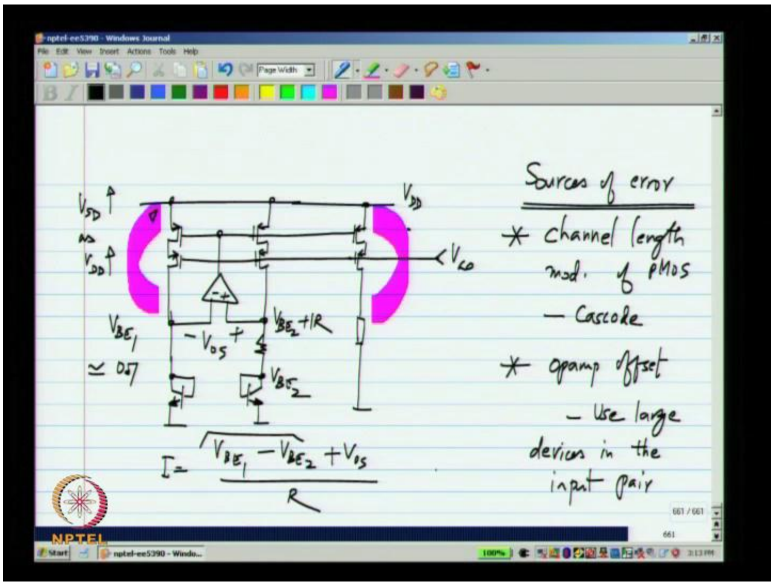
Analog Integrated Circuit Design
Prof. Nagendra Krishnapura
Department of Electrical Engineering
Indian Institute of Technology, Madras

Lecture - 54
Bias Current Generation and Band gap Reference

Hello and welcome lecture 54 of Analog Integrated Circuit Design, we are going discussing how to generate current references such that a transistor which is bias with this current as a constant g m. We will look at it little more that is sources of where are in such references, and move on to voltage references.

(Refer Slide Time: 00:32)



We have looked at the PTAT cell what we had was these two PMOS transistor generating the reference current for bipolar transistors, and the entire circuit is self biased. Now; obviously, the implication here is that, this current is replicated somewhere else and used, and at many places for us, now one of the important things in a reference is that it should also not be very sensitive to the supply voltage. Now, in this particular circuit the voltage is here V B E 1 and V B E 2 plus I R which is the same V B E 1, these are almost constant there be a 0.7 volts as so.

So, as the V D D changes the source drain voltage of these PMOS transistors V S D increases as V D D increases. So; that means, that this could be operating from a very different source drain voltage from this transistor, which is actually sourcing the current into the desired circuit. So; that means, that the actual current that is flowing here is different from the self bias current that we generated, so that can also lead to errors.

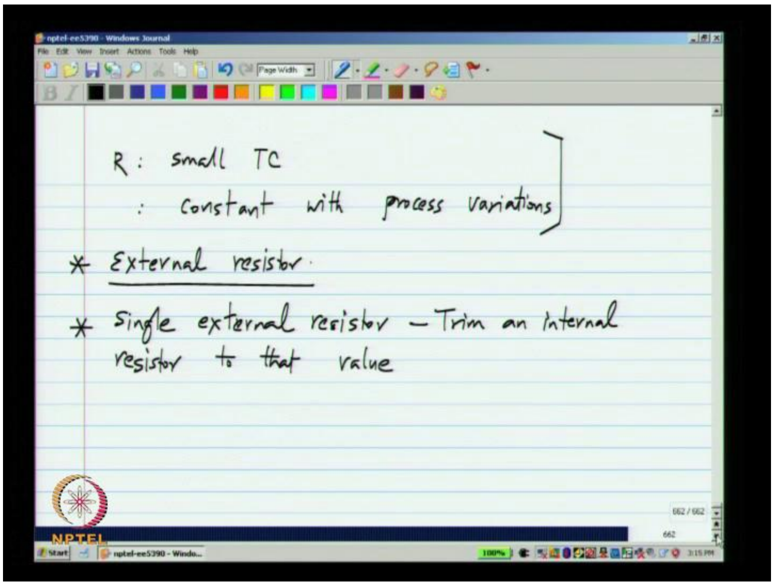
So, the easy way to fix this is by making this current mirrors more accurate, and less sensitive to V D D that is by making them case code current mirrors, what we can do is to introduce case code devices here. So, in this case a current mirror also would be a case code, and we have to bias the certain V G naught, which can also be generated using self bias that is mirror this and backup, just have to make sure that the circuit starts up correctly.

