

Nanoelectronics: Devices and Materials
Prof. K. N. Bhat
Centre for Nano Science and Engineering
Indian Institute of Science, Bangalore

Lecture – 18
Velocity Saturation. Ballistic transport, and Velocity Overshoot Effects and
Injection Velocity

We start our third session on the carrier transport in nano MOSFET devices and MOSFET non classical MOSFETs.

(Refer Slide Time: 00:24)



NON-classical MOSFETs

Module-1 (Lecture -3)

Carrier Transport in Nano MOSFET (continued)

Professor K. N. Bhat
Centre for Nano Science and Engineering
(CeNSE)
Indian Institute of Science
Bangalore-560 012
Email : knbhat@gmail.com

 NPTEL

1


(Refer Slide Time: 00:30)

Effects of Short channel lengths on carrier velocity and carrier transport when $V_{GS} > V_{th}$

Short channel lengths lead to higher E-field along the channel :

Increases the carrier velocity leading to

- 1. velocity saturation**
- 2. Ballistic transport and Velocity overshoot effects**
- 3. Injection Velocity**



2

So, we continue on some effects of short channel lengths on carrier velocity we just began on this last time, the 3 aspects with will be discussing will be 1 is the velocity saturation effect on the drain current and then the a quick look at what is the implication now ballistic transport and velocity over shoot effect whether we should worry about that that also we will discuss. For silicon we will see how good is important is the velocity overshoot is there then we will also discuss injection velocity it is ultimately, which is the 1 which controls the drain current in short channel devices.

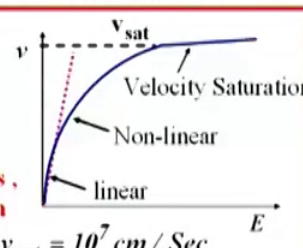
(Refer Slide Time: 01:21)

Carrier Velocity Vs Electric Field

At low fields the drift velocity 'v' is small compared to thermal velocity v_{th} . Assuming a constant collision time, τ_m


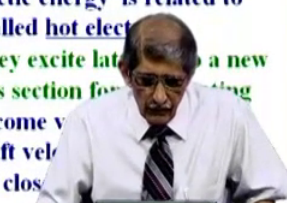
$$v = \frac{qE}{m^*} \tau_m = \mu E. \quad \mu = \frac{q\tau_m}{m^*}$$

At higher fields, collision time reduces, velocity 'v' increases sub-linearly with electric field E



At high fields, 'v' becomes comparable to v_{th} , total energy of electron increases significantly. As kinetic energy is related to Temperature ($m^*v^2/2=kT$), they are called hot electrons.

When they collide with lattice atoms they excite lattice atoms to a new vibrational mode which has large cross section for scattering of hot electrons. Thus, lattice atoms become very effective in scattering mobile electrons, causing drift velocity to saturate at a limited value v_{sat} , which is close to v_{th} .

Now, before we go into that velocity saturation, we take a quick look at the velocity versus electric field characteristics of the carriers. I used the word carriers because the

sort occurs hold good both for electrons and holes what I am showing is mainly for electrons right now, similar are given can be extended to the holes also. Now in thermal equilibrium when you do not apply an electric field, you know the pre electrons are moving in random directions because they have the thermal energy they move with the thermal velocity which is about 10^7 centimeter per second ok.

Now, even though they are moving there is no current flow, because they are in random motion and each of them cancel there in all the directions. So, there is no net current flow now when a apply an electric field at the electrons have an additional component in the direction of the electric field, that is it moves towards the plus terminal opposite to electric field direction. So, it has a component of velocity. So, the carriers move in that particular direction directed by the electric field. If it is whole it will move in the direction of electric field plus to minus if it is electrons they move from minus to the plus.

So, now these is this component is superimposed on the thermal velocity. So, you get the current due to the velocity of these electrons, which can be expressed in terms of electric field and collision time what is collision time when the electrons move in a particular direction they get scattered by the lattice atoms, they get scattered by various effects. So, in a time τ there will be collision. So, the velocity or I am sorry the electrons get accelerated by this force q into E q into E is the force on the electrons E is electric field divided by m^* , m^* is effective mass of electrons because you talk of moment of electrons in the crystal not in the free space . So, this is the acceleration, acceleration into time of collision gives you the velocity; a acceleration to time you see the velocity of the carriers at the time of collision.

So, there is a maximum velocity that gets; you can say there are different electrons between separate different acceleration and or different collision lengths and collision times. So, what we talk of this velocity will be an average value for all the electrons put together which gives rise to current. Now this thing whatever we write here v usually express as μ into electric field were μ is cyclic $q \tau$ $q \tau$ m divided by m from here, I just pull out this electric field term out and call this term as the μ that is a mobility. Now you can see the here there was an assumption that when the change the electric field the collision time was not changing. So, that is why this mobility is constant. So, v is proportional to electric field and velocity field characteristics be linear when electric

field is low. At higher electric fields the acceleration is more, the time of collision will reduce because the distance between collisions is same the time of collision or collision time will get reduced.

So; that means, the mobility actually gets reduced, but electric field is increasing. So, the velocity field characteristics does not increase linearly, but it becomes non-linear this is because of collision time production. So, even though electric field is increased it does not increase velocity does not increase as much as you think it should. When you go still to high electric fields you can see that the velocity become higher and higher and it will become comparable to the thermal velocity which is there even the thermal energy.

So, total energy of the electrons now becomes thermal energy plus this additional energy which is absorbed from the electric field. Now the kinetic energy which you call it has half $m v^2$ is related to the thermal energy or kT . So, once you say that the kinetic energy has gone up because of the thermal energy and the energy absorbed by the electric field, in fact, finally, the energy kinetic energy net energy this is through the energy absorbed from the electric field.

So, you can say that it looks as if the temperature of the crystal has gone up or temperature of the electrons as gone up, even though crystal temperature is thermal room temperature the electron energy is higher by this amount. So, you will say that they are hot electrons; a transfer not really hot it looks as if they are hot because the energy is much more than that of thermal electrons.

So, when they collide. So, you can see now these hot electrons have additional energy over and above the regular thermal energy. So, when they collide with the lattice atoms they excite the lattice into a new vibration mode, which has large cross section for intercepting the more electrons. In other words what I am trying to point out is these electrons when they collide with these lattice atoms because they have large energy; they create extra vibration energy in the lattice. So, if there is extra vibration energy the chance of collision of these electrons becomes more for example, if the atom is moving like that you know the cross section of that movement is with larger. So, what happens is almost all the electrons are collide and lose energy.

