# Fundamentals of Semiconductor Devices

Prof. Digbijoy N. Nath
Centre for Nano Science and Engineering
Indian Institute of Science, Bangalore

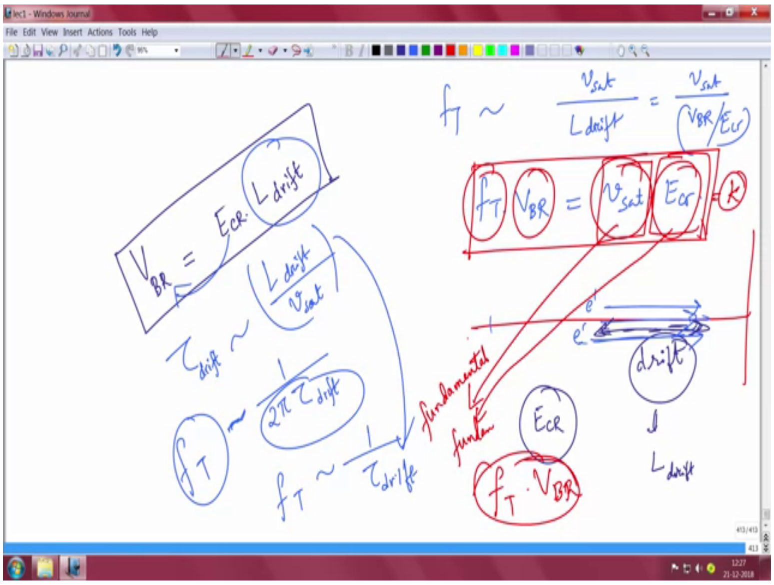
## Lecture – 56
## Transistors for RF (contd.) and transistors for Memory

Welcome back. So, we have been discussing about RF Power Transistor, introduced the basic concepts there. I told you in today's lecture we shall discuss a very important figure of merit for RF transistor its called Johnson Figure of Merit. And we will see that, you know a how are the materials and devices with a better junction's figure of merit and how it is derived.

And from there we will conclude RF transistors and we will move to things like memory or you know logic devices where we use transistors. Remember, we are discussing about the applications of transistor in some broad areas: primarily power switching, RF power amplification, memory and logic. There are many other applications also, but these are the main power applications that you are talking about here.

So, let us continue with the discussion on the RF transistor, and we will come to the white board to discuss about Johnson figure of merit.

(Refer Slide Time: 01:17)



Johnson figure of merit and Johnson figure of merit is very important for RF devices, it tells you how good an RF device is.

So, if you think of any transistor, you thinking of any transistor say this is gate and this is drain by the way you know sorry; if this is gate and this is drain. This region, this region actually in the channel this is the channel probably, this is called the drift region and your voltage is typically blocked in this drift region only. As you deplete the gate the depletion extends sideways towards the drain and so that the blocks the voltage.

This is the drift region that we talk about the electron has to drift through this region. The material has a mental breakdown field that is E critical, that is a fundamental property we cannot do many thing about it. If this is the maximum field that we can sustain and this is the drift region that the thickness of, the drift region here is call that is a you know L drift. This is the drift region that it has to go 1 micron 10 micron whatever you decide that, ok. Then the breakdown voltage that you are that you are holding between this the V breakdown; the breakdown voltage that you can actually have is given by the product of critical electric field times the drift region. This is your breakdown voltage, fine.

Now, in this drift region the electrons have to transit, the electron will move across the drift region right; through the channel and come after the drain. So, there is a time taken where electron to drift. What is the time taken by the electron to drift? The time the electron has to move from the gate to drain there is the source here, but electron while injected from the source it has to move from in the gate drain region, which is the most crucial region for blocking voltage. There is the time that is there and this is typically given by the distance the drift distance divided by the velocity which they move and the maximum velocity which it can move is the velocity of saturation, simple. So, this is the minimum time it will take to move there ok.