I would not show the details of that, the only thing is it should be such that these transistors remain in saturation. So, in that case what happens is this voltage can be very different from this, but the currents will not be, so different because of the case coding effect that we have seen earlier. Another source of error is the offset of this OPAM. We said that this V B E 2 plus I R will be force equal to V B E 1 by the action of negative feedback.

But, in presence of offset there will be a difference V O S between these two, so this current finally, will be V B E 1 minus V B E 2 plus V O S divided by R. So, this will not be exactly proportional to absolute temperature, even if this difference is exactly proportional to absolute temperature. So, we have to make sure that this V O S is much smaller than V B E 1 minus V B E 2, and how to do that basically in this half amp we have to use a large enough input differential pair, so that offset of this differential pair is rather small.

So, the sources of error could be a channel length modulation of PMOS transistors for this use case code, and OPAMP offset and for this use large devices in the input pair. So, these are the common precautions that one vertex, while designing circuit of the sort.

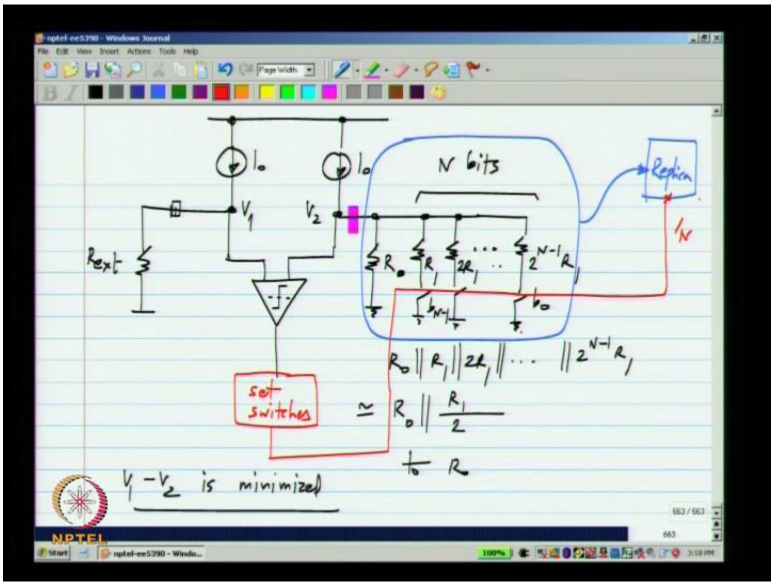
(Refer Slide Time: 06:02)



Now, we also require that this R have a small temperature co efficient, otherwise it is temperature co efficient will appear in the temperature co efficient of the current. And also it be constant with process variations a both of these are relatively easy to do, if we have an external register. External register will be constant with process variations, and also we can have by an external register, which has very small temperature co efficient.

But, one of the problem says that first off all you may not want to use an external register, and especially many times you have a number of such circuits on your integrated circuit. So, you certainly do not want to use many such external registers, one for each one because, that would become to expansive and the number of pins, and it will increase the number pins unnecessarily and, so on. Again we can exploit matching between on chip registers to overcome this, we can use the single external register. And you trim an internal register to that value, there are number of ways of doing it I will just outline one of the possible ways.

