Digital Integrated Circuits

Dr. Amitava Dasgupta
Department of Electrical Engineering
Indian Institute of Technology, Madras

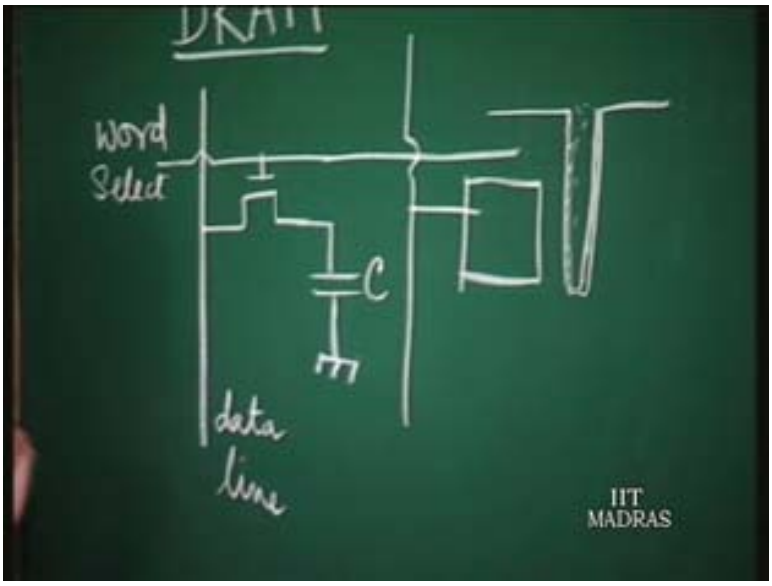
Lecture No # 35

DRAM-CMOS and BiCMOS

We have been discussing memories and last few classes we discussed the static RAM. Today's class we shall take up the other form of memory which is the dynamic RAM as you know is very popular nowadays, the reason being that you can get very large capacity of memories in a single chip. We shall take up what is dynamic RAM. How do you actually realize it and what are the properties of such a memory? The basic cell of a dynamic RAM as it is nowadays is just you have simple pass transistor followed by a capacitance, this is the cell.

The memory basically have storage. The storage is done in the capacitor that is the charge is stored in the capacitor and depending on whether the capacitance is charged or not, you have stored either a one or a zero. If the capacitance is charged you say you have stored one, if the capacitance is in the discharged condition you can say zero is stored in the memory cell. Basically you can see that it is a very simple structure compared to a static RAM where you had 6 transistors per cell or 4 transistors per cell. Here you just have one single transistor. The area occupied by the cell should be much smaller with the result that you can have a larger capacity of memory. Of course there is a capacitor here, one has to remember that and we know actually in integrated circuit the capacitor usually takes up a lot of area. That is the problem but nowadays with the development of technology you have what is called trench capacitors that is basically you dig a trench in the semi-conductor like this.

(Refer Slide Time: 06:07)



Then you grow a thin oxide and then you fill it up again with different material say poly silicon. This capacitance here, the oxide is the insulator. On two sides we have silicon basically this forms a capacitance and here the surface area taken up by the capacitance is not large. Although you have a large area because of the very nature of the capacitance which is in the form of a trench. You are actually not taking up lot of area in fact in modern IC technology, you have this single transistor and the capacitance integrated in the sense that the transistor actually sits on the top of the capacitor. Basically the area requirement is quiet small.

What you do is basically here you have a select line, this is the word select and this is the data line. When you select this word and this word select line goes high, this transistor which is nothing but a pass transistor turns on and the capacitance is connected to the data line and the capacitance whatever if it is basically connected to the data line and on the data line, you have the information corresponding to the voltage across the capacitance. That is how you read that capacitance. That is of course you have a large number of such cells and again just in the same way as we have already discussed for other memories that you have an array and on this particular line, you will have different cells located. Once you energize this word select, all the cells on this word line is going to be selected and then again you have many columns and then you have to again select one of all these columns and that is how you particularly zero in on one cell in the particular array but this is how you actually use a dynamic RAM.

