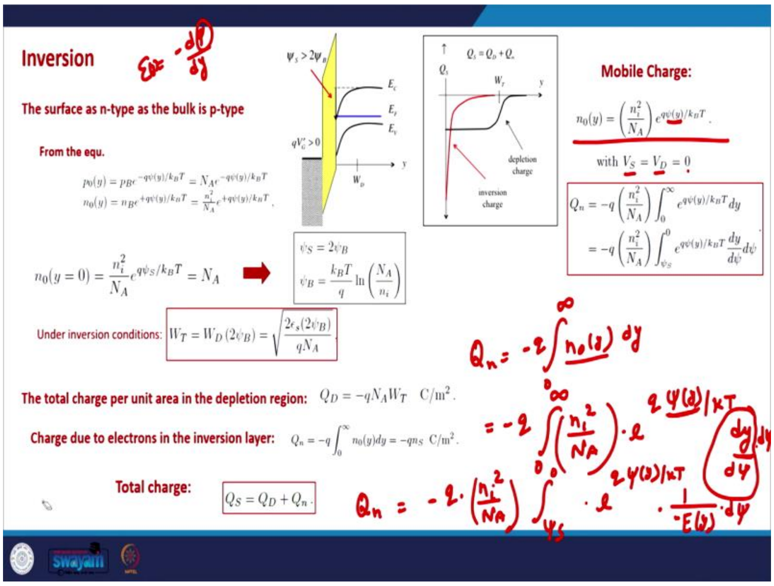
(Refer Slide Time: 30:44)



So, the total charge in the semiconductor thus this Q_S, which is essentially this S denotes the semiconductor charge is the sum of the depletion charge which is given by $-qN_A$ times W_D. And at the onset of the inversion this W_D becomes W_T plus this mobile charge or the inversion charge which is given by this quantity, which is defined as this Q n is defined as $-q \int n_0(y)dy$. And this integral is defined as $-qn_S$ which is the surface charge density and this is coulomb per meter square essentially those units are coulomb per meter square.

So, the total semiconductor charge is the sum of these two charges and these charges are again very important if we try to understand the electrostatics of the MOSFET. So, this is pretty much straight forward and we have already done a derivation for this.

(Refer Slide Time: 31:47)



Now, let us focus a bit on the mobile charge in the semiconductor. The mobile charge in the semiconductor is given by or let us start with this expression which is essentially $n_0(y)$ at any point is given by n_i^2/N_A times $\exp(q\Psi(y)/kT)$. So, this these mobile charge carriers are dependent on the potential at that point in the semiconductor this we have seen I am just review revising it so that this the importance of this point is clear.

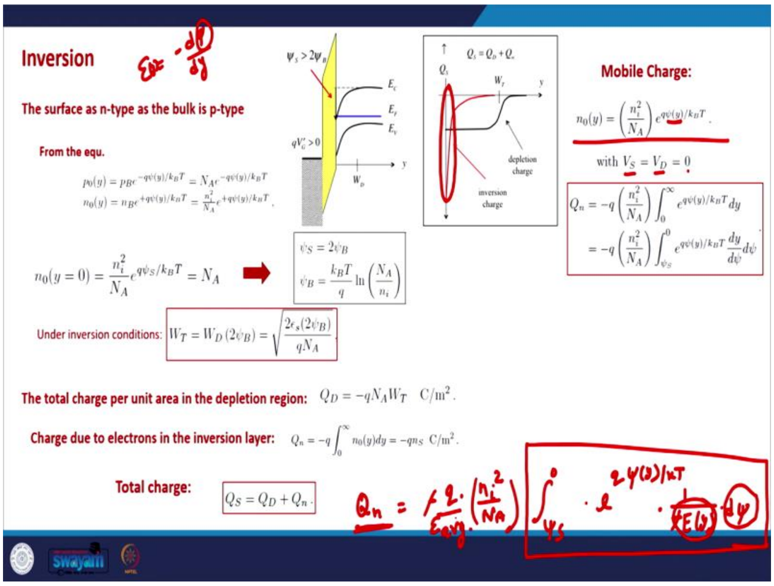
So, the mobile charge at any point where the potential is $\Psi(y)$ is given by this expression and as I told you in the beginning as well that we assume the V_S is equal to V_D is equal to 0 which means that there is no voltage on the drain and the source, we are just focusing on the gate oxide and semiconductor region.

And in that case the total charge per unit area will be the derivation of this quantity Q_n will be -q times this Q_n will be $-q \int n_0(y)dy$ integration from the surface of semiconductor to deep inside the semiconductor and if we put the formula of this it becomes -q times 0 to ∞, $n_i^2/N_A \int \exp(q\Psi(y)/kT)\, dy$. So, this is a constant can it can come out, apart from that the other point that is important here is that the exponent is $\Psi(y)$ which is a function of y and the variable over which we are integrating is dy.

So, we need to change the variables. So, what we do is we divide by $d\Psi$ and multiply by $d\Psi$. So, this thing dy by $d\Psi$ is actually the inverse of the electric field because the electric field is $-d\Psi/dy$. Because psi is the electrical potential. So, what we can say is that this is minus so this is the electric field at any point y. So, this is $-1/E(y)$. So, Q_n is -q

n_i^2/N_A, the limits will change accordingly as well because the point at y equal to 0, the surface potential is Ψ_S at y equal to ∞ the surface potential is 0 and inside we will $q\Psi(y)/kT$ this will be -1/ E(y) into dΨ.

