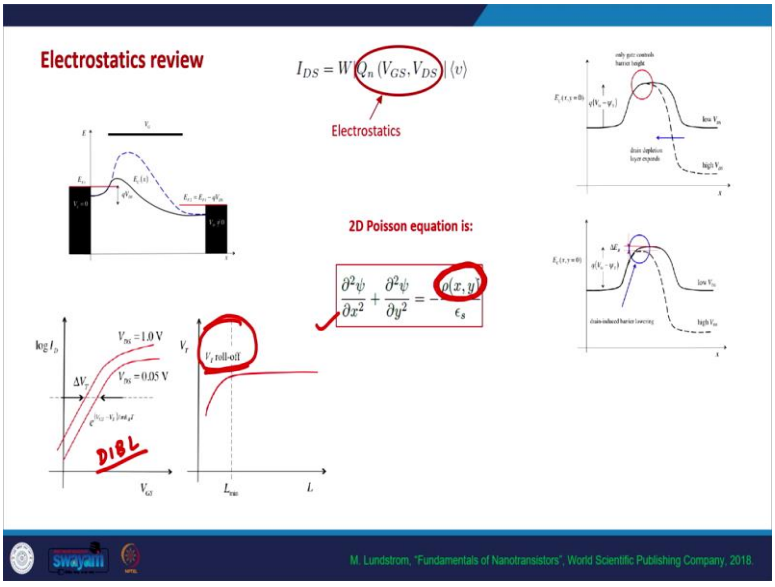
Physics of Nanoscale Devices
Prof. Vishvendra Singh Poonia
Department of Electronics and Communication Engineering
Indian Institute of Technology, Roorkee

Lecture - 54
MOSFET: 2D Electrostatics and Quantum Confinement.

Hello everyone, today we will finish our discussion on 2D Electrostatics of MOSFET and we will start understanding or we will start a discussion on Quantum Confinement in MOSFETs.

(Refer Slide Time: 00:42)



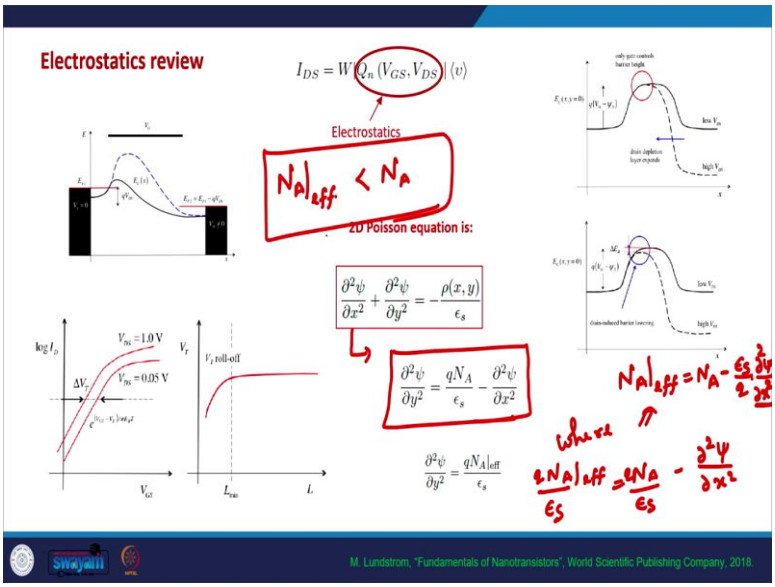
So, as all of you may recall that we have been studying the effect of small channels what happens, when we make this MOSFETs very small. In that case, we observe that there is this DIBL effect which is the Drain Induced Barrier Lowering.

And there is this threshold voltage roll off as well, which is the reduction of the threshold voltage as the channel length is reduced. And this is this happens because of multiple factors one of the factors is DIBL other factors might be the that we need small voltage to deplete the channel because the source and drain depletion regions are becoming significant part of the channel.

So, in order to account for all these things we need to understand the 2D electrostatics and that is why we start with the 2D Poisson equation. The 2D Poisson equation in the

MOSFET looks like this we need to take the derivative of potential as a function of with respect to both x and y and we need to find out the we need to know the charge density both as a function of x and y in the MOSFET.

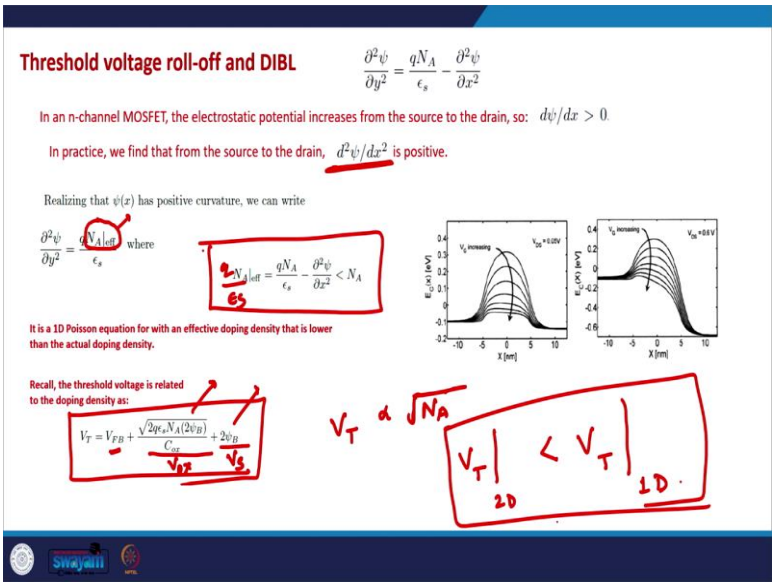
(Refer Slide Time: 01:59)