So, the maximum energy that is required by that is equal to is decided by that allowed velocity that is achieved by that, and that velocity is equal to a thermal velocity or close

to that. See when the velocity has become equal to thermal velocity have close to that the energy is almost much large compared to thermal energy. So, at that point what is available of required that is that if closed to thermal velocity it is not exactly equal to thermal velocity, but very close to that slightly less than that still we can say we saturated that velocity which at which the maximum velocity is reached its a saturation velocity.

So, what happens is if I keep on increasing electric field initially linear non-linear linear because of the collision time decreases finally, velocity saturates because all the energy it has required within that collision time is lost to the lattice. So, within that collision the time it has whatever a velocity is required is very close to the thermal velocity, which is the limiting velocity of the electrons. It cannot go beyond that because entire thing is absorbed by the lattice. So, this is the principle of velocity saturation is basic semiconductor devices this discussion comes for completeness sake, I just brought this (Refer Time: 09:38) into. So, what happens in the case of MOSFET? So, remember that linear, non-linear, saturation. If the field is and that saturation velocity is about 10 to the power of 7 centimeter per second for electrons or holes it is about 6 times 10 to power 6 centimeter per second.

(Refer Slide Time: 00:10)

Velocity Saturation Effect in Short channel MOSFET

Electric field is no longer low, when L is short

$$E_{av} = (V_{GS} - V_{Th}) / L$$

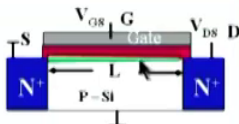
In silicon, for electrons at Fields $E > 30$ kV/cm, drift velocity saturates at its scattering-limited value v_{sat} , which is close to v_{th}

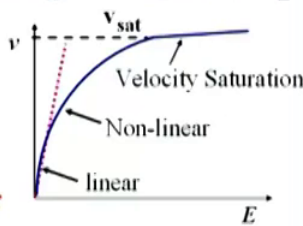
Electron velocity saturates at


$$v_{sat} = 10^7 \text{ cm / Sec}$$

Hole velocity saturates at

$$v_{sat} = 6 \times 10^6 \text{ cm / Sec}$$






4

Now if you take a look at the velocity saturation effects on the short channel MOSFET what will be the electric field? The electric field along divide direction that is from the source to the drain in that region, there is a voltage drop equal to gate voltage minus

threshold voltage that is a channel potential if you have discuss earlier. So, voltage of across the channel from between the drain end at the source end once saturation current has reach if this for $V_{GS} - V_{th}$. Now you can call or you can think of how much will be the electric field that is whatever voltage drop with their divided by channel length that average one treatment.

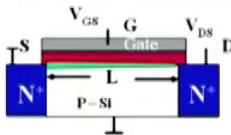
So, now, you will as the short the channel length becomes shorter and shorter electric field becomes larger and larger; that means, if the electric field is larger than about 30 kilo volts per centimeter for electrons that is velocity saturates to 10^7 . So, what your challenge is this 30 KV per centimeter can be reached very easily see for example, even if you take a number like one microon channel length, if the volt $V_{GS} - V_{th}$ is 3 volts; 3 volts divided by 10^4 that is 30 KV per centimeter. So, even there we can have saturation, we can see if the channel length is smaller even if the threshold $V_{GS} - V_{th}$ is 2 volts 1 volt, you will have the velocity saturation . So, when you go down to point 1 point 2 micro manual length definitely there we velocity saturation and that value be 10^7 per centimeter or electrons and whole velocity saturates at slightly smaller value because of it higher effective mos ok.

Now, let us see how this will change the drain current. If we are operating in this region where the velocity proportional to electric field, the Ohms law holds good and there he said it is a square law, I_{ds} is equal to $(V_{GS} - V_{th})^2$ into some constant which depends upon $\mu_n C_{oxide} w$ by L here let us see that. So, due to high electric field I just take a look at the drain current in the scaled down MOSFETs that is short channel lengths everything reduced in size.

(Refer Slide Time: 12:59)

Drain Current in scaled down MOSFETs

Due to the high value of $E(y)$ along the channel, velocity of carriers = v_{sat}



Velocity saturation model for I_{DS} :

$$I_{DS} = W Q_n(0) v_{sat} = WC_{ox} (V_{GS} - V_{Th}) v_{sat}$$

Transconductance $g_m = \frac{dI_{DS}}{dV_{GS}} = WC_{ox} v_{sat}$

I_{DS} and g_m are independent of Channel length, L

Transit time $\tau_t = \frac{L}{v_{sat}}$. **Cut off frequency** $f_T = \frac{1}{2\pi\tau_t} = \frac{1}{2\pi} \frac{v_{sat}}{L}$

NPTEL

Due to the high value of electric field in the y direction, that is along the channel length the velocity of carriers will be saturation velocity.

Now, we have assumed that electric field is very high everywhere; that means, velocity saturation will be there everywhere. So, if you take the current at the source end itself whatever current enters here will flow through the thing. So, we have written this formula the drain current is width into charge into velocity $W Q$ into v , W is the width of the channel Q is the electronic charge here, 0 is implying that x or y equal to 0 here. So, Q_n . So, $W Q$ into v_{sat} is actually the drain current, this is the velocity saturation model very simple. Now if you write substitute for the inversion charge, inversion charge is C_{oxide} into V_G minus $V_{threshold}$ that we are put here.

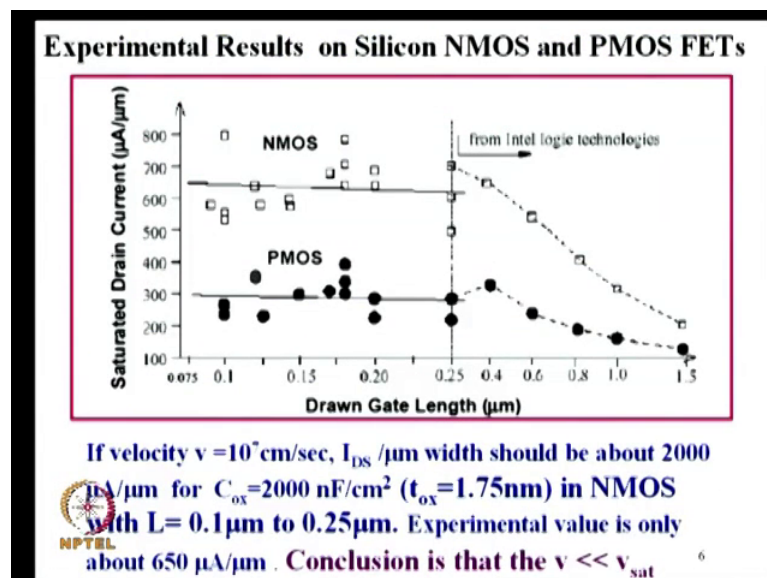
So, the drain current now you becomes WC_{oxide} into V_G minus $V_{threshold}$ into v_{sat} . That is at this point the drain current does not depend upon the channel length when the lower field case you remember that the drain current is $WC_{oxide} \mu$ divided by L into V_G minus $V_{threshold}$ square. Now here it is not square law, it is I_D increases linearly with the V_G minus $V_{threshold}$ and it is independent of the channel length because the current is the because the velocity is saturated, velocity no longer depends upon depends on different no longer dependence upon the channel length. So, it is divided.