(Refer Slide Time: 35:24)

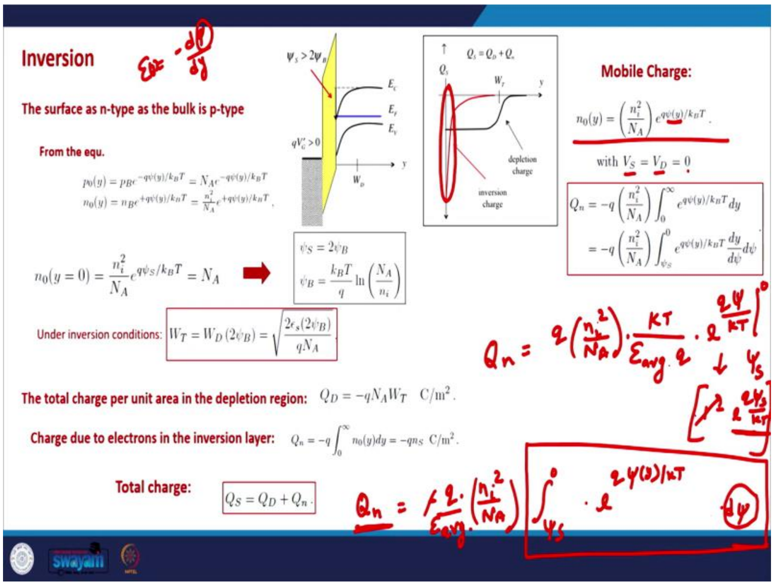


So, this is what we have now. And if we assume this integration actually this integral is not so easy as you can also see because we have exponential function of Ψ. But after changing the this variable we have made it simple, but still this -1/ E(y) is there, so this minus can cancel this.

But if we assume that electric field in the region of interest in this region that is this inversion region very close to the interface, this is not changing abruptly. Generally, because generally, we are interested in the mobile charge in a very short region, if we assume that electric field is almost constant in that region which is a fair assumption in many cases and this E(y) if we replace this E(y) by the average value of the electric field in that region, then we can take it out because it becomes a constant than.

So, this will become this E(y) will come out and this will be q by $E_{average}$, so that way this integration becomes easy.

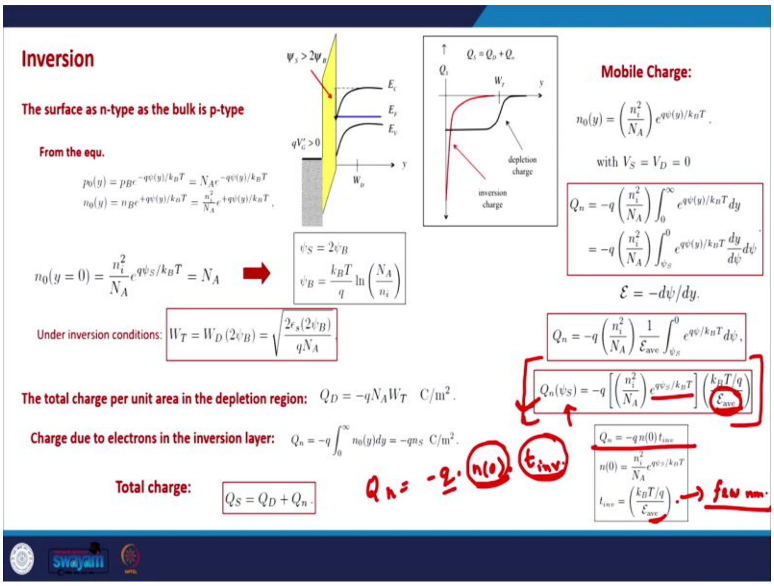
(Refer Slide Time: 36:40)



Now, we are just we just need to sort of integrate it with respect to psi y and on integration what we obtain is Q_n is q n_i^2/N_A times $1/E_{average}$, the integral gives the same exponential $q\Psi(y)/kT$ divided by q and kT will be in the limits are from Ψ_S to 0.

So, under these limits this will be; so with 0 it is it is exponential to the power 0 is 1 and Ψ_S. So, it will be this just this part will be 1- $\exp(q\Psi_S/kT)$, but generally this second part is extremely large as compared to this 1.

(Refer Slide Time: 37:35)



So, we can ignore 1 and in that case ultimately what we obtain is this, this the Q_n is -qn_i^2/N_A times $\exp(q\Psi_S/kT)$ times kT/q divided by $E_{average}$ and this $E_{average}$ is the electric average electric field in the inversion layer.

So, now this expression is quite, I would say is a systematic expression because we can compare this to Q_n is equal to -q times n(0) times $t_{inversion}$. If we compare this expression to this expression because in this expression q is the electron charge unit n(0) is the charge per unit volume in the depletion in the inversion layer and t is the thickness of the inversion layer.

And if we do that then these are I would say these are not exact expressions, but these are so to say phenomenologically this is how we try to understand the thickness of the inversion there. And if we compare this expression with this expression that we just derived this n(0) is equal to n_i^2/N_A times $\exp(q\Psi_S/kT)$ and this thickness of the inversion layer becomes kT/q divided by $E_{average}$.

And this is just a few nano meters actually. If we do for a typical electric field, if you do the calculation it is just a few nano meter of thickness of the inversion layer. So, this is how we understand the mobile charge and the inversion region in the MOSFET ok. Hope this is clear to all of you and if you closely have a look here in various conditions.

So, this expression is fairly general and in various conditions we just need to find out the average electric field and if we know that we can find out the mobile charge in various conditions. And what are the various conditions so. So, this mobile charge can be there even below threshold which means even before the onset of the strong inversion there might be some electrons at the interface.

But these electrons are less as compared to the depletion charge. So, generally we ignore that, but we which was in this expression we can find out the mobile charge below threshold as well as above threshold. The only difference will be in the average electric field in the semiconductor. So, I hope this idea of inversion or electrostatics in inversion regime is clear and we will begin with this point in the next class.

Thank you and see you in the next class.