(Refer Slide Time: 07:48)



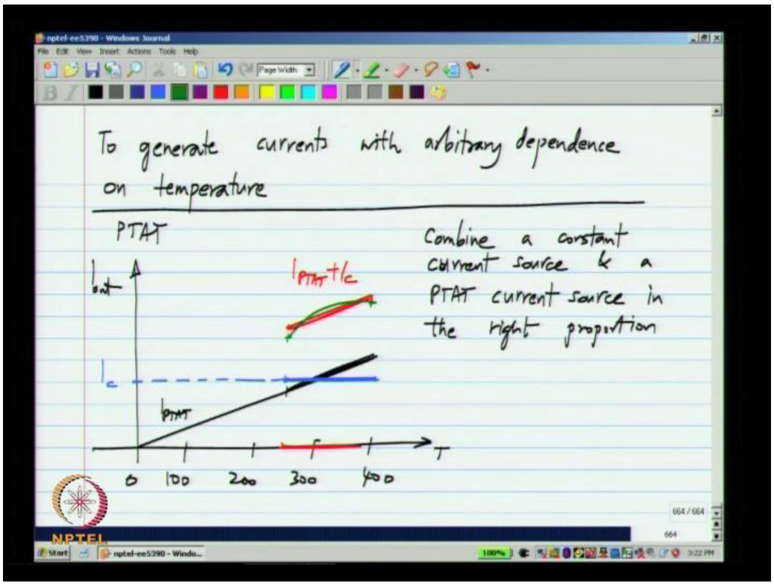
What you do is, let say we have an external register R e x t and this is the ground, and this is an on chip current source that will passes certain current. And you can pass the same current through an on chip register, which is variable and how do you make it variable one of the commonly used west is to use a set of parallel registers, which can be switched well I will call this R naught and this R 1. Let say 2 R 1 and, so on up to 2 to the N minus 1 R 1 that is N bits of an resolution, and this is b 0 up to b N minus 1.

So, by controlling this switches you can change the value of the resistance that appears between that point and ground. And it can change from a minimum value of R naught parallel of all these resisters, which is basically R naught in parallel with approximately R 1 by 2, all the way to R naught an all the switches are open. Then what you do is you compare these two voltages in a comparator, and based on the output of the comparator you set the switches, you can for instance 0 binary search.

Such that finally, the difference V 1 minus V 2 is minimized, what then happens is that, the effective resistance between this point and ground will be as close as it can be two the resistance between that point and ground or the external resistance. Now, after that any time you want a resistance on chip that is accurately defined, what you do is you replicate these entire thing. This will be a replica, and it will also have N bits of control and the same N bits are used here, and replica could also be scaled if you wish.

So, this is one common way of realizing a numbers on chip, which are lock to an external register. So, this also can possibly used to realize R, so there are number of techniques like this which if you think about the situation will become obvious to you or if you look at the literature you can study those things.

(Refer Slide Time: 11:20)



Now, one of the other requirements may be to generate currents with some arbitrary dependence on temperature that is we have know how to make a current that is portioned to absolute temperature. But, you may be required to have a current that has some dependence like this or like that or something, this may be to compensate something of rather. Now, many times what you can do is to combine a constant current source, and a proportional to absolute temperature current source in the right proportion to do this.

So, what I mean is, so let me say that this is the absolute temperature axis plotted all the way from 0 Kelvin, of course we will not operate anywhere near that, and our operating region will be usually somewhere over here from 0 to 100 degree c or may be even minus 40 d c to 100 c and, so on, it will over there. Now, if you have a proportional absolute temperature current source, it is variation will be linear with temperature, and in the ranges are that we are interested in it will like that.