Now why is it called a dynamic RAM? It is called a dynamic RAM because the mechanism of information storage here is charge stored across the capacitance and when charge is stored across the capacitance, there will always be some leakage paths. For example you have an MOS transistor here, you have the junction here that is the source to the substrate junction of this MOS transistor. This can be a leakage path and you also have other leakage paths what will happen is slowly the information stored across the capacitance is going to degrade with time and with the result that if you wait

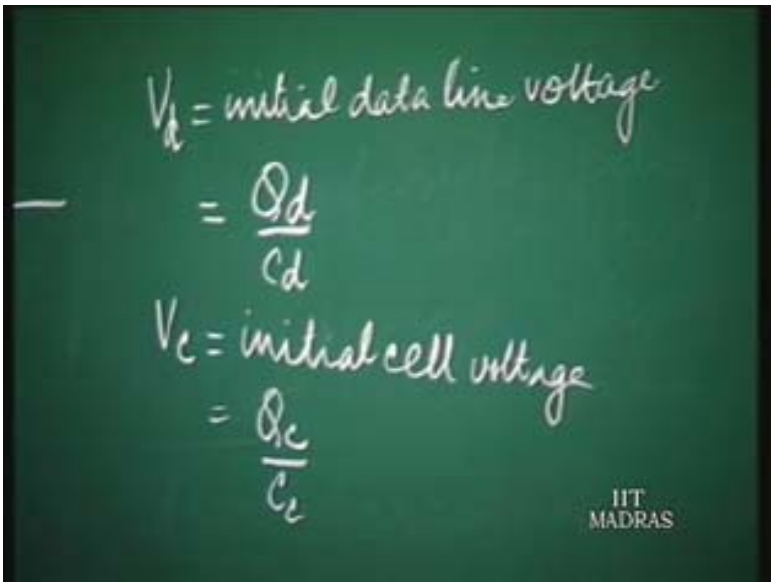
for a long time after you have written into it to read from the cell, the information may have leaked away. That is if you have charged this capacitance, slowly this charge will leak off and after a long time you have lost enough charge what is going to happen is you may read zero instead of the one which was stored here, that is the dynamic nature.

Basically what you have to do is if you have stored a one, you have to refresh the memory from time to time. That is one drawback of a dynamic RAM compared to a static RAM because of the dynamic nature of the storage you have to refresh the memory from time to time that is one. The other problem is because we want this cell to be small it must take up a small area. This capacitance again must necessarily be of a small value, although you are making a trench capacitance you cannot have a very large value of capacitance whereas this data line capacitance is going to be a very large capacitance because sitting on this data line, a large number of such cells. This data line capacitance is going to be very large.

Basically if you think of the mechanism that what you are doing is suppose you initially charge the data line to some value say $v_{d.}$ which is the initial data line voltage and then what you do is you select a particular cell. Now suppose we say that this data line voltage is equal to $v_{dd.}$ by 2 say that is if you are operating from a 5 volt power supply this is 2.5 volts and then when this capacitance is charged, the voltage across the capacitances is 5 volts, if it is discharged it is zero volts. When it is connected to the data line, if it is fully charged it is going to increase the data line voltage and if it was zero volts it is going to pull down the data line voltage.