So, if we do that what we see is that we can write down this 2D Poisson equation just like a 1D Poisson equation, but now on the right hand side instead of qN_A/ϵ_S we need to write $qN_{A\text{-effective}}/\epsilon_S$, where this $qN_{A\text{-effective}}/\epsilon_S$ is $(qN_A/\epsilon_S).(\delta^2\Psi/\ \delta x^2)$ which means that $N_{A\text{-effective}}$ is $N_A - (\epsilon_S/q)\ (\delta^2\Psi/\ \delta x^2)$.

And what we also saw last time is that this quantity $(\delta^2\Psi/\ \delta x^2)$ is a positive quantity in the MOSFET. So, which essentially means that $N_{A\text{-effective}}$ is smaller than N_A the doping of the channel.

(Refer Slide Time: 03:13)



So, what it essentially says is that we can consider the 2D electrostatics just like the 1D electrostatics, but what we need to consider is that we need to consider a modified doping of the semiconductor material in the channel. And that is an important point and I advised you all of you to look into how various device parameters change if we make this change if. Instead of N_A we replace the N_A by $N_{A\text{-effective}}$ how does various parameters change.

So, here we will see the threshold voltage of the device the threshold voltage is given by this formula if all of you remember it the threshold voltage is a combination of the flat band voltage the voltage drop across the oxide this is the Vox and the voltage drop across the semiconductor.

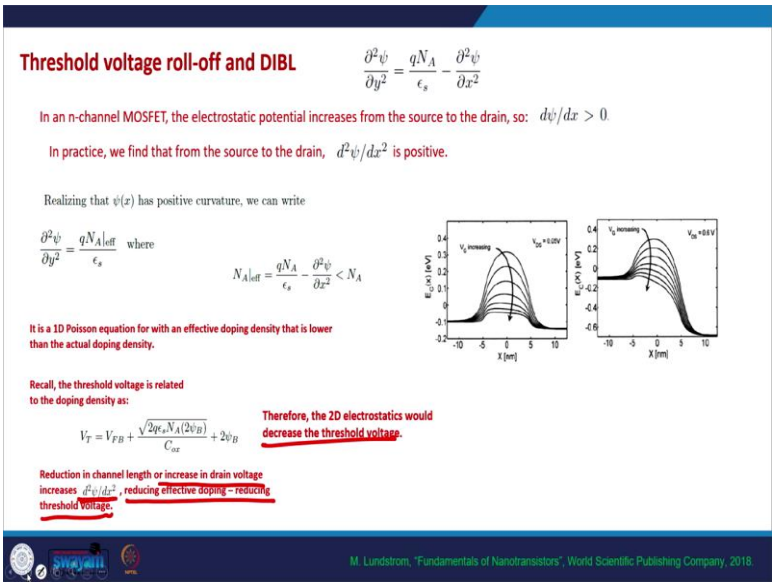
So, at the threshold voltage the voltage drop across the semiconductor is just $2\Psi_B$ and the voltage drop across the oxide is given by the is determined by the depletion charge divided by the oxide capacitance. The depletion charge as you see here or the threshold voltage itself depends on the square root of the doping of the semiconductor.

Now, if we consider the 2D electrostatics and what we would need to do is we would need to replace N_A by $N_{A\text{-effective}}$ and $N_{A\text{-effective}}$ is smaller than N_A as seen from here. So, what it means is now this V_T while we consider the 2D electrostatics will be smaller than V_T when we only consider the 1D electrostatics.

So, what it means is that for small channels where we need to consider the 2D electrostatics because this is a significant parameter there, this is a significant number there. This is not significant in long channel, so we can ignore this; we can do a reasonably well treatment with 1D electrostatics. But in small channel we need to consider this, so we need to do a 2D electrostatic analysis.

So, the threshold voltage essentially will drop down. So, this also explains the I would say the DIBL or the threshold voltage roll off that we started discussing with. And similarly various other parameters will also change the depletion width will change all of these things all of those parameters you can yourself figure out ok.