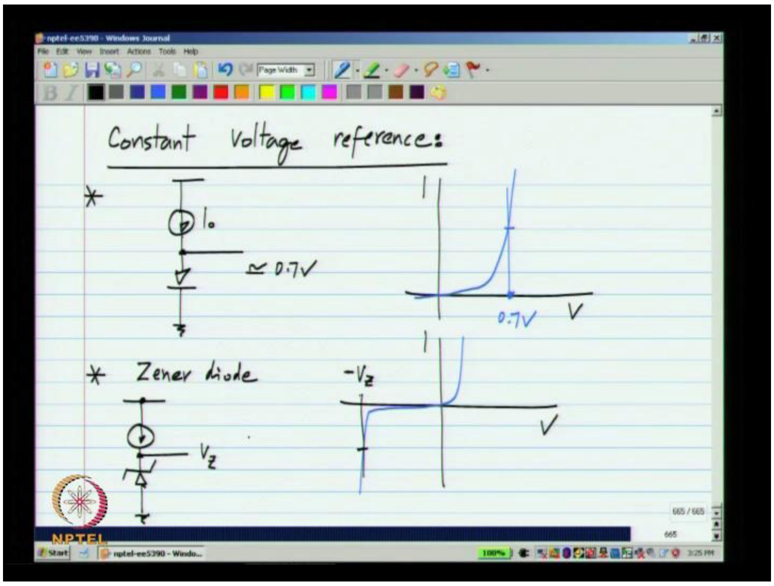
And let say you have constant current source that is constant with temperature, then in the range that where interested in it will be like that. If I sum these two what I will get will be something like that, this will be I P TAT plus I C, and this is not linear that is this is a straight line, but it is not passing through the origin. Now, this could be the approximation for a desired curve, which is like that or something.

So, what you can do is if your given a current versus temperature curve, you can look at the end points and see how it is varying, and try to generate it using a combination of I P TAT and a constant current, for many times this a crowed approximation that will surprise. And by the same token you can also generate current which have a negative temperature coefficient, instead of adding a P TAT current source to a constant current source. We can subtract the P TAT current source from the constant source, then the resulting current will have a negative temperature coefficient.

And by adjusting the proportions you may be able to get something that approximates the requirement that you have, so this is the another trick that you can use. Similarly, the circuit that generates a current, which keeps the g m of a mouse transistor constant, will have some complicated temperature dependents by combining that with the constant current source you may be able to get other dependences over temperature.

So, we have now discussed generating current references to some extent, and the current references over of the type which will keep the g m of certain transistors constant. And to generate constant current, we need a constant voltage as I discussed earlier, so what will now look at is how to generate a constant voltage on chip. Again when I say constant voltage it should not be dependent on temperature, and preferably also not dependent on process and power supply voltage and, so on.

(Refer Slide Time: 15:53)



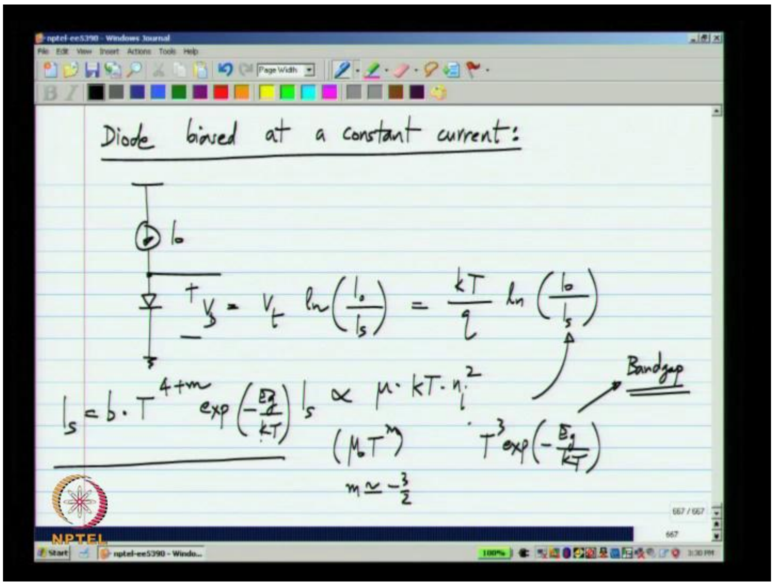
Now, there are many constant voltage references in a crewed sense that are use some times, first off all if we look at the I V characteristic of a diode, it is rather steep. And you know that for operating point calculations it is common to assume that the forward drop of the diode is just 0.7 volts, this is just an approximation. But, you can think of this as something that is roughly constant. So, if you pass a current through a diode you will get approximately 0.7 volts and you can think of that as a voltage reference, it terms out the this is too crowd it will vary with temperature it will also vary with the current and, so on. But, to some approximation it works, and especially with discrete circuits you also know that, there is something known as zener diode. Essentially it is same as diode, which is operated in it is reverse breakdown region.