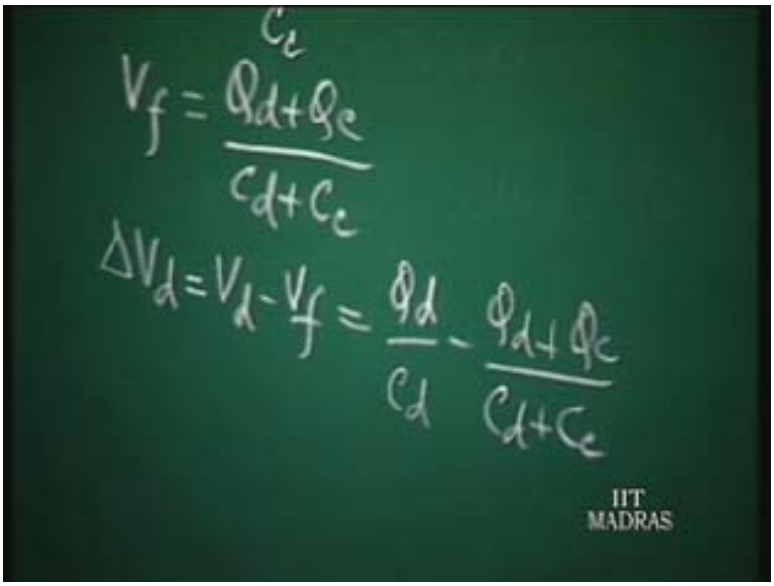
Now let us just do a small calculation and see what is going to be the change in the data line voltage. Now this $v_{d.}$ we can say is equal to $Q_{d.}$ by $c_{d.}$ where $Q_{d.}$ is the total charge, data line charge we can say and $c_{d.}$ is the data line capacitance. Similarly if we say that $v_{c.}$ is the initial cell voltage or that is basically the capacitance voltage that is equal to $Q_{c.}$ by $c_{c.}$. That is the total charge across the capacitance and $c_{c.}$ is the cell capacitance that is basically the capacitance here.

(Refer Slide Time: 10:10)



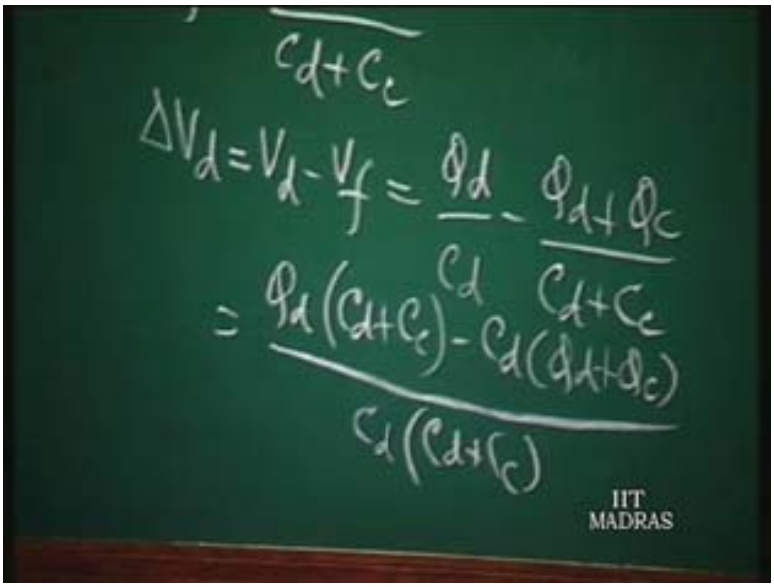
Now what is going to be the final voltage? When you connect the two when the cell is selected, it is going to be the total charge by the total capacitance. This is the final voltage that is basically when you connect them, you will have a common voltage across the capacitance and the data line because basically if you consider this transistor to be a short, there will be a common voltage across the capacitance and the data line. This is the final voltage, the common voltage. The delta $v_{d.}$ or the change in the data line voltage is equal to, we can say is equal to $v_{d.}$ minus $v_{f.}$. This is equal to... what is $v_{d.}$? $Q_{d.}$ by $c_{d.}$ minus $Q_{d.}$ plus $Q_{c.}$ by $c_{d.}$ plus $c_{c.}$.

(Refer Slide Time: 11:39)