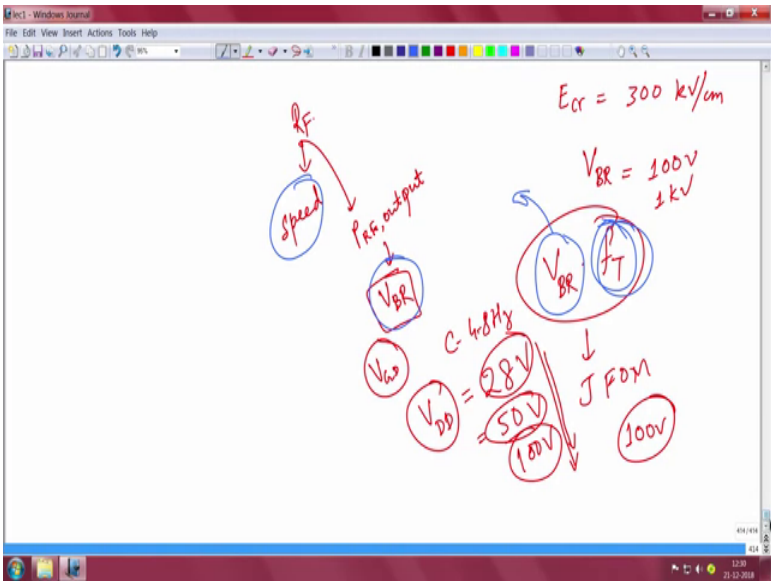
Now, you know the current gain cutoff frequency which is f T the speed; the current gain cutoff frequency which is speed is actually given by this is the speed the first is actually given by 1 by 2 pi times the time it takes to drift, you can say. How much time the electron takes to drift, is inversely proportional to speed. If it takes a longer time the speed is slow, if it takes a smaller time the speed is fast. So, f T is actually directly proportional to 1 by you can say that time taken to drift ok.

Now, I will come again here. So, if I from here if I put that then f T if I use this expression here then f T is actually v sat by L drift right. And I can put L drift from here as V BR by E CR; so I can say this is v sat divided by V BR right into by the breakdown field. So, what I can do that is that sorry, what I can do that is I can write it better I can say f T into V BR is equal to v sat into E critical.

Now, this is actually the Johnson figure of merit. So, the frequency at which you can operate times the breakdowns voltage is equal to self saturation velocity of the electrons times the critical electric field. Now the saturation like saturation velocity of electron is a fundamental material l property, you cannot change that right saturation silicon will be constant. Similarly fundamentally the critical breakdown field also is the fundamental property you cannot change its constant.

So, in other words the product of the saturation velocity and the critical electric field is a constant is a constant which you cannot change. So, it means that your product of frequency at which you operate the device, I mean not operate the maximum frequency that, we can get in a device time is the breakdown voltage; the maximum voltage at which you can operate the device that product is going to be constant and its limited by the product of saturation velocity and the electric field and this is called the Johnson figure of merit, right.

(Refer Slide Time: 05:49)



You see the electric field is fundamental property. So for example, in silicon it is around 300 kilo volt per centimeter, but the breakdown voltage is not fundamental. If you use a thicker silicon you will block more you can block 100 volt, you can block 1 kilo volt wherever few kilo volts, it depends on how thick you want to make the silicon, right this is not a fundamental property.

So, the maximum voltage that we can get the breakdown voltage we can get times the maximum speed you can get that product is constant ok. Which means this is called Johnson figure of merit – JFOM; and a higher Johnson figure of merit means that the transistor is better as an RF device. Because in RF device why are it is important and RF device you want the speed, of course the RF device has to work at speed you want to work at 4 gigahertz, 8 gigahertz, 12 gigahertz whatever.

So you need speed, and also you need the higher output power a higher RF output power. If we do not get higher power then there is no point, you need to get higher works in the RF. And to get higher output power you need to actually have a higher breakdown voltage in the device.