And this characteristic, in this region is rather strep and if you bios somewhere there, the voltage will be by and large independent of the current and that is known as the zener voltage. Typically it is denoted by special symbol, but basically it is a reverse bois diode, which is operated in the reverse break down region, now this also has a certain temperature coefficient. Now, it terms out the there are two different breakdown mechanisms, the avalanche breakdown mechanism and the tunneling breakdown mechanism.

And one of them has a positive temperature coefficient, and the other one has a negative coefficient. But, there is a problem with making this CMOS integrated circuits, first offal you design the junction in a CMOS I C, so that the reverse breakdown is usually at a voltage that is higher than the power supply, this is for safety. And you normally do not design special junctions, which can breakdown at a voltage below the supply voltage.

So, clearly if you the diodes breakdown only beyond the supply voltage, we cannot use it to realize a voltage reference. So, what we have to do is to come up with the circuit that has 0 temperature coefficient, and uses their components that we have own chip, so what will do is first look at the diodes behavior more carefully that is I said that, the diode gives approximately 0.7 volts, but depends on temperature. So, what will do is we look at how exactly it depends on temperature, and perhaps try to cancel that one.

(Refer Slide Time: 19:00)

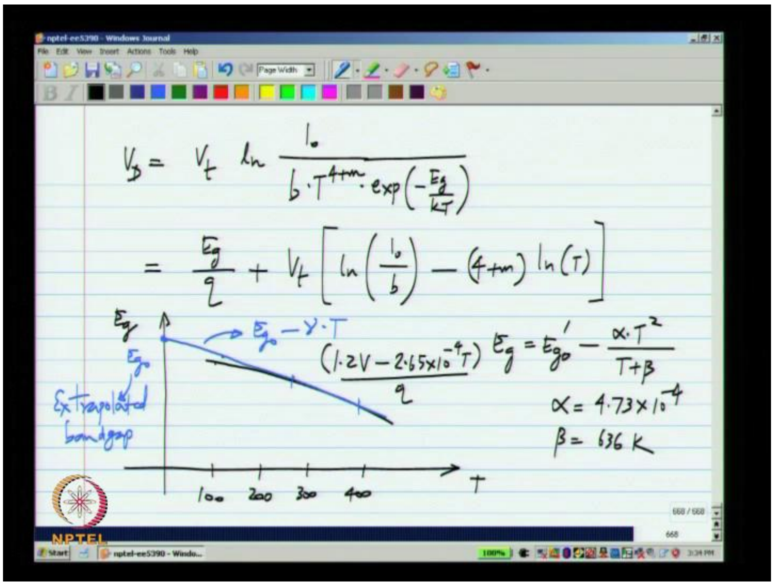


I will call this V D which is V t log I naught by I s which of course, is k T by q l n I naught by I s. Now, just looking at this it looks as though the diode voltage is proportional to absolute temperature, but of course, I s it turns out is very strongly dependent on temperature, I s proportional to mobility and k T and n i square where n i square is square of the intrinsic carrier concentration. And mu itself is dependent on temperature as mu naught T to the m where m is approximately minus 1.5 and we have k T which is of course, proportional to temperature.

And this n i square it is also dependent on temperature as T q exponential minus E g by k T, where E g is the band gap. And of course, k T is proportional to temperature here, but there is a dependence on T cube, which makes I s very strongly dependent on temperature it increases with temperature. So, the net result is that the diode drop V D actually reduces with temperature.