Transconductance defined as $\Delta I_{DS} / \Delta V_{GS}$; when you differentiate that this term goes off and you will get $W C_{ox} v_{sat}$. So, you can see couple of things here to be noted your drain current depends from the channel width of course, but do not want to make this large the maximum that you can get will be decided by the saturation velocity and of course, larger t_{ox} larger will be the current; that means, thinner the oxide more will be the current. So, it will be the transconductance. In all the integrated circuits you would like to have best transconductance that means, the velocity is to be highest and you get that why do you need best transconductance because, the driving capability for a given change in gate voltage is maximum, if transconductance is maximum ok.

So, I_{DS} and g_m are independent of the channel length that is that 1 thing to note it out noted the transit time there is a time required for the electrons move from the source to the drain, that is length divided by velocity and the transit time will control how much will be a cutoff frequency that is $1 / (2 \pi \tau_t)$ is be the ω_{fd} from by 2 point τ_t . So, higher frequency transistors will have of course, the shorter channel length and because of shorter channel length, you will get faster feed of operation not because of improved velocity ok.

Now, let us see whether you get the drain current independent of channel length when you go to smaller channel length.

(Refer Slide Time: 17:21)



There is some experimental result which is available the as I reported in one of thesis in Stanford University Pgd PhD thesis, I think that this result are given by Intel; very interesting you see you keep on reducing the channel length x axis you see the channel length, y axis give us you the saturated drain current $i_{d sat}$ that is after the drain current is large enough.

So, as we keep on increase reducing the channel length there for different MOSFETs sort channel sort channel sort channel you can see the drain current keeps on increasing because you are not in the velocity saturation region the drain current is nu in this region is controlled by $\mu_n C_{oxide} w$ by l into $V_{GS} - V_{threshold}$ whole square. So, because it is inversely proportional to channel length, the drain current keeps on increasing velocity is the channel length. This is for p channel device, n channel device the bottom curve if for the p channel devices, but once you go down to channel length which are something like 0.2 microns and below, these are all the experimental point they are all scattered, but on an average you can see that is reging constant it is independent of the drain channel length which almost shows that this is correct drain current is independent of the channel length you can see that its independent of channel length.

Both N channel MOSFET and P channel MOSFET show the same characteristics. There is a difference in these two numbers because of the difference in the mobility and the difference in the saturation velocity you can see, but if you calculate this I think you have to watch very carefully, these devices have an oxide which is about 1.75 nanometers around that for these channel lengths or 1.5 inches, I have taken this number 1.75 and corresponding oxide capacitance per centimeter square is it is 2000 nano F per centimeter square, I am trying to find out if a substitute this C_{oxide} and if a substitute this v velocity saturation, do I get this number per electrons. See what you are telling is if there is velocity saturation, I can substitute that me take $V_{GS} - V_{threshold}$ is 1 volt, C_{oxide} is 2000 nano farad per centimeter square which is 2×10^{-6} farad per centimeter square and v_{sat} is 10^7 .

So, 10^{-6} into 10^7 that is 10^1 , 10^2 with there. So, this is this product on this side is 20, initially they define the current as drain current per micrometer width of the channel I have taken for they are given there, in the (Refer Time: 20:55) per micrometer, w is 1 micrometer. So, when you substitute this as 2000 nano farads, this is

10 to power 7 product is 20 , 20 into 10 to power of minus 4 on micrometers 10 to power of minus 4 centimeter. So, that 20 into 10 to power minus 4 is actually 2000 micro ampere per micrometer. 20 into 10 to power of minus 4 is 2000 into 10 to power of minus 6 , there is 2000 micro amperes or 2 milli ampere per micrometer.

So, what we are telling is if there were velocity saturation, you would have got this current has 2000 micro amperes per micrometer what we get here is less than that, which would be that there are 2 things; one is what you say is correct that channel the drain current is independent of the channel length, but the velocity is not velocity saturation there is a different velocity which is smaller than the velocity saturation. So, that is what we get from here.

Now, let us see what the model that we are taken here, what you have done is you have taken the charge at the source end and taken the velocities is velocity saturation. Now that would mean the current here is whatever the charge is present into velocity saturation, if you go down along the channel the velocity cannot increase beyond that point velocity is maximum is 10 to power 7 . So, if you say this v velocity saturation here also there is velocity saturation here also the velocity saturation what about charging? That would imply that charge also with constant, because ID if it is thousand micro ampere per centimeter square it is a same thing everywhere because current continuity.

So, what you are trying to point out is if you assume this model, which you have derived written based on velocity saturation here, that to mean that throughout the v velocity saturation that would mean that charge is throughout same, which implies there is no field in this region. If there is field in this region plus here minus here the charge will be keeping on decreasing, if charge is keeping on decreasing velocity will be keeping on increasing if the charge is not decreasing charge is constant there is no voltage round.

So, this argument that we were put forward is not quite correct. So, what we are telling is velocity saturation may not take place here because charge is maximum; velocity is less than velocity saturation. As move towards the drain end charge keeps on decreasing because there is a voltage drop here, because charge keeps on decreasing 1 of the keep on increasing you may have a velocity saturation towards drain end not at the source end.

So, what we are trying to point out is at the source end the velocity is less than the velocity saturation. So, I should calculate the current taking this same formula, but v is less than that velocity saturation, what is that velocity we do not know right now will find out that.

(Refer Slide Time: 24:39)

Velocity saturation throughout the channel ?