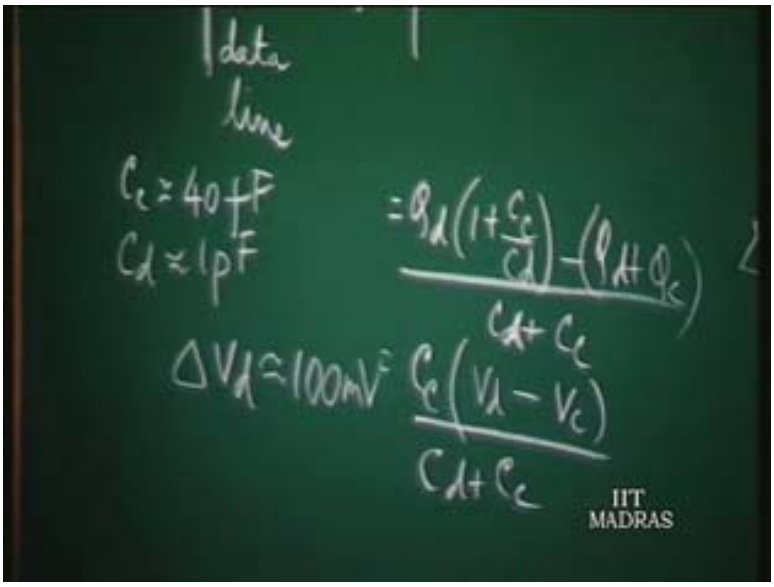
This we can simplify like this, you can write it is equal to just Q_d into c_d plus c_c minus c_d Q_d plus Q_c divided by c_d into c_d plus c_c. Now if you do further analysis, for example if you divide by c_d then you get which is equal to, I am writing it here Q_d into 1 plus c_c plus c_d minus Q_d by Q_c c_d plus c_c. Now Q_d and Q_d will cancel here and what you are left with is Q_d by c_d into c_c, Q_d by c_d is v_d and this is Q_c which is nothing but c_c into v_c. You can take c_c common here, you get Q_d by c_d is v_d minus here you get v_c by c_d plus c_c. What you get is v_d is the initial data line voltage, this is the difference the change in the data line voltage after the cell is connected to the data line.

(Refer Slide Time: 12:23)



The v_d - v_c is the initial difference between the data line voltage and the cell voltage into this c_c by c_d plus c_c.

(Refer Slide Time: 15:02)



Now the cell capacitance is usually much smaller compared to the data line capacitance. Some typical value would be, for example c_c would be of the order of say 40 femto farad whereas c_d would be of the order of say one Pico farad. If you take that values and this v_d - v_c suppose you have initially charged the data line to say 2.5 volts and say basically v_d minus v_c, v_c is say 5 volts or zero volts, the difference v_d minus v_c is 2.5 volts. Basically what you have is c_c by c_d plus c_c, c_c is much smaller than c_d this is 40 by 1 pf is around 125^{th}. This delta v_d under such situation will be almost about 100 milli volts.

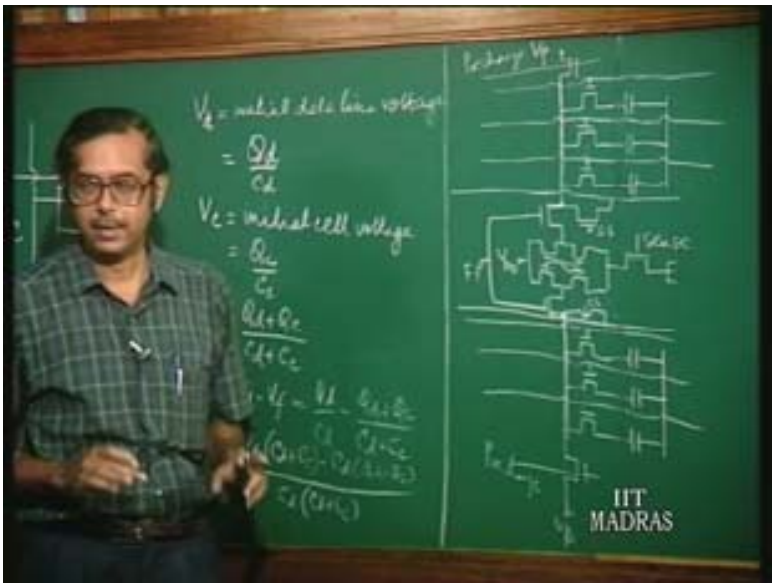
What is the situation? Initially you have charged this to 2.5 volt v_{dd} by 2 and then you have connected the cell irrespective of whether the zero is stored here or a one is stored here. The data line voltage is going to be v_{dd} by 2 plus minus 100 milli volts. If a zero is stored here, it becomes v_{dd} by 2 that is 2.5 changes to 2.4 volts, if it was charged to 5 volts it would just go up to 2.6. You have to detect that very small difference. It is very difficult and the other thing is since this cell is connected to this data line, the cell voltage also becomes the same. Whatever is stored in the cell is lost, initially whatever you had stored whether a zero was stored or a one was stored, now after you have read from the cell it is no longer there because the voltage has become the same as the data line voltage, irrespective of zero or one it has now become very close to 2.5 volts.