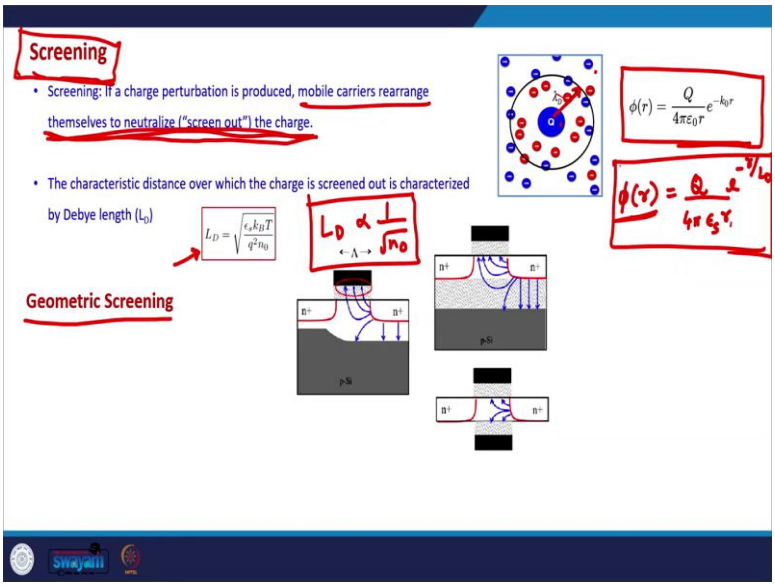
(Refer Slide Time: 05:54)



So, essentially the 2D electrostatics would decrease the threshold voltage and reduction in channel length or increase in the drain voltage this they increase this parameter ($\delta^2\Psi/\delta x^2$) and that is why they would reduce the effective doping in the semiconductor and that is why this threshold voltage roll off will be can be explained using the 2D electrostatics ok.

Now, this is sort of the shortcomings of the 2D electrostatics or short comings of the short channels. Now we will see how do we, essentially mitigate these unwanted effects. This threshold voltage roll off is an unwanted effect we will see how to mitigate this.

(Refer Slide Time: 06:53)



And in order to see that we need to understand an important concept the concept known as screening. This is an important concept in electrostatics and generally it is there in when the charges when there are lot of mobile charge carriers. It is prominent in liquid state, even in solid state where the free charge carriers are there.

So, what is screening? So, what screening says is that in a system of charges in a system of charges, which means in a system where we have mobile charge carriers, if we put an additional charge or if we create a charge perturbation. Then what happens is because of that charge perturbation or that additional charge the mobile charge carriers rearrange themselves in such a way that they will neutralize or screen out that perturbation after a certain distance.

So, in this figure here as you can see this is a system of mobile charge carriers. The blue ones are the positive charge carriers and the red ones are the negative charge carriers. What we do is we put some positive charge in the system, then what happens is that the negative charge carriers organize themselves or they rearrange themselves in the system in such a way. That after a certain distance, if we look after a certain distance the effect of this positive charge and the collective effect of the negative charge will neutralize each other.

Which means that we will not feel any potential due to this positive charge, after a certain distance. And this is shown by this circle here that in the circle as if we see there

is this large positive charge +q, but in order to neutralize this the negative charges have rearranged themselves.

In such a way that after this distance this plus q charges as if screened out. So, this also happens in solids because in solids especially in semiconductors and metals there are lot of free charge carriers electrons and holes and if we create any charge perturbation in that case that charge perturbation is screened out ok.

And that is why we cannot write the simple coulomb potential formula in these systems, where mobile charge carriers are there. And we need to write down the screened coulomb potential formula and this is how it looks like.

So, the potential due to any charge $\Phi(r)$ at a distance r is given by the Coulombs law, which says that the potential is charge divided by $4\pi\epsilon_o r$. Epsilon naught is there because we are considering the free space, if we need to consider the semiconductor instead of epsilon naught we need to replace it by epsilon s ok.

This is the simple coulomb potential formula, but in presence of mobile charge carriers, we would additionally have an exponential term which is $\exp(-r/L_D)$ or $\exp(-k_o r)$ where k_o is $1/L_D$. And this L_D is known as the Debye length of the material.

So, what it means is that with each Debye length the potential due to the charges due to a charge perturbation or due to an additional charge, decays by 1/e. In addition to its in addition to the coulomb potential formula, there is an additional decay due to the screening of the charge.

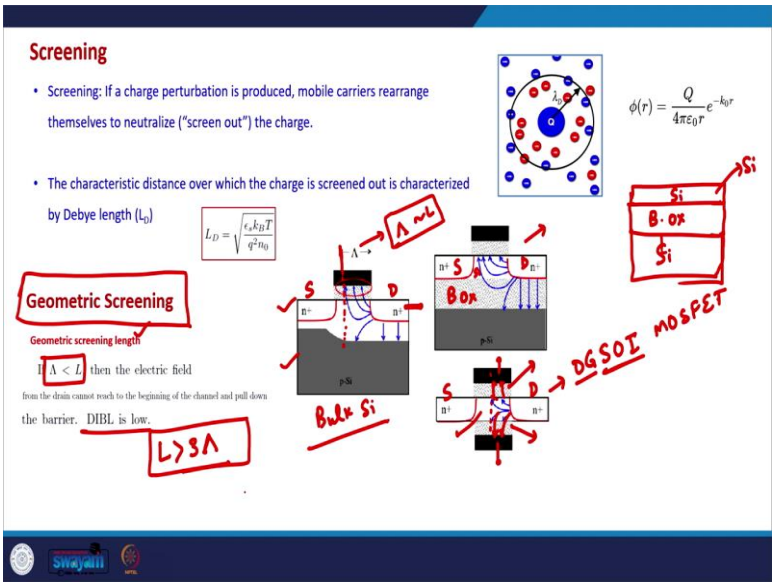
And after a certain distance this potential due to a charge will be almost 0 will be insignificant, after a few Debye lengths ok. So, this notion is important because this is true for a charge perturbation, but this is also true for any electrical field that is applied to a material.