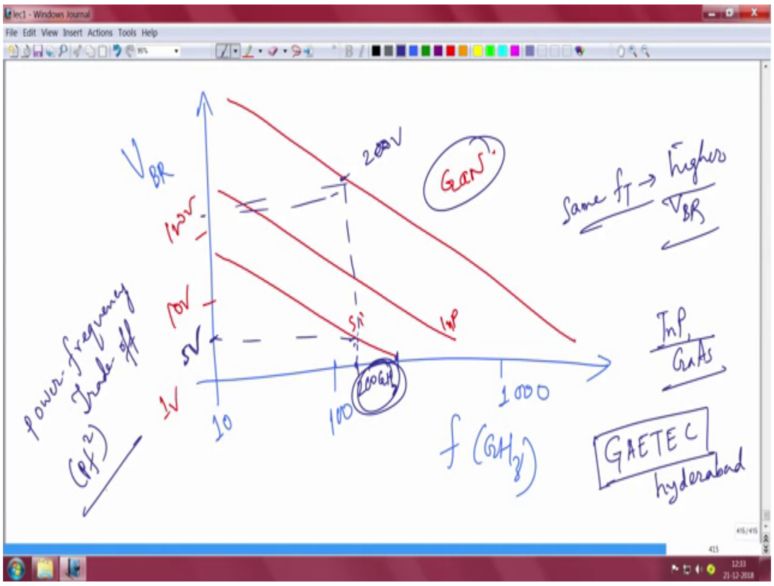
A higher breakdown voltage in the device means that your get drain voltage can be high and so you will be able to block larger voltage. It is not like a power switching device, but nevertheless you see when you operate the power transistor as a you know in a C band for example, in 4 to 8 gigahertz or whatever 8 to 12 gigahertz anything; you have something called V DD that is your drain bias or the supply bias that supply bias will decide how much output power you have.

Typically, the supply bias could be 28 volt, then you will get some power. You can also put the supply via says 50 volt or 100 volt; in case of 50 volt supply power your output RF power will be also more than in the case of 28 and so on; and 100 volt will be it will be even more. So, you want that you want to bias the device at a higher voltage; bias [FL] the supply the V DD should be at a higher voltage so that you can get better RF output power.

But, to apply a higher voltage you also breakdown needs to be high. For example, if your breakdown of the device is 100 volt you cannot apply the device an 100 volt you have to apply that less like 50 volt or whatever, right. So, you need to have a higher breakdown voltage and also you need to have a higher speed. So, those are two critical parameters just trade off with each other. A higher breakdown voltage will lead to a smaller speed and a higher speed will lead to a lower breakdown voltage. So, you have to trade off between the two. And that is a very important thing that you know people tend to forget, but it is very important actually.

So, I will show you a picture here, I will draw a picture here. So, what will happen actually is that I will change the slide.

(Refer Slide Time: 08:38)



And what you can do is that you can plot the breakdown voltage on the y axis, on the x axis I can plot the frequency. So, suppose this is 10 gigahertz this is in gigahertz and suppose this is 100 gigahertz this is 1000 gigahertz 1 terahertz there is no transistor that will probably give a gain at 1 terahertz almost; very very you know very few indium phosphide is too close to 1 terahertz.

So, you know you will have silicon that will probably go like this right then there is indium phosphide, this is indium phosphide, this is silicon then there is GaN ok. This is breakdown voltage maybe 1 volt 10 volt whatever 1000 volt and so on ok. So, I am not drawing accurately, but if you draw accurately you will see that you know; if you want to get you know if you want to get frequency response of say you know 200 volt 200 gigahertz cutoff. If you want to get a 200 gigahertz cutoff this is not 200 gigahertz sorry this will probably be around 6-700 gigahertz, in this case.

Suppose you want to have a 200 gigahertz transistor, you want to operate 200 gigahertz you know the maximum operation frequency is 200 gigahertz then only you will get a few volt; maybe you know you will get some 5 volt or so. That means, at 200 gigahertz if you want to get a cut off a 200 gigahertz, your silicon transistor cannot have a breakdown more than 5 volt which means your device has to be bias that like 2 volt or 1 volt and that will be your output power will be very very low.