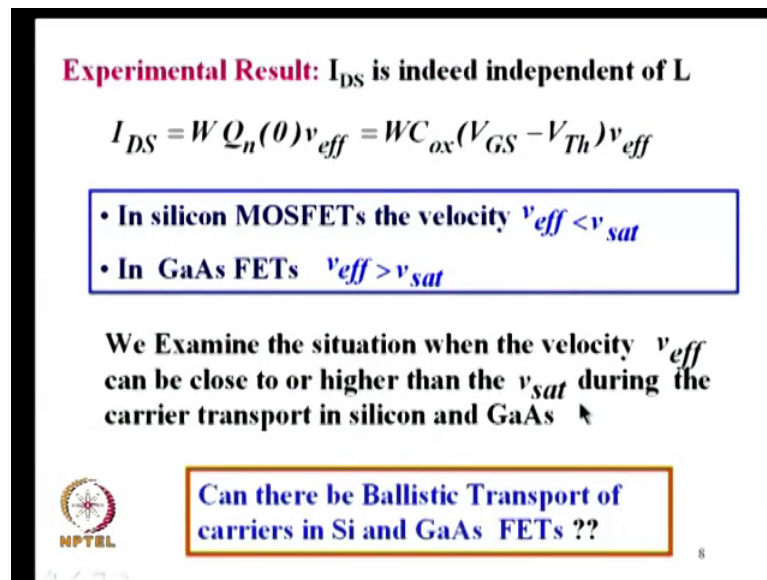
- The drain current formula in the velocity saturation model assumes that the electron velocity = v_s throughout the channel .
- This assumes that the charge Q_n does not vary throughout the channel . This is possible if there is no voltage drop or at least the voltage drop is negligibly small. This would imply that the electric field along the channel is small..
- This would lead to the conclusion that the velocity saturation is absent near the source end of the channel where the electric field is low.

The diagram shows a cross-section of a MOSFET channel. The source (S) and drain (D) are represented by blue blocks labeled N+. The gate (G) is a red block above the channel, with gate voltage V_{GS} applied. The drain voltage V_{DS} is applied across the channel. The channel length is L, and the substrate is P-Si.

• Velocity saturation can exist only in regions closer towards the drain end of the channel where the Q_n is low and the field is high .

So, whatever I have mentioned I have summed up and put it here, the drain current formula in the velocity saturation model assumed that the electron velocity is V_s throughout the channel this assume that charge does not vary throughout the channel which is not correct. So, from what we have discussed we you know the charge does not vary means, this is possible there is no voltage drop, but there is a voltage drop across the channel if there current row. So, this will be lead to the conclusion that velocity saturation is absent near the source end of the channel where the electric field is low. Lower than the sat lower than saturation that you give you or so, and velocity saturation can exist only near the drain end.

(Refer Slide Time: 25:32)




Experimental Result: I_{DS} is indeed independent of L

$$I_{DS} = W Q_n(0) v_{eff} = WC_{ox}(V_{GS} - V_{Th}) v_{eff}$$

- In silicon MOSFETs the velocity $v_{eff} < v_{sat}$
- In GaAs FETs $v_{eff} > v_{sat}$

We Examine the situation when the velocity v_{eff} can be close to or higher than the v_{sat} during the carrier transport in silicon and GaAs

Can there be Ballistic Transport of carriers in Si and GaAs FETs ??



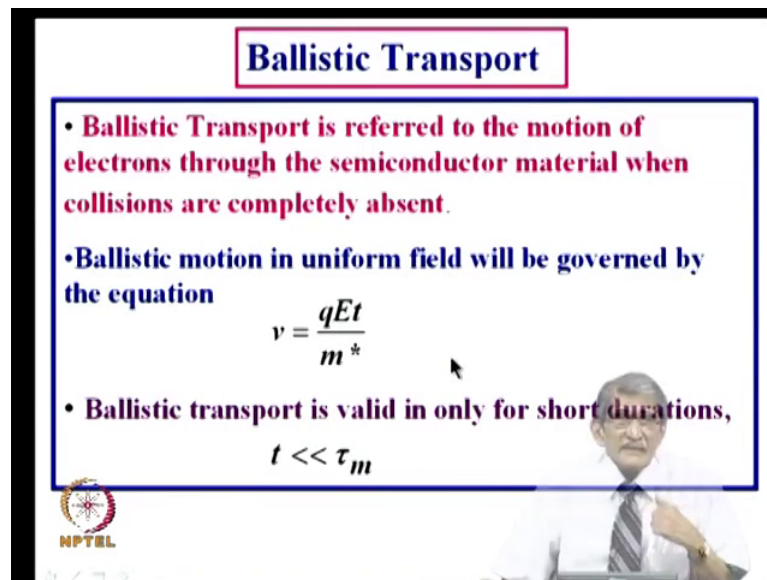
8

Now, let us examine in that quickly go through some of these things so, this is a basic thing that is involved. So, experimental results ideas is indeed independent of one we are very happy, but these some low volts good except the charge when you take a resource end the velocity is not allow saturation, but some effective velocity which is lower than saturation velocity, but it is a function of saturation velocity and some other factor comes into picture. So, WC oxide into V G S clear and I am substituting here. So, all that we replace this by v effective.

Now, in V silicon from the result that your seen this v effective less than the saturation velocity. I will not go through this at this stage, but I will just point out that if you take gallium arsenide devices that effective velocity will be even more than the saturation velocity, because of what is known as the velocity over suit effect the electrons are injected to the channel suddenly to a high field it goes through a transient phase where velocity gets much more than saturation velocity then comes back to the saturation velocity. So, that effective dominating gallium arsenide were is in silicon that is very marginal. So, it is always less than this.

Now, let us examine these there is one more term that I bring in here that the ballistic transport can there be ballistic transport of carriers in silicon in gallium arsenic question mark will not discuss that now in silicon.

(Refer Slide Time: 27:21)



Ballistic Transport

- **Ballistic Transport is referred to the motion of electrons through the semiconductor material when collisions are completely absent.**
- **Ballistic motion in uniform field will be governed by the equation**
$$v = \frac{qEt}{m^*}$$
- **Ballistic transport is valid in only for short durations,**
$$t \ll \tau_m$$

MPTEL

What is ballistic transport? Normally the electrons get accelerated they get scattered to collisions with lattice atom or impurity ions everything whatever it be waving scattering centers.

Now, the ballistic transport is referred to the motion of electrons through the semiconductor material, when collisions are completely absent. When will be the collisions absent? If the collisions centers are far away. So, ballistic motion in the; if in the uniform electric field will be governed by the equation, the velocity there is acceleration into time. So, what you are implying is if there is no collision if the electron moves without collisions the acceleration will be, forced by the mass there q into electric field divided by mass. So, for example, if the electron is moving in vacuum there is no collision that is ballistic transport, the velocity will keep on increasing there is a vacuum tubes the velocity of electrons is much more than that in semiconductors, there say people are thinking of bringing back the vacuum tube for high power high frequency operations, there is one activity going on when to bring in vacuum tubes into very high frequencies there are sequences.