So, when we apply an electric field in a material and if the material has mobile charge carriers, then the mobile charge carriers sort of reorganize or rearrange in such a way that the electric field lines are terminated up to a certain distance. And that is what is the technique that we use in order to, neutralize the bad effects of the 2D electrostatics.

And this technique is known as the geometric screening. Before going into the, this idea of geometric screening this Debye length for the solid states materials for the semiconductors is given by this formula. And as you can see the Debye length is inversely proportional to the number of free charge carriers in the semiconductor.

So, if there are more number of charges present in the semiconductor, the Debye length will be less; which means that the charge will be screened in a smaller distance. And the same thing is true for any electric field that is applied to the material and that is that actually gives us an interesting technique which is known as the geometric screening technique. So, it is illustrated here in these figures.

(Refer Slide Time: 12:51)



So, essentially what happens in a in a small MOSFET in a nano MOSFET. That if we apply any V_{DS} that any drain voltage, the drain voltage the electric field due to the drain impacts the barrier height, which is which means that, the drain electric field is having its impact on the source side as well ok.

So, what we can do is that whenever a drain voltage is applied whenever we apply a voltage on the drain terminal it will create electric field lines, because a positive voltage will create electric field lines starting from the drain terminal. And the nearby mobile charge carriers will rearrange themselves in such a way that these lines terminate to them ok.

So, what we need to do is, we need to have more number of mobile charge carriers very close to the drain side. So, that the electric field lines emanating from the drain terminal terminate quickly terminate at a short distance. So, the distance up to which these electric field lines persist is known as the screening length. So, in this figure here as we can see that the screening length is represented by this capital lambda.

So, here as you can see the screening length is almost equal to the channel length and this is the case of the bulk MOSFET. So, in a bulk MOSFET what happens is if we make a nano MOSFET the screening length, which means the length up to which the electric field lines from the drain will terminate is quite comparable to the channel length.

So, what it means is that the drain electric field will have its impact on the source terminal as well ok. So, the geometric screening technique is that we try to make a MOSFET geometry such that, these lines terminate very close to the drain side itself.

So, as you can if you see here in these figures this is the geometry in which the electric field lines from the drain terminal, this is the drain, this is the source are terminating; almost midway terminating very close to the drain terminal. And this geometry of the MOSFET is known as the double gate SOI MOSFET, DG SOI MOSFET. DG means Double Gate, so we have gate on top and we also have a second gate below the channel.

And the channel is just as small piece of silicon, small piece of semiconductor. On both sides of the semiconductor we have insulator or this oxide layer. So, this is known as the SOI device Silicon On Insulator.

So, in SOI devices what we do is, instead of using a bulk silicon in the body we use a very thin layer of silicon very thin layer of semiconductor. And then we have the oxide layer which is known as the buried oxide or Box. This is a very thin layer of silicon and this is grown on top of a silicon.

So, instead of using just a bulk silicon we use a thin layer of silicon buried oxide and silicon. In double gate structures what we do is we put a gate on this side as well like this. So, this is a double gate SOI MOSFET in which we have a source, we have a drain, in between we have a small semiconductor channel, on both sides of the semiconductor channel we have the oxide and the metallic gate.

So, in this case what happens is that the electric field lines that start from the drain terminal, they will terminate to the to these gate to both of these gates very close to the drain terminal. The reason for that is that this gate has lot of electrons and these electrons reorganize them themselves such that they will sort of sink most of the electric field lines to them.

So; that means, that in a in smaller devices if we can enclose the channel by gate on maximum sides on more number of sides in conventional MOSFET, we have the gate only on one side. But if we can somehow enclose the channel from more number of sides, then we can terminate the electric field lines starting from the gate very close to the gate itself.

So, that this technique is known as the geometric screening technique. And this is the DG SOI structure MOSFET this is the bulk silicon and this is the single gate SOI MOSFET. In this case what is there is source drain silicon oxide buried oxide and p silicon. This is a single gate structure.

So, as you can see as compared to the single gate SOI structure this double gate SOI structure is more effective. So, that is why we need to use these SOI MOSFETs if we want to make MOSFETs less than 10 nanometer or around 10 nanometers of channel length.

So, these kind of techniques need to be used. This is known as the geometric screening technique and this is used to neutralize the effect of the 2D electrostatics or I would say the mal effects of the bad effects of the 2D electrostatics.