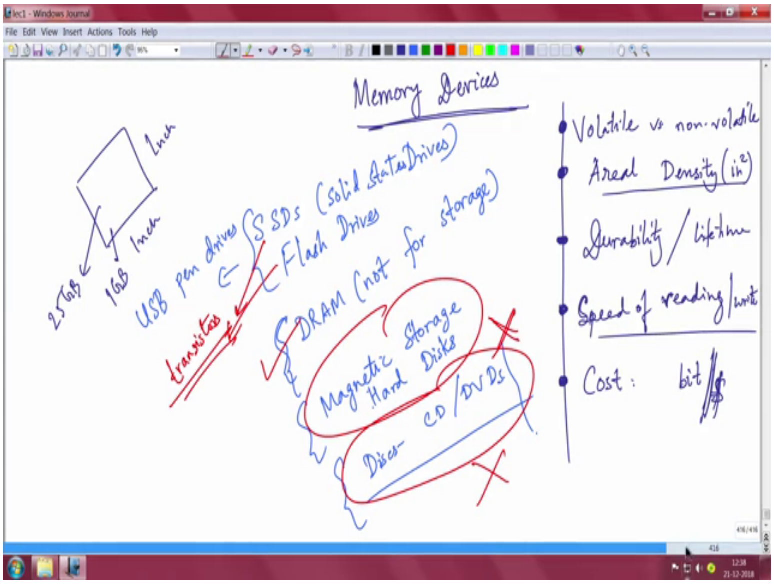
In the same time in gallium nitride you can actually bias the device that much higher because at given 200 this thing you will get probably like 100 volt or so ok; maybe 200 volt or 200 volt or something. That means, at 200 gigahertz cutoff frequency you can still get a gallium nitride device that can be close to 100 volt, and you might be able to bias it higher. So, we will get the higher output power.

So, you know these devices are actually better than silicon in terms of Johnson figure of merit in that for a same the frequency you will get a higher breakdown voltage; a higher breakdown voltage ok. So, dominant technologies I told you indium phosphide, gallium arsenide they are used very widely in; they may not be able to give the best of power, but they have operated very high frequency.

So, in India also you know if you want to look up for example, there is this company government actually company it is called GAETEC its situated in Hyderabad. You can Google up today GAETEC Hyderabad, actually is Gallium Arsenide Enabling Technology something like that. It basically makes this gallium arsenide mmics or this monolytical integrate circuits. All the transistors, gallium, arsenide, even they work on gallium nitrides; you will see that they make this device for strategic and other space applications. These are indigenous gallium arsenide gallium nitride devices that are being made and these are useful for RF applications ok. So, that is very important for RF applications.

So, now, you know that you know there is a trade off in the RF devices in terms of breakdown and not on resistance, but the frequency. And this is also called the Johnson figure of merit is also called power frequency trade off; power frequency trade off or pf square trade off ok. So, that brings us to an end to the discussion on RF transistors.

(Refer Slide Time: 11:56)



In the next we will move towards is memory; memory devices. So, what kind of memory devices used transistors and semiconductors? I am sure you probably know of many of the memory devices definitely, but you know many of the details might missing. Some of the important parameters of a memory device; what are the what are the target criteria in which you distinguish the memory devices or you classify or you rank the memory devices in terms of that performance and other things.

One is called volatile versus nonvolatile memory. You know what is a volatile and what is nonvolatile memory; whether we can erase the content versus whether it can you know if it will eventually you know; whether it is a memory the information will be there or with the where the memory information that you stored here will eventually go away ou know. I will come to that again later; one of the thing is called Areal Density.

Areal density actually some information how many kilobytes or megabytes or gigabytes of data can you store per in square of physical space. Like you know I have a memory that is 1 inch by 1 inch something, pen drive, hard drive whatever. How much memory can I store? Can I stored 25 GB or maybe I can store 1 GB only or so on and so forth ok. That is called areal density you want a very high areal density device. Do you know what is the maximum areal density, today we will come to that.

You want a very high real density, so that it very small location very small size, you are able to store very large amount of information that is called areal density. One of the important parameters of course, is durability and life time; how long the device will work what is the durability you know how robust it is in a way.

Another thing very important is speed of reading and writing. You have to read from the memory, you have to write to the memory; right when you store when you store information on the memory it is writing know and when you retrieve it is called reading. With what speed can you do it, it cannot be like you know you wait for many hours to write that the just they copy a single file that does not happened; the speed is very important. And of course, everything else remaining same what remains important is cost which is what is the price per unit Dollar, how many bits can you buy for how many per unit Dollar; per unit Dollar how many bits are you storing ok. These are very important things.