These are the problems that is the data line voltage change is very small and number two while reading the memory, the contents of the memory is lost. How do you overcome these two problems? Any reading mechanism must be able to overcome these problems. The way to do it is basically again when you are reading from the memory, you have to have again a mechanism like you are detecting not the absolute value of the

voltage but a differential voltage that is whether it is increasing above 2.5 or going below 2.5.
A mechanism of that sort, a differential voltage and also you have to ensure that while you are reading, you also refresh the memory or you basically rewrite into the memory. That is also necessary because the reading mechanism must not spoil whatever is contained in the memory. Just we shall just take up DRAM circuit which will do this. Basically you can consider it like this. You have what is called a pre-charge mechanism and then you have large number of memory cells like this. These are connected to the word lines. What I am drawing is just a column in the memory. You have a sense amplifier, the sense amplifier is nothing but again what you have is the back to back inverters in the flip flop configurations that is the output of one inverter feed going as the input to the other inverter.

That is basically the sense amplifier then you have a sense input here. This is v_{dd} say, then you have control to connect the data line to this sense amplifier call this FF, FF stands for flip flop. Then again you have another set of such cells, these are the another set of word line, this is another pre charged voltage v_{p} and this is the pre charge signal and then this is the sense input here and then you will have a sort of column. This is column select signal, I just put cs here also. This point is to be connected to, there is a column line here, this point is connected to this line and this is another column select.

(Refer Slide Time: 22:38)



Now basically what you have is I just explain again. This is one column of the memory and these are the word lines. Suppose you have 128 cells in a column, what you are doing is you are dividing the column into 2 parts, 64 cells on top and 64 cells at the bottom. You have 2 halves, 64 cells and 64 cells and this cell which you see is just like the cell which we have already seen. First is you pre charge these data lines. You have

two halves of the data line, one is the bottom half and the other is the top half. You first pre charge this data line say to v_{dd} by 2 that is done through this. This is the pre charge voltage. When the pre charge signal comes, both the data lines are pre charged to v_{dd} by 2.
Then what happens is of all the other signals, the sense is off initially and all this FF is also off and the cs that is the column select is also off. Now what happens is one of the cell, either you are selecting a cell in the upper half or in the bottom half, only one word line can go high, if you are selecting a cell in the upper half that means all the cells in the bottom half or not selected. Say suppose you have charged to 2.5 volts, this data line voltage of the bottom half is going to remain at 2.5 volts but in the upper half, if a one is stored in that memory cell, this upper half data line voltage is going to go up by say 100 milli volts as we have seen. It is going to go up to 2.6 volts. The upper half data line voltage and the lower half data line voltage there is going to be a difference. If zero was stored this may go down by 100 milli volts to 2.4 volts.

Then what you do is you energize this flip flop. There are two pass transistors here, they connect the data lines to this sense amplifier. These are pass transistors (Refer Slide Time: 24:12). What you do is first you select the cell, let the differential voltage develop and then you connect the data line to the sense amplifier as well as you energize the sense amplifier by making this transistor go high. As long as this sense input is low there is no current flowing, this sense amplifier is not operational. There is no current flowing. Only when this transistor turns on then the sense amplifier is activated and then depending on the difference in voltage at these two points, the sense amplifier output is going to go to one of the two stable states.

If the upper half voltage is higher, the upper line is going to be pulled up all the way to 5 volts and the lower voltage is going to go down all the way to 0 volts because in that part of the characteristics, it's an inverter because the gain is very high and they are connected with positive feedback. It is going to pull up one side to v_{dd} and the other side is going to be pulled down to 0. In the same process what you are doing is suppose we had selected one cell in the upper half. The upper line voltage had gone up by 100 milli volts.