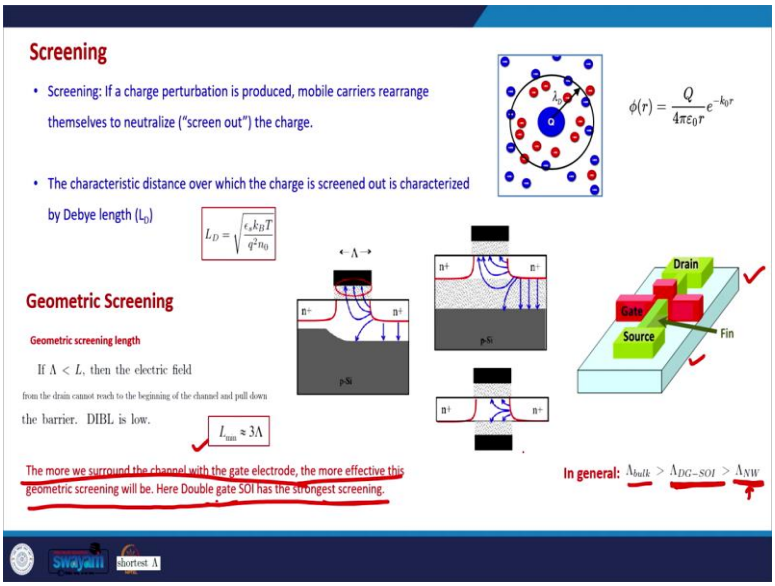
So, in this case the threshold voltage is not changed a lot and we can maintain the consistency of the threshold voltages across various devices. So, as I just told you a few minutes back that the length up to which the electric field lines are terminated is known as the screening length. And generally we try to make sure that the screening length is smaller than the channel length.

And in that case the DIBL will be low because the electric field due to the drain will not have any impact on the source side. So, the DIBL will be low ideally we would like to have this screening length or length should be greater than thrice of the screening length,

actually the channel total channel length should be greater than the thrice of the screening length.

So, this is generally one of the standards or one of the points that is kept in mind by device designers, while designing small MOSFETs while designing the nano MOSFETs.

(Refer Slide Time: 20:08)



So, in order to achieve this nowadays and some of you might be aware of this that instead of using these planar MOSFET structures. Nowadays in nano MOSFETs these kind of structures are getting are being used and these are known as FinFETs.

So, the principle is the basic principle is almost the same. Here we have a channel between the source and the drain these are the contacts to the channel and the channel is now vertical it is like a fin and the gate is now on the three side of the channel as you can see, on this left side top and right side.

So, that way the gate will be able to give a very good geometric screening to the drain electric field. So, we can make very small channel lengths in this kind of geometry. So, that is why the FinFETs are have had become extremely popular, in few nanometers MOSFET device architectures.

And there can be the gate can be on 2 sides, 3 sides or even on 4 sides. So, that is why this nano wire MOSFETs in which the gate is all around the nano wire, are even more

are even better than these double gate or triple gate structures. Because in that case the screening length is even smaller ok.

So, these are some of the novel techniques that are being used in nano MOSFETs in order to mitigate the impacts of the small channel effects ok. So, this is a typical guideline that the device designers follow that the screening length should be or the channel length should be, at least thrice of the screening length. And the basic principle is that the more we surround the channel with the gate electrode the more effective this screening will become.

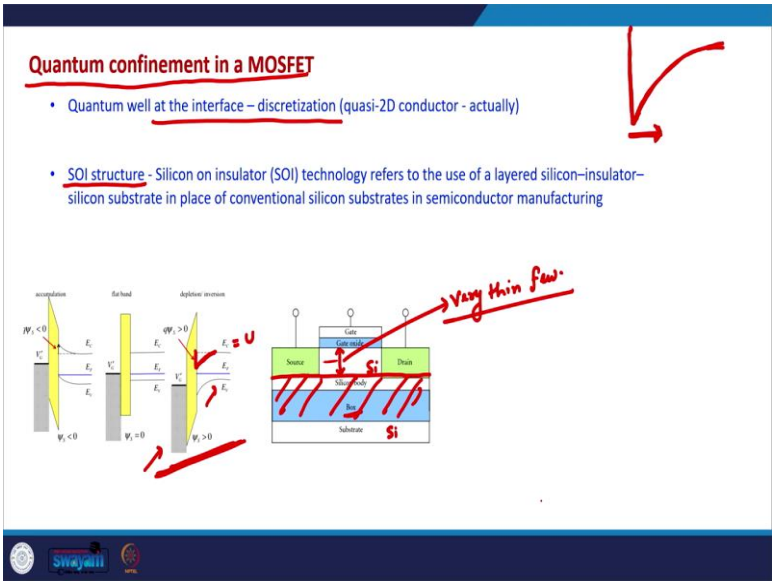
So, therefore, the DG SOI structure has the strongest screening in all these three, but this will have even better screening and a nano wire MOSFET will have even better than the FinFET screening.

So, in general the screening length of the bulk MOSFET is greater than the screening length of the DG SOI MOSFET and that is greater than the nano wire MOSFET. In the nano wire MOSFET the channel is just a nano wire and on all sides of the nano wire there is a gate.