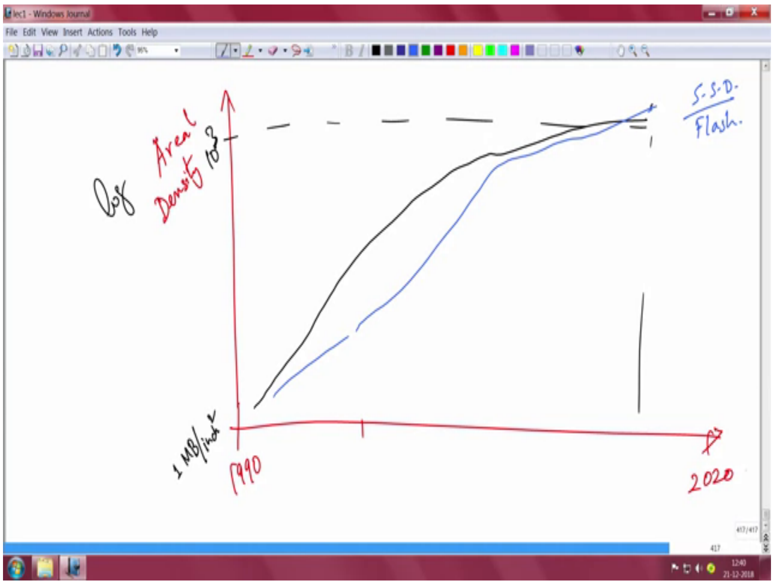
So, you know some of the common memory devices are of course, SSDs which are called Solid State Drives which are very upcoming and very solid state drives which are very widely used now, and also we called them flash drives the same thing ok. We call them flash drives flash drives are something like your USB thumb drives, the pen drives that you have the pen drives that you have or the you know memory cards in the in your camera and other things that you have.

And of course, the second one is the there are many storage device like DRAM and DRAM is not for storage by the way this is for computer booting up where you store cash and all files right. DRAM is not the storage per se, but it is for computer booting up and all this dynamic access random access memory you call them. Then there are this magnetic memories magnetic storage devices or magnetic memory devices which are like the hard disk. The additional hard disc that you have inside computer that that plugged you know, you know the additional hard disc that are inside the computer many of the you know previous generation hard discs that you can connect which I plugged supply right.

Then other things are discs; discs are like CDs and DVDs which have become extinct already, you know although we have seen their rise, but they have also become extent. Out of these in they might be other one (Refer Time: 15:46) and all. How many of them use transistor do you know? This do not used transistors, this do not use transistors, this also do not used transistors, magnetic storage devices used magnetic materials and other fundamentals it does not use transistors.

Transistors is a used for DRAM for example, although it is not for storage ok, it is a memory device not for storage exactly, it is for computer booting have and all and solid state drives, flash drives they also used transistors. Exactly is not the transistor, but it is a transistor structure you will see that and why it is call a flash memory right. I told you we talking about the different kinds of you know the technologies that are using the transistors; we do not have to go into the basics of magnetic storage and stuff like that right.

(Refer Slide Time: 16:24)



.

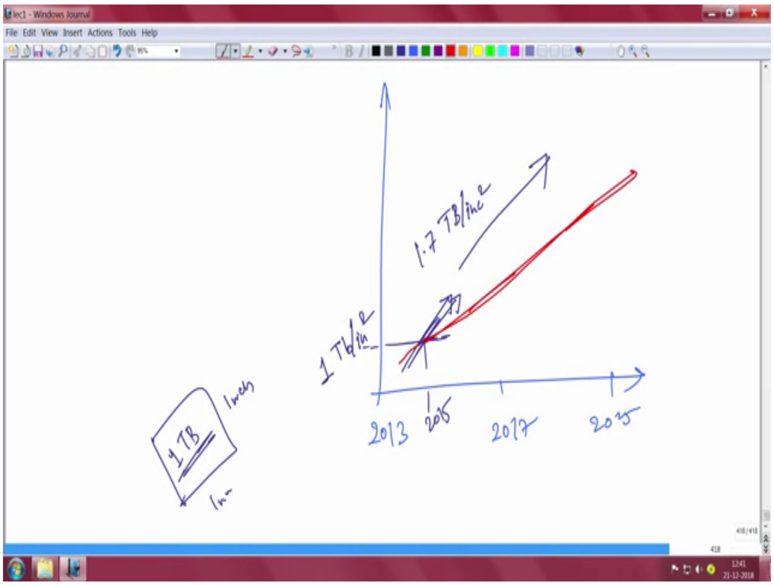
But you know there is something called you can plug the areal density of I told you areal density is the number of bits or bytes per you know square in you can store, this is the areal and density and this is your year right so, from 1990 all the way to 2020 right; so all the way till 2020.