So, but if the time is t if you take, the velocity at the time if acceleration into time t . Now ballistic transport will take place. So, long as the time is less than the collision time that the collision, now let us see what it is in silicon.

(Refer Slide Time: 29:22)

Ballistic Transport in Silicon . This can take place
for a duration $t \ll \tau_m = 3 \times 10^{-13}$ Sec

If the carriers reach velocity $v = 10^7$ cm / Sec then
the critical distance for pure ballistic transport in Si at
300K is $l \ll v\tau_m = 3 \times 10^{-13} \times 10^7 = 30$ nm

**∴ In Silicon MOSFET, pure ballistic
transport will not take place until the channel
length is well below 30 nm (=0.03 micron)**

NPTEL 10

This is something people have simulated from Monte Carlo simulation etcetera they have seen that, for silicon the average collision time is about 3 into 10 to power of minus 13 seconds, that is point 3 peco second very small within that time the collision takes place. So, you will have ballistic transport in their time less than that collision time, that is no collision you can use this equation to compute the that is what you actually do. So, what is that length let us see. Let us say the electronic move with the velocity of 10 to over 7 due to ballistic action, average velocity 10 to over 7 and the time t 3 into 10 to over minus 13 what is the length rabled? There is 3 into 10 over minus 13 into 10 to power of 7, there is 13 nano meter; that means, if you assume that maximum velocity is 3 to 10 to power of 7 is present for the electrons and within that collision time it would have traveled it distance distance equal to 30 nanometer; that means, there will be ballistic transport if the channel length is less than 30 nanometer, if the very much less than 30 nanometer, if the channel length is 30 nanometers and above it will govern with the scattering that is the mobility and saturation velocity ok.

So, what we are saying is there will not be any ballistic transport there is in silicon, till it is about much less than 30 nanometer say 10 nanometers 30 nanometers if you go you can say there is ballistic transport will be there. So, you may not have velocity is much greater than greater than velocity saturation, at the source and that is a meaning of that.

So, in silicon MOSFET pure ballistic transport will not take place until the channel length is well below 30 nanometers we examined that if there is a chance of velocity higher than the saturation velocity there is there is not possible. If you take gallium arsenide again the simulation results have shown that the collision time there is 5 picoseconds much larger, say if the velocity 10^7 I am taking the maximum velocity it will not be that much than the length traveled by the electron electrons before the collision has taken basis velocity in to this time, that if 500 nanometers. What we are telling is in gallium arsenide the collision will be absent or scattering will be absent till it has moved 500 nanometers 0.5 micrometers, which that would imply that there will be ballistic transport in channelings which are 0.5 microns.

So, if the carriers are injected into that gallium arsenide channel, in that short period of time before it gets scattered if it is suddenly injected into the channel where their high field is present it will get accelerated to high velocities is higher than that of 10^7 that means, in gallium arsenide you will have velocity overshoot effects like a transient effects in the IC circuit etcetera, then the overshoot effects will be there because before it comes to a halt, the velocity goes over and above the saturation velocity and that make you rise to your injection velocity much more than the saturation velocity.

So, that is why in the case of gallium arsenide devices, you may encounter drain currents which indicated that the injection velocity of electron see much more than the saturation velocity, that is because of the velocity overshoot effect. I have put it very briefly without going to details of that the meaning is sudden transient effects in gallium arsenide leads to that overshoot effect because that has at least few picoseconds of time whereas, in silicon you do not see that because it is less than much lower than picosecond, that is about 0.3 picoseconds. So, do not really see the overshoot effects.


So, do not go into the discussion on the gallium arsenide at this moment, when I discuss gallium arsenide devices I may bring this pattern. Let us focus on silicon because we are happy with silicon devices and we are talking of all devices in silicon.

(Refer Slide Time: 34:15)

Conclusions on drain saturation current $I_{DS(sat)}$

- In GaAs FETs the carrier velocity in the channel may exceed v_{sat} due to the onset of ballistic transport and velocity overshoot effects for $L < 0.5\mu\text{m}$
- In Si FET (a) Ballistic transport can not take place unless $L \ll 30\text{nm}$. (b) I_{DS} is limited by the carrier INJECTION VELOCITY v_{inj} at the source end of the channel .

$$I_{DS} = W Q_n(0) v_{inj}$$
$$v_{inj} < v_{sat}$$

 This I_{DS} is independent of Channel length, L

12

So, the conclusion from this discussion is, in gallium arsenide if it is the carry velocity in the channel may exceed saturation velocity, because the collision takes place after few picoseconds due to the onset of ballistic transport and velocity over shoot effects for small even up to 0.5 up to 0.5 micron channel length in silicon fets ballistic transport cannot take place pick even when channel length is 30 micro 30 nanometer. For 50 nanometers that is estimated by taking such velocity 10 to over 7 it will be lower than that. So, much smaller channel lengths are required to see ballistic transport or velocity over suit effects in silicon.

So, drain current is limited by not by the saturation velocity not by velocity ballistic transport not by velocity over suit effect, but the velocity with which it is injected that is something else is controlling in the injection velocity. So, at what velocity the carriers are injected you will see what it is I will not get down into the detailed discussion on that.

So, drain current if it is controlled by injection velocity which is smaller than that velocity saturation, I can say that take still write ids is $W Q_n$ at 0 into injection velocity and that is lower than that then I will get the drain current which is actually smaller than what you would get with the velocity saturation, but this injection velocity does not depend upon channel length, it would depend upon the condition sat the source end let us

see what it is I will not derive it, but I will just show that rather deriving is bit of involved thing there is a detailed paper written by Lundström and Ren.

(Refer Slide Time: 36:44)

Ballistic Transport and Injection Velocity

Considering the effects of scattering near the source end it has been shown that

$$\frac{1}{v_{inj}} = \frac{1}{v_{th}} + \frac{1}{\mu_n E(0^+)}$$

$v_{inj} = v_T = 10^7 \text{ cm/sec} = \text{ballistic transport}$

Low field mobility should be high to achieve High Performance MOSFETs

Reference: Lundstrom & Ren, "Essential Physics of carrier transport in Nano scale MOSFETs" IEEE TED-Vol-49 pp.133-141, January 2002

13

On the physics of carrier transport in nanoscale MOSFETs published in January 2002 in which what they have said is, what the carriers are injected from the source to the channel. They are able to inject from the source to the channel by gaining an energy to raise above that barrier ok.