Although these structures are difficult to fabricate, but they give us better control on the channel in small devices. So, that is all about the 2D electrostatics and we can also do phenomenological analysis of the screening length in the solids although we will not go into that here, because of the time constraints.

But in principle these are the basic concepts that we need to keep in mind while designing a nano MOSFET ok. So, with this we will now see another important consideration.

(Refer Slide Time: 23:21)



We will now to now need to understand another important effect which is the quantum confinement in a MOSFET. What is a what is quantum confinement in MOSFETs? Just take a moment and think about it, where is the quantum confinement or where is the confinement happening in the MOSFETs.

If you recall that the MOSFETs in the depletion region, by now we are familiar with this kind of structure. So, in the inversion and in the depletion region this is how the bands look like and the band profile is quite similar to the or this band profile is actually exactly same as the potential energy profile of the electrons.

So, as you see that on the interface which is also the channel region we have this kind of potential profile in the MOSFETs and these are the regular MOSFET these are even the bulk MOSFETs. So, what it means is that in the inversion regime the MOSFETs have this kind of potential energy profile at the interface of the oxide and the semiconductor which means that this is and this is also the region of the inversion layer.

So, most of the electrons stay in this kind of potential profile in MOSFETs even in the bulk MOSFETs, this happens at the interface of the oxide and the semiconductor. And this is actually if you recall our discussion of particle in a box, the basic principle is if the electron is confined it will lead to discretization of the energy levels.

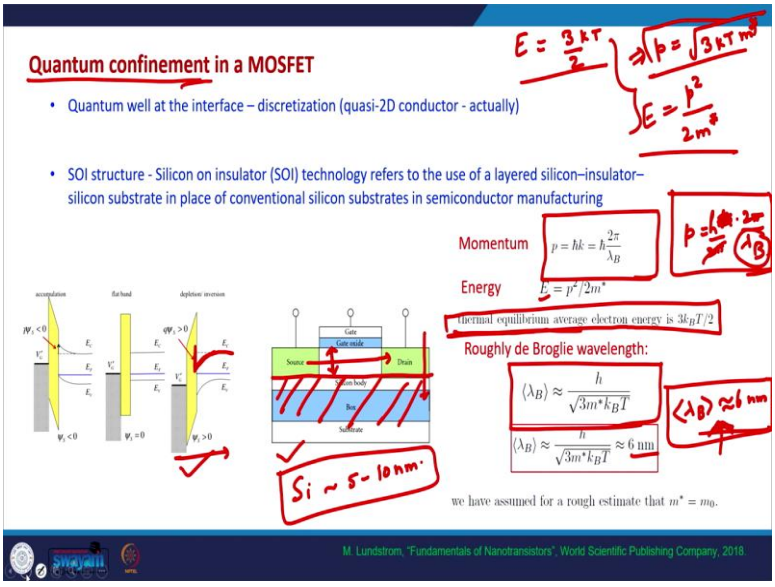
And this is a confinement. Confinement means that now the potential is sort of constraining the movement constraining the, the movement of electrons in this direction. Apart from this what we just saw in SOI structures, the SOI structures are like this. We

have a source, we have a drain, we have a gate oxide gate and in SOI structure we have a silicon semiconductor silicon here.

Then we have buried oxide which is Box and then we have again semiconductor. Most of the times this silicon layer this semiconductor layer is extremely thin. So, which means that this buried oxide is up to here. So, all of this is buried oxide and this layer is extremely thin few nanometers only. So, in this case also the electrons will be confined between the two oxides in a way.

So, this is a potential confinement this is the confinement due to the potential energy of the electrons this kind of confinement, this is a geometric confinement. So, in SOI MOSFETs there is a geometric confinement as well in regular MOSFETs there is this confinement due to the potential profile of the electrons. Now we need to see if we need to consider quantum mechanics or not in these two cases or when do we need to consider the quantum mechanical effects.

(Refer Slide Time: 27:12)



And in order to find out that let us do a simple a quick calculation, when do we need to consider the quantum confinement in the MOSFETs or when do we need to apply quantum mechanics in these confinement scenarios.

So, there are two scenarios one is because of the potential energy second is, because of the geometry and both of these are prominent in nano MOSFETs. So, let us see when

should we consider the quantum confinement in MOSFETs. For a quantum mechanical particle or for any particle the momentum is given by this relationship, p is actually equal to $\hbar$ times $2\pi/\lambda_B$. This is this straight forward De Broglie relationship.

Now, the electrons have the energy equal to $p^2/2m^*$ in the devices, where m* is the effective mass of the electrons. If we consider a very general case and if we consider the energy of the electrons, equal to just the thermal energy in thermal equilibrium. So, which means if we consider E to be just equal to the thermal energy which is 3kT/2 and this comes from the classical thermodynamics.

This is just for a rough calculation in that case the momentum is from these two relationships the momentum will be given by from these two $\sqrt{(3kTm^*)}$ ok. And if we put this value of the momentum in this De Broglie relationship we would be able to find out The De Broglie wavelength of the electrons, when the electronic energy is just the thermal energy of at a temperature T.