So, your magnetic you know hard drive and your solid-state drives have gone up in steadily over these years increasing in their areal density. So, one of the curves is going like this for example, right one of the curves is going to this is your hard drive for example. So, the hard drive has gone from like you know this is at 1990 it was close to one maybe megabyte 1 megabyte per inch square or less than that probably at the area. And now eventually it has this is a log scale by the way this is a log scale it has gone up crazy high.

So, now you are talking about maybe you know few 100 gigabytes per square inch you know few 100 giga 10 to the power 3, 10 to the power 3 gigabytes per square inch it is like almost you know 1000 gigabytes or 1 terabytes per square inch. And then you have of course, solid state drives that have decreased that was going at very slow phase, but suddenly just caught up and now it is almost over taking hard drive.

Solid state drives of flash drives or flash drives and now starting to overtake the hard solid the magnetic hard drive storage in terms of areal density as to how many you know bytes of bits you can store per square inch there is a crazy number right.

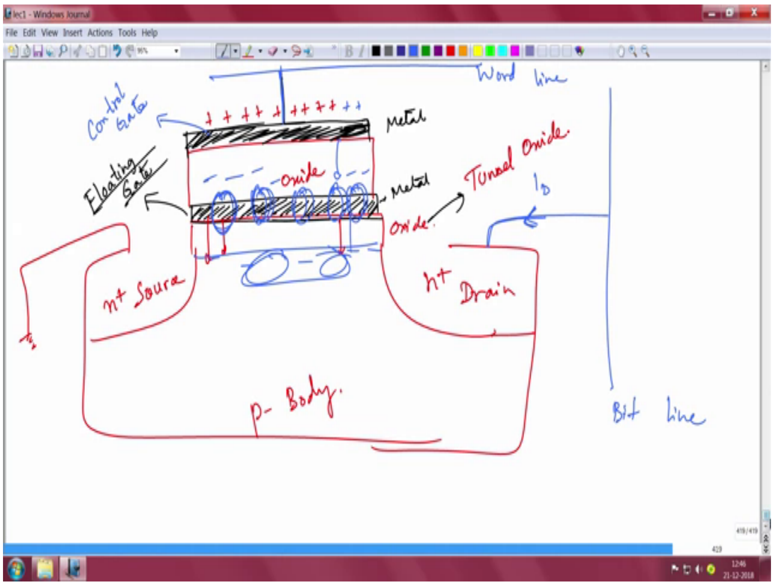
(Refer Slide Time: 18:11)



You might on to zoom it and see, if you want to zoom it and see then you know in 2013 and this is 2000 you know 25 for example, we will have data till around 2017, you know your magnetic hard drive is doing like this right. But a solid state drive your solid state drive is going crazy already it is exceeding the magnetic drive and it is supposed to go towards even higher ok. And what is the, this is the areal density? So, this point for example, this point where they are crossing is around 1 terabyte per inch square; that means, in an 1 square inch by 1 inch you know area you are able to store 1 terabyte of data.

So, the your this is around 2014-2015 data couple of years back. So, your solid state drives are now actually able to store more than your magnetic hard drives in terms of areal density, it has crossed up to I think 1 point almost 7 terabyte per inch square of data it can store right well magnetic memory devices will go grow, but it will not grow as fast as your solid state hard drives and memory devices right. So, now this is about the basic idea that you know this is the thing.

(Refer Slide Time: 19:28)



But how does this devices work. As I told you have you know flash memory devices and then you have DRAM devices that are used in that are that are made from semiconductor I know your transistor technologies for example. So, flash is used in as I told you like pen drives and USB drives and so on and DRAM is the main memory in PCs and work stations helps in computer booting and other things you have the read write sort of form right.

So, I will come to the working of the flash which is more important in an more apprising in these days. So, flash is a very dominant technology and its made of externally a silicon and MOSFET sort of a transistor, but its slightly different architecture. Whereas, the slightly different architecture what I mean is that you actually have you will see that the structure is very similar, you actually have an n plus source and you actually have an n plus drain which is just like your silicon MOSFET that you have studied a MOS capacitor that you have studied..