Now this line is going up to 5 volts and basically what you are doing is you are rewriting into that particular cell. That cell itself is being returned back to 5 volts. Now if this side had gone down by 100 milli volts, so 2.4 volts. What is going to happen is this lower half is going to go up to 5 volts, the upper half is going to go to zero. Basically you are writing back into the cell. You are writing a zero back into that particular cell. The lower half no cell is selected, it is not going to affect the contents of any cell in the lower half. Basically by the sense amplifier it is doing two functions. One is it is magnifying that small difference in voltage. You are going to have 5 volts on one side and 0 volt on the other side, its amplifying the difference and also it is rewriting into that particular cell, the voltage of which was affected during the reading mechanism and then you do the next thing.

Once you have developed that 5 volts at 0 volt difference, there are two transistors here. This is one column line, the two lines here. You have a number of columns situated like this and this is one particular column. You have many similar columns with a sense amplifier and then this is the column select, all of these columns will have the input here.

Again you have a decoder and the output of the decoder is going to go to this column select. Basically you select one of the so many columns and only one column's output is going to be connected to this output line. This is a pass transistor again, what this does is it connects the output of the flip flop or the sense amplifier to the column line. Then once you have developed that large voltage difference, you choose one particular column through that y decoder and then that is connected to the output lines. Basically one has to be careful in that there has to be a proper sequencing of all this. First you select the particular cell in a column, let the small difference develop then only you energize the sense amplifier. You should not do that before because if you energize the sense amplifier before what is going to happen is it is going to go to one of the two stable states.

Once it goes to one of the two stable states, a small voltage difference of 100 milli volts is not going to affect that stable state. First you have to develop that small difference and then you have to energize the sense amplifier and then once you have energized the sense amplifier, the two voltage levels are developed that is 5 volts and 0 volts then only you do choose the particular column. You don't choose the column beforehand because the capacitance, if you choose it beforehand the column is connected to the data line and also the column capacitance will come into picture which is again going to affect the differential voltage. Once you developed that 5 volts to 0 volts then only connect that particular column to this output line because this line also has a large capacitance because all the columns are sitting on that line. Because you require a definite sequencing of operations, it makes the dynamic RAM comparatively slower compared to the static RAM. This is how you operate a dynamic RAM. As we said the dynamic RAM, the problem is you require to refresh the memory from time to time.

Obviously if you read from a particular row say suppose you have energized this particular first row say, row zero what is happening is in all the columns the row zero is energized and the row zero value is actually read by the sense amplifier in all the columns and the sense amplifier actually refreshes that memory location. But it may happen that when you are reading, you may not have read a particular row for a long time. All the cells in that particular row, if it was charged to 5 volts, the charges may go down and you actually have lost that whatever was stored in that memory. What you have to do is deliberately you have to refresh the memory location. How do you do that? The way to refresh a memory is by reading that memory.

Once you read a particular cell, we actually refresh that particular cell. What you have to do is you have to deliberately read every cell after some interval of time and in that process you basically refresh that particular memory cell. In fact what you do is once you read that means you have to energize sequentially all the rows one after another. Once you energize a row zero, on all the columns you were actually refreshing that particular cell corresponding to row zero in each of the columns. If you have a square array of n bits say, you have root n times number of cells in each column and you have root n number of columns.

Basically you require root n refresh cycles to refresh all the memory locations. Basically to read each row, energize each row all the cells are refreshed. That is how you refresh the memory just by reading mechanism. This is the configuration for a normal MOS type dynamic RAM, we shall now take up another example of a dynamic RAM where which uses the BiCMOS concept that is using bi polar transistors. Now obviously the reason to use bi polar transistor is to speed up operation. I shall just give a reference and we shall take up the structure they have used and just see how it actually improves the operation. Its kitsukawa and others IEEE again same journal of solid state circuits, volume 25, number 5, pages 1102 to 1111 October 1990. (Refer Slide Time: 34:28) That is the reference.