So, by making this replacement λ_B turns out to be equal to h bar. So, $\hbar$ is $h/2\pi$, so if we replace $\hbar$ by $h/2\pi$ this goes away λ_B is just $h/\sqrt{(3kTm^*)}$.

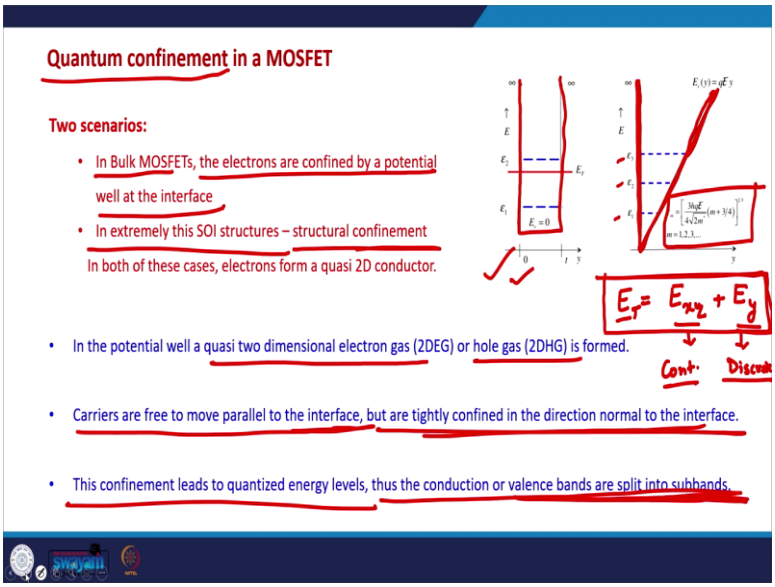
And at room temperature if we do this rough calculation what it turns out to be is 6 nanometer. So, lambda B on an average turns out to be around 6 nanometer. So, what it says is that at room temperature by a very rough calculation the De Broglie wavelength of the electrons is around 6 nanometers.

What it means is that if the electrons are confined in this order of distance, then we need to consider the wave nature of the electrons; we need to consider the quantum mechanical effects, we need to consider the discretization of energy levels. So, what it means is that if this confinement due to this potential energy profile, if this distance is around this value 5 to 10 nanometers or below that value smaller than that. In that case we need to consider the discretization of the electronic energy levels.

And similarly in SOI structures if the silicon this is the oxide if the silicon layer the semiconductor layer is 5 to 10 nanometers in that case also we need to consider the quantum mechanical effects. And the basic principle is that the confinement leads to discretization.

So, what it means is that the energy of the electrons will be discretized in this direction the direction of the confinement. So, in these 3D devices the electrons and holes at in these regions will behave like a quasi 2D systems.

(Refer Slide Time: 31:54)



They will behave like a 2D electron gas or 2D hole gas. So, these are the two scenarios in the bulk MOSFETs the electrons are confined by a potential well at the interface and the potential well looks like this as we have just seen.

In extremely thin SOI structures ET SOI structures, there is a structural confinement because the silicon the semiconductor layer is extremely thin. And in that case the confinement is like this. There is this rectangular potential well in bulk MOSFETs we have the triangular potential well.

In, so this analysis we have already done the similar analysis can be done about the energy states of the electron because of this kind of confinement, because of the electron confinement in a triangular box. And this gives us various discrete energy levels of electrons by this relationship.

So, this you can do this as a, homework exercise this we already know these various energy levels. So, in these potential well the electrons and holes will behave like quasi 2D electron gas or quasi 2D hole gas which means they are free to move in two Directions they are free to move in x and z directions.

Because if we look at the geometry of the devices this is the y direction this is the y direction and the confinement is in the y direction in x direction which is this direction they are free to move and in the z direction which is perpendicular to this surface is also the free movement of the electrons and holes.

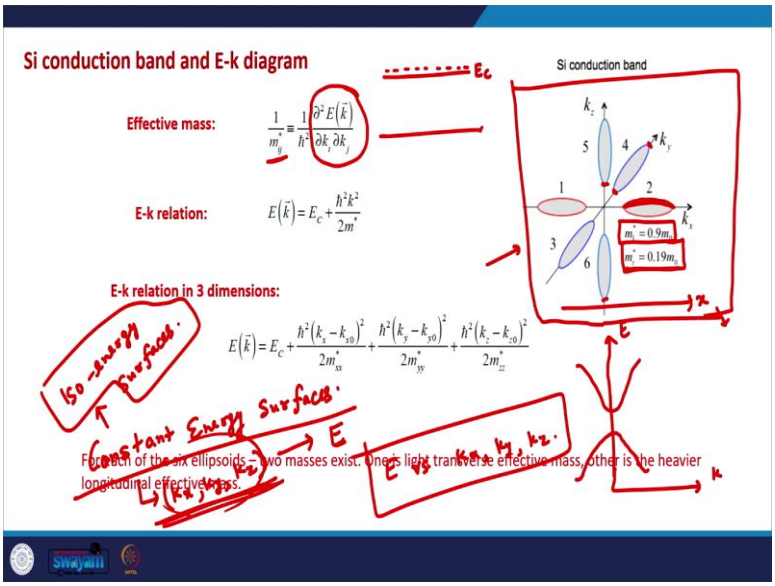
So, essentially in almost all the MOSFETs because this, the thickness this thickness of the inversion layer is only few nanometers. So, this is of the order of the De Broglie wavelength of the electrons even in the bulk MOSFETs. So, in the in almost all the MOSFETs we need to consider the quantum confinement of the charge carriers and what does quantum confinement do? Quantum confinement leads to creation of quasi 2D electron gas and quasi 2D hole gas.