And it is a p type body which is very similar right and then of course, you are making a contacted here to the ground over here I am talking about the flash how a flash works actually ok. The simple architecture how a flash work that, you are not biasing the drain you are actually pumping drain current here because there is a line that is called the bit line ok. That is the bit line that will read and write the data that is the bit line that will pump some current that will get some current out of this drain.

Now, you have your oxide here you have an oxide here right and then you have a metal gate you have a metal gate here. Now you might said this is just like the silicon MOS capacitor MOS ferrite, I will come to the difference. It is your metal gate here, but the metal gate actually is floating which is very strange right. On top of that there is also an another oxide which is strange and then on top of that you have another metal gate right. This is your another metal gate, this is also metal gate, but this is a isolated metal and this is connect is metal gate. This metal gate is connected to the word line, you know word line and bit line this is the bit rise and this is the word lighting there right.

So, this is connected to the word line, this is this is your actually we call this top metal as the control gate because that is why you are applying the word line right control gate and that is connected to the word line. But this over here sandwich between two oxide layer is actually called floating gate, floating gate the floating gate actually is floating.

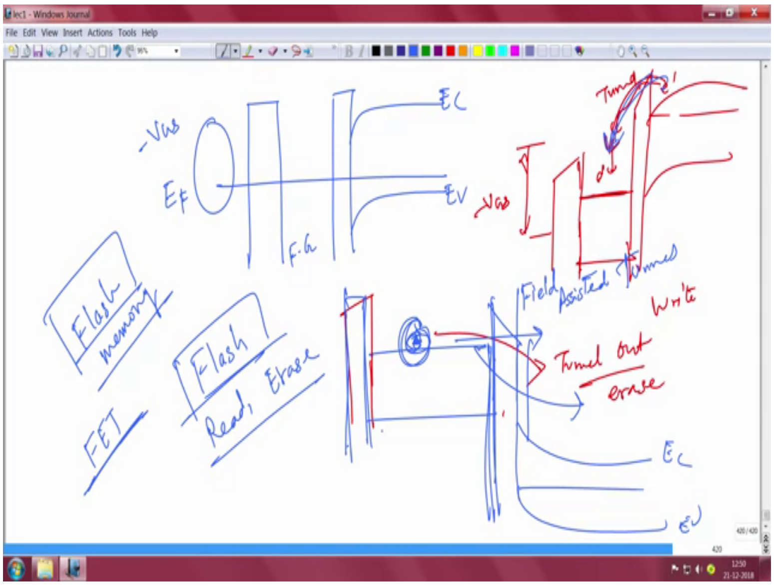
So, it is actually not do anything, I will come to that and this oxide in the lower part this oxide is actually very thin it is called tunnel oxide carriers in tunnel produce oxide actually. So, how does it work now? Actually the way it works is that you are charging and discharging this capacitance ok, you are discharging and charging this capacitance which can which have field assisted tunneling you will actually get carrier tunneling across this ok.

This is the thin oxide, you are able to tunnel through the carrier the carrier through the thin oxide and that tunneling happens very fast. Your reading and writing or your charging and discharging of this capacitance will depend on this tunneling current that is there I will come to that quickly. And this tunneling current happens this charging discharging happens very quickly in a flash that is why it is called flash memory. It happens like a flash ok, the way it works is that you know even the control gate if you induce a positive charge then you know you will eventually have some negative charges here right.

So, on the see on the floating gate this is the floating gate right this is a floating gate, on the floating gate you will have no net charge only polarization charges induced by the charge on the top gate ok. And you will have a reduction in the electron charge in the channel due to charge if you have charge if you are inducing charge on the in the floating gate if you induce charge on the floating gate, then you will reduce the charge on the by charged conservation principle, you will reduce the charge in the channel; you will reduce the charge in the channel right.

So, essentially what you are trying to do is that by this word line you are putting this charge or taking the charge here from the control gate you are able to change the charge state of the channel by changing the charge state of the floating gate here, this is the floating gate here and extra charge stored can be tunnel through this oxide how. So, that you have to draw the band diagram right. So, I will quickly draw the band diagram here that you will see.