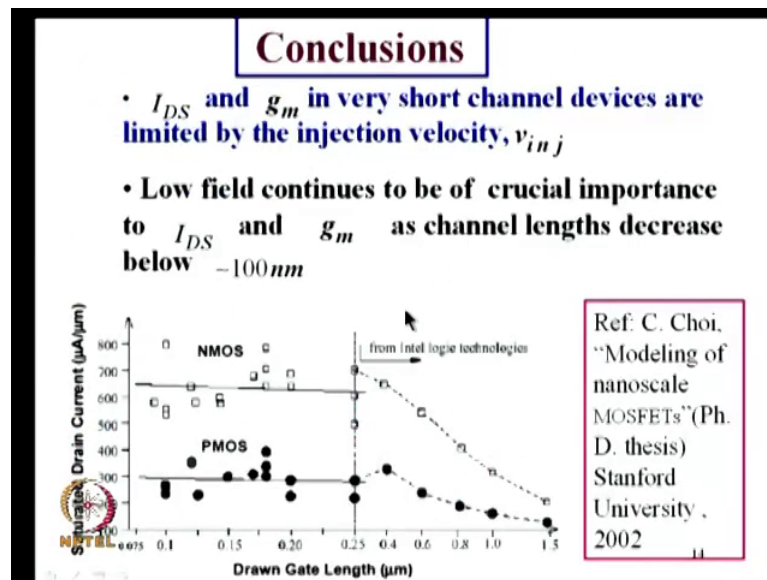
So, they have got the sub thermal energy. In fact, they would be injected at a velocity close to 10^7 there. So, you would think that the earlier analysis what we said is injection velocity v_{inj} thermal velocity, that is what you get there at the point of injection at the point when electrons have just landed in a channel, but the moment they land in a channel they start moving within a very short length they suffer collisions, they suffer collision they start scatter backward. So, the average velocity of that beyond that point within that short distance is controlled by the scattering mechanism, I am not showing the derivation. So, that is controlled. So, if you go through the analysis what happens is, if there is no scattering at all you do get the thermal velocity. Because of the scattering the velocity of the electrons is controlled at that point by combination of this injection velocity thermal velocity, and the velocity with which it moves that is mobility into the electric field penalties is mobility to the electric field that the low field the

electric field is not high at the source end, electric field high only at the drain end if it is high of course, saturation velocity you will get its not high.

So, whatever electric field at E plus 0 plus means just at the source end that is a velocity that you would get. So, a combined effect of that with skipping all the derivation, I would suggest that you go through this favor for that because that itself would have taken me about half an hour to discuss that. So, injection velocity is controlled by these equations that is thermal velocity which is 10^7 centimeter per second, if there is no collision at all after injection that would have been 10^7 centimeter per second. But because of the combined effect of the scattering you get injection velocity less than the thermal velocity. You can see that if the mobility very very large it could be that in that electric field, you will have it will like putting 2 resistance in parallel the resistance is govern by the resistance which is lower ok.

So, here if this is larger than this one if the μ into e larger than thermal velocity v injection with thermal velocity if thermal velocity is larger than the second term here denominator injection velocity govern by this mobility it will actually combined effect it will leather this nor that. So, what turns out from that is this for the conclusion for this or discussion is not ending here, considering these both these effects together what do you say that you can still right the drain current is equal to WQ into injection velocity, injection velocity is the combined effect of these two terms. If I have high mobility in the channel, you may get thermal velocity itself very close to saturation velocity, but it will only will less than that. But if the mobility low then it will far below the thermal velocity saturation velocity. So, you can see that.

(Refer Slide Time: 41:13)

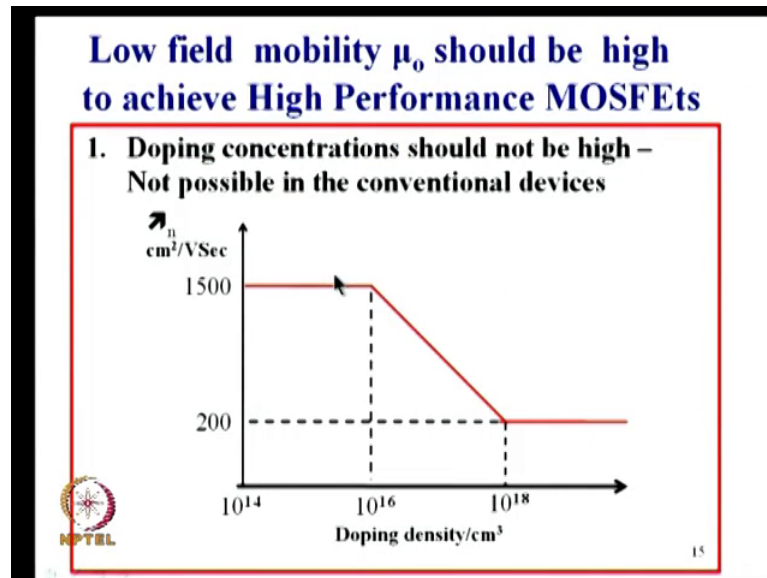


It is particular term drain current that you gets saturation drain current will be in the very short channel devices are limited by the injection velocity, which is controlled by both saturation velocity or thermal velocity and the mobility and you can see that both the terms which determine the injection velocity are independent of the channel length.

So, injection velocities does not depend upon channel length therefore, if you are computing this current using the injection velocity, the drain current will be independent of the channel length. So, what we say here wholes good. So, you get the drain current which is independent of channel length then you go to channel length which are 0.2 0.1 microns length or below, but the current is lower than what you would get with the saturation velocity, because injection velocity smaller than the saturation velocity and this is actually you can see the references is this is a research scholar, modeling of Nanoscale MOSFETs PhD thesis Stanford university 2002. There is a time when that model also was given by lunchroom etcetera in the Ford University.