This is different way of sensing, I just give you how they do it. What they have is again the cell is quite similar, you have a transistor and a capacitance and then this is one particular cell. Again I will draw the column, all although it is a column you are drawing it like this. What you have is each column will have a sense amplifier, again the sense amplifier is the same sense amplifier. You have the CMOS sense amplifier that is 2 CMOS inverters, the output of one going in as input to the other and these outputs are connected here. This is the sense amplifier then this is called the CMOS sense amplifier as well as the rewrite amplifier. Basically it does two functions, it senses what is the voltage in the cell and at the same time rewrites back into the particular cell then you have a pre charge circuit. This is the pre charge circuit. The function is to pre charge the two lines, the bit lines. This is pre charge voltage $\text{phi}_{p.}$, these are two MOS transistors. When $\text{phi}_{p.}$ is high, these two transistors turn on and $v_{p.}$ is passed on these two lines.

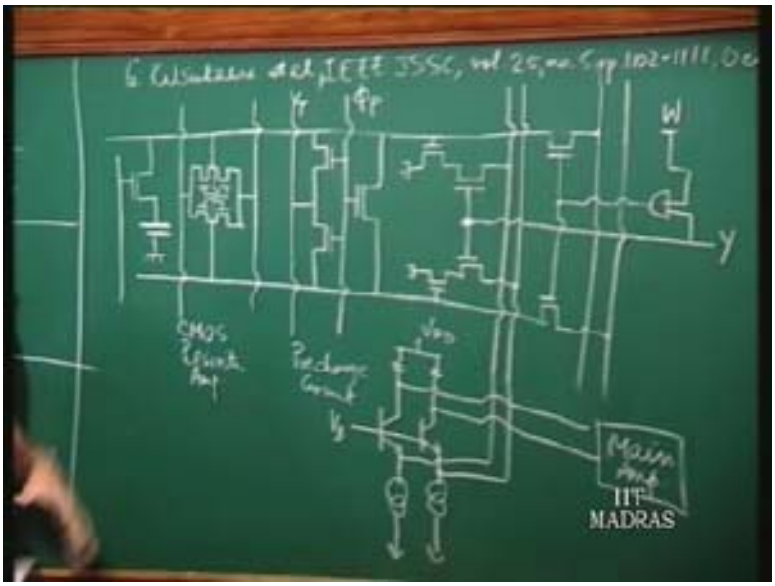
There is another transistor which is basically connected across the lines and when this is on, the extra transistor is there to equalize the voltages between the two lines, just acting as a sort of connection between the two lines so that it ensures that the two line voltages are the same. Then you have the mechanism for reading. You have two transistors here, these two lines carry the output of each column. This is the y select line. When you have two more lines this is for the writing mechanism, you have two transistors sitting here. One is connected here and the other is connected here. This is energized only when you want to write into this cell. This is write and this is the AND function of this and this (Refer Slide Time: 39:01). Then you have the amplifier which

is used to amplify the write signals. The basic structure is like this, you have two bi polar transistors. This is a constant based voltage here and then these two lines are connected to the emitters of the two transistors. This is V_{DD}, these two go to the main amplifier.

Now this is the circuit that I will explain. As I said this is a particular column, you have so many columns all sitting on this set of lines. You have all this columns connected to this set of lines and in this column, you have a number of such cells. I have just shown one cell but you are going to have a number of such cells. This is the word line which selects a particular cell. Suppose you select this particular cell, what happens is if you select the cell, this particular voltage on this line is going to get affected, this part of the bit line (Refer Slide Time: 41:01).

Again just like the previous circuit which we have discussed, the line voltages are pre charged to a particular value with the help of this pre charge circuit.

(Refer Slide Time: 40:19)



This pre charge circuit using these three transistors the pre charges, the two bit lines, two common voltage say V_{DD} by 2. Now you select a particular cell on this column and what is going to happen is the voltage in this upper line of this column is going to get affected. Again as we have seen this voltage difference is going to be pretty small, small voltage difference. This line voltage may go up by a 100 milli volts or go down by a 100 milli volts but it creates a difference between the two bit lines then what you have to do is you have to energize the sense amplifier. Basically you have the two signals here phi_s and phi_s prime connected to this sense amplifier. Obviously if phi_s is high and phi_s prime is low then only the sense amplifier is going to get activated. If it is other way around then it is not going to get activated it's not operation. Once it gets

activated after this creates a small differential voltage, the sense amplifier is activated which creates a large difference in voltage between the two lines and then what happens is this line voltages are different.