(Refer Slide Time: 24:20)



So, in equilibrium in equilibrium I will draw the band diagram ok. So, this is your Fermi level; this is the Fermi level and you have your floating gate in between.

So, this is your oxide and you have your conduction band and you have your valence band this is your p type substrate this is a floating gate. So, I will call it FG, it is your FG. Now, when you are writing the up when you are writing on this device what happens is that you are applying a negative this is the main gate control gate you will apply a negative voltage on the gate.

So, what will happen is that your device, your band diagram will look like this ok; your band diagram will look like this. So, we applying a negative voltage here minus V Gs and your bands will bend like this so bands will bind like that. So, what will happen is that you will be able to tunnel inject an electron, you will tunnel inject an electron and it will form in the level here ok, you will tunnel inject an electron here right that is called the write operation right.

And when you store you remove the charge it will store basically, what will happen is that you will have a you will have a when you remove that bias you will be able to store the electron will be stored here ok, the electron will be stored here the electron will be stored here. And when you write when you want to you know erase the thing then you apply a opposite polarity on the bias on the gate.

So what will happen is that you will something like that then this will have a field assisted tunneling because the field will become like that the bands will go like that this is E C this is E V, your applied opposite polarity thing this is a field very thin field and this is the field assisted tunneling will happen. Field assisted tunneling is also called Fowler-Nordheim tunneling ok.

This electron, this electron will be able to tunnel out once the electron tunnels out the information stored is lost and that is called erasing. So, this movement of electron by tunneling to here inside you know and then going outside here by erasing this tunneling in an out of the electron with the floating gate. And the tunnel oxide it happens in a flash very quickly and that is how you read and erase data, you can read and erase data ok; you can erase and erase and erase data.

So, this charge this electron that is store unless you apply field here it will not be able to tunnel out itself because a tunneling oxides it will cannot tunnel out unless you apply a field. So, this electron will remain there as long as you want in a way ok.

But again there is a slight discharge might happen, but that is the different field of study altogether. So this is the basic operation principle of flash memory and it uses a p type silicon or n type I mean this n np silicon. And it is very fast and you can store information very quickly you can read and write very fast you can have very dense information because your transistor and MOSFET fabrication techniques are very advanced you can write a light lot of this transistor in a very small area.

So, that is the reason why you can have very high density of information that is stored within actually the within a small area ok, unlike the magnetic hard disk memory magnetic hard drive memory is that you have are actually not transistor based. Those are magnetic material based you have some other mechanism to read and write memory there we will not going to talk about that ok.

But this is an main application of a semiconductor memory, the flash memory which uses a floating gate a tunnel oxide to essentially tunnel the electron in and out to read and write the data and this is unique in a application you know, it cannot do it in non semiconductor based devices very well.

So, this is a broad application of a transistor as a memory devices. So, what remains now is transistors high speed logic I have already discussed high the transistors that are useful for high speed logic which is a silicon MOSFET that we will discuss in the next class perhaps so we will end the class today here. We have discussed the final part of RF transistors Johnson figure of merit hm.

And also I told you about the power frequency you know the limit that you have an different materials that you can use for performance like indium phosphide gallium nitride. Then we move ahead and we went to the flash memory and different kinds of memory technologies, I told you that flash memory is one of the very apprising and expanding rapidly expanding memory device market.

Silicon MOS structures which are floating gate in between the gate oxide and tunneling oxide enables this you tunnel the electron in and out or through the tunnel oxide and that is how you read and write the data, it is its it happens in a flash like very quick that is why it is called flash memory.

So, what remains now is probably a couple of more lectures to end this course module fundamentals of semiconductor devices. So, in the next lecture we will quickly go through what is needed in a CMOS logic you already know the MOSFET MOS transistor that is what basically CMOD logic is all about. I told you we will quickly run through some of the techniques used in device fabrication, what are all practical or experimental techniques used in making the transistors diodes LEDs and so on.

May be in the last lecture we will try to solve some of the question papers and some questions semiconductor numericals and questions that are asked in some common exams like GATE or NET or something else to give you an idea how semiconductor questions has have to be attended ok. So, with that we will end the class here.

Thank you for your time.