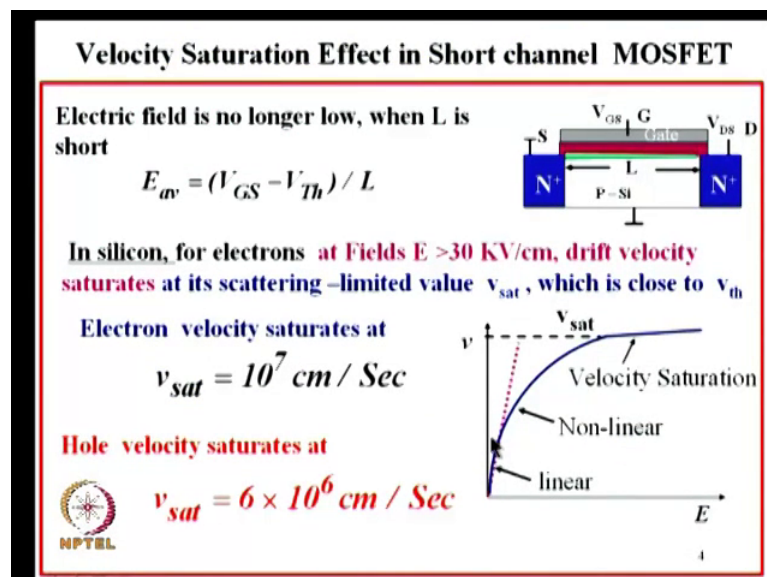
So, what is our end this from this discussion what we see is, you need to have very high mobility for electrons because after all thermal velocity is fixed by temperature 10 to power 7, you need to get higher and higher mobility what the way you can get higher mobility what are the factors which affect the mobility that will see now.

(Refer Slide Time: 43:09)



So, low mobility from field mobility should be high, see μ is actually the mobility is the low electric field you go back to the μ_0 mobility is actually and just the slope here.

(Refer Slide Time: 43:23)



Low field region and the source end the field is low as you go towards the drain and the field becomes high, the velocity naturally go up there. So, now, going back to this I am skipping all these things which you are discuss now. So, you are looking at either material switch you are give high mobility or if you want to stick with silicon, what is the way you can increase the mobility what is the what are the factors which increase the

mobility or limit the mobility. If you plot mobility μ_n on this is actually μ it has (Refer Time: 44:07) this is actually the arrow is μ , I think because I change in the program, this is the mobility is the y axis and the doping concentration. You if recall to reduce the short channel effects you always increased mobility in the conventional MOSFET.

So, if we increase the mobility, the I am sorry if you increase the doping I am sorry let me restate, the conventional MOSFETs you increase the doping to reduce the short channel effects and also reduce the oxide thickness to compensate for that. So, if you increase the doping the mobility is almost flat I have plot at the this is actually approximate a straight line it will not be the state will be smoothening here smoothening here. So, when the doping is sub to about 10^{16} to over 10^{18} mobility of electrons in Silicon 1500 centimeter squared per volt second. But once you reach about 10^{17} to 10^{18} doping, the mobility starts following mobility starts following because in the lower doping concentration regions the scattering center are only lattice atoms, but when you go to higher doping the dopants they in there in the ionized state, the ionized dopants you rise to additional scattering centers. The low doping mobility they are not the there is scattering, but it is not high compared to the lattice scattering,, but when you go to higher doping concentrations additional lattice scattering due to the ionized dopants takes place mobility falls.

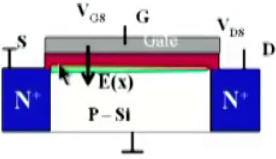
You can see the assistant of fall in mobility what are 1500 centimeter square per volt second and become even as low as 200. So, you can see if you look at take a look at though doping concentrations this term this is this is not 1500 less be 200 when you go to 10^{17} to 10^{18} . In 10^{16} to 10^{17} it may be about 600 or 700. So, moral of the stories you cannot go to such to high doping concentrations in the classical MOSFET you have to have high doping concentrations to reduce that short channel effects, what are the other factors coming into picture? The other factor which leads to reduction in the low field and they say that field, that field is actually along the y direction, but in the case of MOSFET there is electrical field in vertical direction there is normal direction that is what we talked here was the I am sorry electric field in that direction then he what we talked up here.

(Refer Slide Time: 47:05)

Low field mobility μ_0 should be high to achieve High Performance MOSFETs

2. Electric Field $E(x)$ normal to the channel must be kept low – In conventional scaled devices this is invariably high due to high doping concentration and ultra thin gate oxide

An Approximate Expression for the mobility versus Normal electric field E_{eff} is as follows

$$\mu_{eff} = 3.25 \times 10^4 E_{eff}^{-0.33}$$


Baccarani and Wordeman, "Transconductance degradation in thin oxide MOSFETs" IEEE Transactions on Electron Devices, ED-30, page 1295, (1983)

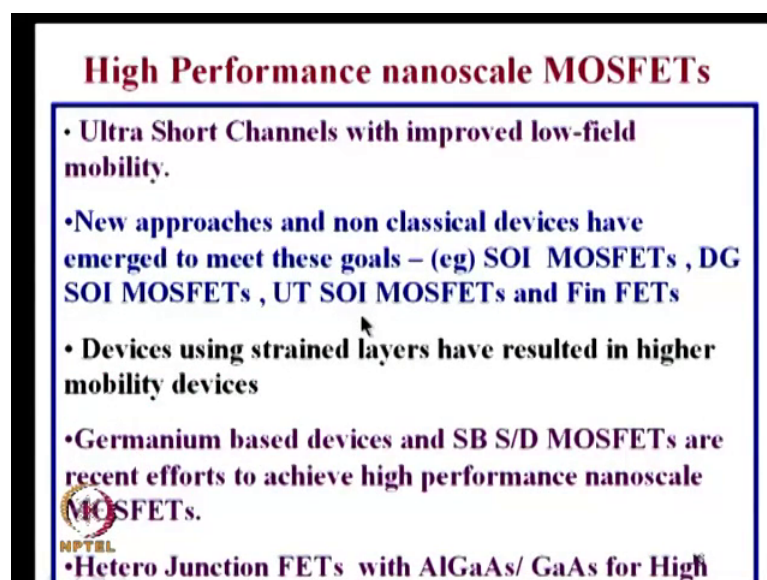
Now, what we talked of here is there is also electric field in the vertical direction that normal direction to the channel. So, electric field in the normal direction normal to the channel must be kept also low, but you know when you make the oxide thickness small and the channel doping high, the when you deplete the channel there would be very high electric field. Oxide thickness is small the depletion layer charge is high. So, $q d$ by c oxide is the electric field or whatever charge is there that give is a its a electric field, $q d$ by ϵ_{oxide} that it is a electric field. Q_d by oxide voltage drop across the thing though if the c oxide is large and $q d$ is large or if $q d$ is large electric field also become large.

So, what you are trying to point out is you will have because of the high dope here and thin oxide field in the large in the vertical direction. So, in the electric field if large in the vertical direction the mobility of this channels along the y direction gets affected why you can imagine, but electric field in the direction vertical direction like this normal direction to the channel if the electrons are moving in the direction, if the field is like that it try to keep it hold it for a while in that region. So, there is a restraining force which keep back in that position that is a quality way of understanding; that means, actually the mobility of the electrons moving in the direction gets affected if the field is large in the direction. So, you must avoid the field met large in the x direction normal direction.