Now we come to this part here. What is the structure here is you have two transistors here. The gate of these two transistors are connected to the bit lines. What is going to happen is one of the transistors, the gate voltage is going to be very small less than the threshold voltage. The other transistors, the gate voltage is going to be very high. Then you have these two transistors, these two transistors are nothing just like pass transistors. The y is connected here, it basically selects a particular column. This is a column sitting on these two lines. You have number of columns, only one of the columns is connected to this particular line. Suppose this column is selected, since one of the transistors is selected what is going to happen is a current is going to flow in one of these two lines. It is a drain current so it is going to flow the other way.

Here these are very small current sources, 50 micro ampere current sources. Now the emitter current of this transistor is equal to this 50 micro ampere plus the current flowing on that line. You have two transistors here which has a common base voltage. The base voltage is the same $v_{B.}$ which means that the emitter voltage is almost constant. You see for a bi polar transistor even with large change in current, the base to emitter voltage difference is very small. What is going to happen is suppose one of these lines is going to carry current. For one of the bi polar transistors, the current is going to be 50 micro amperes. For the other bi polar transistors it is going to be 50 micro ampere plus the current which is drawn by the MOS transistor which may be as large as about one milli ampere.

It is 50 microampere and one milli ampere that is the emitter current of this bi polar transistor. The emitter current is almost equal to the collector current, these are resistance here. The drop across one of the resistances is going to be very large, the other drop is going to be very small. You get a differential voltage here and this differential voltage is fed to a main amplifier which is another bi polar differential amplifier which is similar to the structure which we have seen in the case of a static RAM. This is how you detect these voltages. Now what is the advantage of using a bipolar transistor here? The advantage here is that although you get a large voltage difference at the output you see on this line across which all the columns are connected, the voltage is going to change by a very small amount. Capacitance of this line is going to be very large because many columns are sitting on this lines, it is going to be very large capacitance. In fact for this particular circuit it was about a Pico farad capacitance.

Now this capacitance if you have to charge it through a large voltage it would take some time but here since the emitter voltage of the bi polar transistor is held almost constant, what you are sensing is basically a current. This line voltage, the voltage across this line is almost a constant, it is not changing. The delay is going to be much

less. If it was a voltage which you are sensing then what would have happened is that voltage had to go all the way, if we have to change from 0 to 5 volts say then it would take a long time because the line capacitance had to be charged through that voltage.

Although you get a large change in voltage at the output, you see across this line the voltage change is minimal. Whatever is the v_{be} of this voltage of this transistor which is taken is very small because of that it is going to be a much faster mechanism. By this way you can sense what is stored in that memory cell. That is how you incorporate a bi polar transistor to improve the speed of the dynamic RAM. Also here this part of the circuit is of course to write into that cell. If you select y and you select w that is write, whatever is on this lines goes into the data line and you can write into that memory cell. This is the circuit for a dynamic RAM with a bipolar transistors. The idea is almost the same. Again you have to use a differential voltage you cannot sense the actual voltage in the dynamic RAM directly because once you connect it to the bit line, it creates a very marginal change in the bit line voltages.

You have to compare it with another voltage. The difference in the two voltages are sensed, you require a sense amplifier to sense the difference in voltage. At the same time to rewrite into the particular cell and here it is just the difference in the reading mechanism. Here you use a bipolar transistor to sense the bit line voltages. Once you have created that big difference in the two bit line voltages, the bi polar transistor advantage is with a small you can have a large difference in current. Even in that case the v_{be} difference is going to be minimal. You can actually sense the current rather than the voltage and the line voltage changes by a very small amount on this line which is going to make it much faster. That is how with a help of the bipolar transistor, you can improve the speed of operation in a dynamic RAM. With that we conclude this discussion on dynamic RAM.