What it means is just to repeat it the carriers are free to move in parallel to the interface in x and z directions, but they are tightly confined in the normal direction to the interface. And that is the case even with the bulk MOSFETs even with the traditional long channel MOSFETs. So, even in the long channel MOSFETs we need to consider the quantum confinement of the electrons and how do we consider this quantum confinement.

So, this confinement leads to quantization of the energy levels in the direction of the confinement and that is why the conduction band or the valence band or the energy bands both of the energy bands are split in the regimes of the confinement. This we have already seen this kind of confinement we have already seen in quasi 2D materials case where the charge carriers the electrons are free to move in two Directions, but they are confined in the third direction and the similar thing will happen here.

So, the electrons and holes will have continuous energy levels in two Directions x and z directions and the discrete energy levels in y directions. And the total energy of the electrons can be written as E_{total} is equal to $E_{xz} + E_y$. This is continuous this is discrete and this we need to keep in mind.

(Refer Slide Time: 36:08)



Before going into the details of this before going into the details of the splitting of the bands or formation of the sub bands, we also need to consider another important concept which is the ISO energy surfaces in the semiconductors.

So, if you remember the E k relationship for the semiconductors for 1D solution the 1D solution of the k p model this is how the E k relationships look like on the y axis we have E on x axis we have k.

And this was the case for 1D materials 1D solids. Generally almost always we have 3D materials and for the 3D materials we cannot plot the E k relationship because that will be a 4D plot because we need to consider E versus k_x, k_y, k_z ok.

So, what we do is what we plot what is known as the constant energy surfaces in order to understand the band structure along various directions, constant energy surfaces and what are the constant energy surfaces. So, we plot all those k points all those reciprocal lattice points k_x, k_y, k_z points for which that give the same energy or that correspond to the a constant energy of the electrons.

So, the which what it means is that for all these points the Schrödinger equation solution will correspond to this energy. And that is known as the, so plot of all these points is known as the constant energy surfaces in 3 directions.

So, what we just need to do is, we need to take any value of E and corresponding to that value of E we need to plot all possible combinations of k_x, k_y, k_z in the reciprocal space. Generally the most important energy for us to consider is the energy at the bottom of the conduction band because that is where most of the transport happens.

So, that is why we plot the constant energy surfaces at the corresponding to the bottom of the conduction band and this is how the constant energy surface for the silicon looks like. You can see that we have ellipsoids 6 ellipsoids in k_x, k_y, k_z space.

And these are all possible combinations of k_x, k_y, k_z points corresponding to a certain energy which is close to the bottom of the conduction band of the silicon. And why is this important because the electrons in the MOSFETs are confined in y direction and they are free to move in the z direction.

So, that is why we need to and from these constant energy surfaces what we see is that the mass of the electrons or the effective mass of the electrons, which is inversely proportional to the curvature of the E k diagram. So, the effective mass of the electron depends on the E k plots ok.

And this effective mass can also be derived from the, these constant energy surfaces. So, we would like to see the mass of the electrons in the direction of the confinement and in the direction of the or in the continuous direction as well. And if we have a closer look here and let us consider the case of x direction just the x direction. In x direction if we consider there would be two effective masses corresponding to the electron one is the longitudinal effective mass and that depends on this curvature.

So, the curvature is this curvature will govern the effective mass in x direction and the second effective mass will be the transverse effective mass, which is because of this curvature. So, as you can see the long in the longitudinal direction the curvature is less. So, the effective mass will be more and in the transverse direction the curvature is more and the effective mass will be less.

So, the longitudinal effective mass of the electrons is 0.9 times the rest mass and the transverse effective mass is the 0.19 times the rest mass ok. So, in the just to quickly review the quantum confinement in MOSFETs, we need to consider the quantum confinements because of the potential well at the interface. And in ET ISO structures

there is a geometrical confinement as well there is a structural confinement as well that is also we need to consider.

And because of these confinements the energy levels will be discrete. So, we need to consider this kind of energy state and do the analysis according to a quasi 2D structure. Now, there is a nuanced point here which is the difference in the effective mass along different directions and that is understand that can be understood from the constant energy surfaces in the silicon in the semiconductors.

So, I will stop here and I will let you think about these concepts the concepts of quantum confinement and the constant energy surfaces or ISO energy surfaces they are also known as ISO energy surfaces ok. With these points we will complete in the next class, we will complete our discussion on the on how to consider the or how to account for the quantum confinement in the MOSFETs.

So, thank you all of you see you in the next class.