Now, an empirical formula is given by these people Baccarani and Wordeman in the year very way back in 1983, why the conventional MOSFET mobility keeps on decreasing mobility in that direction keeps on decreasing, if the E_x we call it $E_{\text{effective}}$ to including all other effect, if that increases that keeps on decreasing. So, what do a way you can do that, you cannot help reducing oxide thickness you want to keep the capacitance large what you can do will be reduce the doping concentration here. So, if you reduce the doping concentration the channel 2 things happen, one the ionized impurity scattering its reduced though mobility reduction does not happen you can go to mobilities which are as high as 10500 in silicon. If you use some other material you can higher mobility is decided by the mobility of that electrons that material and also if the reduce the doping concentration here, the depletion layer charge gets reduced there for the charges in the depletion regards reduce means electric field again is lower in this vertical direction normal direction.

So, if you reduce the doping concentration naturally it get improved more low field mobility low field in the direction, is get improved mobility and you will get improved velocity you can get very close to the saturation velocity thermal velocity. So, the entire drive in the present day devices is to choose these are the reference which I am telling.

(Refer Slide Time: 51:10)



High Performance nanoscale MOSFETs

- Ultra Short Channels with improved low-field mobility.
- New approaches and non classical devices have emerged to meet these goals – (eg) SOI MOSFETs , DG SOI MOSFETs , UT SOI MOSFETs and Fin FETs
- Devices using strained layers have resulted in higher mobility devices
- Germanium based devices and SB S/D MOSFETs are recent efforts to achieve high performance nanoscale MOSFETs.
- Hetero Junction FETs with AlGaAs/ GaAs for High

NPTEL

So, to get high performance what should you do? We just last time which I am discuss here today which has per lot of juice in that high performance nanoscale devices

therefore, should ensure the doping is low. So, that mobility is high they should also ensure that the vertical normal electric field is low. So, that again the mobilities low field mobility is high ok.

So, ultrashort channel mobility devices with improved low field mobility is required for high performance nanoscale MOSFETs. You cannot help reducing the channel length why should reduce the channel length? You saw channel length divided by velocity using the transistor higher frequency; to go to higher frequency even though velocity gets saturated because of reduce channel length; the time required for transport of carries is reduced to frequency is high. So, to take care of this aspect ultimately it improve the mobility, new approaches are used is new approaches are non classical MOSFETs finally, we are getting down to what we wanted to discuss.

So, they have emerged these meet some of these things meet this goal of low mobility, those devices are silicon on insulator MOSFET insulation MOSFET. If you recall the MOSFET that he have got is whole thing is on a bulk region and out of that we discussed started the meaning only the top layer is the one which is being used you know that it is actually mechanical support. So, you can have silicon on insulating substrate a thin layer of silicon used we will see how that works out sorrows to improve the low field mobility. So, because of that you have got double gate MOSFET, ultrathin SOI MOSFET silicon on insulator MOSFET and Fin FET, these are the things which we discussed in the coming next few lectures.

So, that is actually using silicon itself and this trick that we are using these devices is only the controlling g the electric fields controlling the dopants which any which is enable by the SOI device. But now you can make use of the SOI device in addition you introduce strain around the channel what is meant meaning by of strain? If have the channel let me go back to one of those slides here these are channel. Now allowing this channel if I introduce a compressive strain in that direction introduce compress the channel in the direction, what happens is the lattice atoms are little bit broad closed to close to across together.

So, the lattice atoms are bit closer together, the collision distance becomes reduced. If the collision distance becomes reduced collision time is reduced mobility will be reduced if a bring in comparison the mobility will reduced. But if I stretch it if I introduce tensile

stress along the channel then the lattice atoms are slightly pulled it may be delta it may be fraction of an angstrom like that after all distance is about 5 angstroms may be its fraction of that. If you increase the length the spacing within the lattice atoms is increased collision time is increased mobility is increased.

So, in n channel MOSFETs if I have tensile stress strain, if I can introduce then the mobility electron mobility will increase. So, whatever is good for electrons is bad for holes, later on will discuss it again discuss though detail in the holes if I stretch it even though electron mobility increases when you talk of the plus charges exactly of the things take a takes place, I will not discuss that at this time frame. In fact, the whole moment takes place because of the jumping movement of electrons from bond to bond through the bond has to break and jump to the near being neighboring point.

So, whole movement takes place by jumping movement of electrons from one bond to the another bond. So, if a stretch it the electron has to jump on a longer path; that means, the whole moment will be more difficult if a stretch it. So, what we said is the electron mobility increases if a use tense address, but because of that the whole movement becomes difficult whole mobility becomes less if I stretch it. How do you introduce the tensile stress? I can deposit nitride on top of the gate that nitride as for tensile stress it is stretch it along it. So, Intel has made some of the devices where they deposited low pressure nitride on to the on the key or the gate to stretch the channel so that ultra mobility can be increased. And the whole mobility can be increased by compressing the channel how do you compress the channel take a look at this. By introducing nitride on top of the gate I can stretch it I can compress it if I introduce germanium into the source and drain region.

Source is silicon during diffusion you can add little bit germanium also. So, this will be a silicon germanium small amount of germanium concentration, germanium lattice or the atom atomic size is bigger than that of silicon. So, in these regions you have silicon and germanium atom at source and they drain end if you introduce at volume increases; because radius of germanium atom is bigger for it is in the given volume you are trying to introduce top ends which are bigger. So, its tries to compress, it is I am sitting here to fat fellows sit down more both my sites will be compressing me. So, that is what is happening here germanium atom graduate on to source that be compressing the atom.

So, if you do it for n channel devices its mobility decrease or p channel devices the mobility will increase. So, in the p channel device of the had germanium to the source drain to increase a mobility. So, that is a strained layer. So, other thing is occurs totally use different materials like germanium, leave silicon and go to germanium germanium has electron mobility which is about 4000 centimeter square per volt second, which is about 3 times close to 2 to 3 times than that of electron mobility in silicon. Whole mobility of electron whole mobility germanium also it about 3000 centimeter square per volt second, oh both electron mobility and whole mobility were high.

So, today people are talking of making in devices on the germanium MOSFETs and also in addition you can use gallium arsenide as the material for (Refer Time: 59:14) devices to. So, to sum up what is being done is either go to silicon on in slater type of materials there is non-classical MOSFET or and used strain layers or go to totally new materials like germanium or gallium arsenide, to realize he realized conventional type of MOSFET, but new materials. So, with that I conclude and next session we start with the non classical MOSFET on silicon on insulator.

Thank you.