So, putting all of these things together I will write I s as some constant b which includes this bolds ones constant. And this mu naught over there T raise to 4 plus m, we get m form here, 1 from there and 3 from there, so it becomes 4 plus m, and I also have exponential minus E g by k T. So, that represents I s verses temperature, the store is not complete though because, E g itself is dependent on temperature we look at it shortly.

(Refer Slide Time: 21:36)

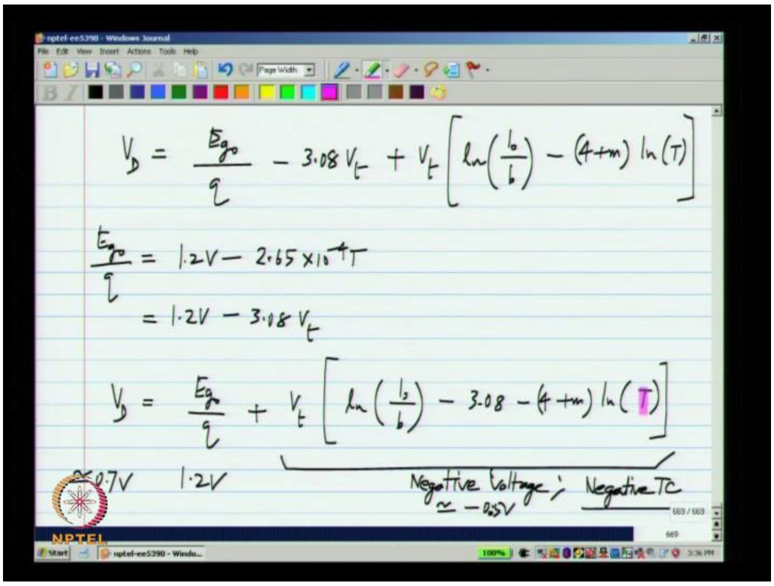


So, the diode drop or the base emitter voltage of bi polar transistor would be V t lan I naught divided by b T to the 4 plus m exponential minus E g by k T. But, if you expand you will get E g by q this by q come from v T being k T divided by q plus v T times log I naught by b it is represents the constant terms over here minus 4 plus m log T. Now, like I said E g itself is dependent on temperature, and it terms out have rather complicated dependents on temperature, which is E g 0 prime minus alpha times square by T plus beta.

And for silicon this alpha is 4.73 times 10 to the minus 4, electron volts for Kelvin and beta is 636 Kelvin. So, if you plot it, it looks something like that, now again we are interested in the temperature range around 250 and 350 or 400 Kelvin, and over that range it can be approximated by a straight line, this is the range of interest. And it can be represented as some E g naught minus gamma times T that is it is a straight line, which starts form E g naught which I will call the extra polated band gap, and then it decreases linearly with temperature.

Now, this approximation is not valid in all region, but in the region that we are interested in it is valid, the two curves become close to each other. And it terms out that these values are E g naught is approximately 1.2 electron volts, rather I will write E g naught as 1.2 volts minus 2.65 times 10 to the minus 4 T divided by q. So, what is on top will have dimensions of volts and if we divided by q you will naturally get the electron volts.

(Refer Slide Time: 25:07)



So, V D can be further rightness E g naught by q and like I said E g naught by q could be sorry this is not this divided by q, but this times q, E g naught by q could be 1.2 volts minus 2.65 times 10 to the minus 4 T. And this because, it is proportional T I will write it in terms of the thermal voltage V t it will be 1.2 volts minus 3.08 V t, so I have minus 3.08 V t plus V t times l n I naught b minus 4 plus m log T. So, this is what we will have, and it is clear that this E g naught by q is approximately 1.2 volts, and we know from experience that V D is about 0.7 volts may 0.8 and, so on.

So, this entire thing will be a negative voltage, so it is approximately minus 0.5 volts are so, and top of it this entire thing will also proportional to V t that is its proportional to absolute temperature, except this last term which also has log T in it. So, this entire term on top of it will also have a negative temperature coefficient that is this entire term will become more negative with increasing temperature.

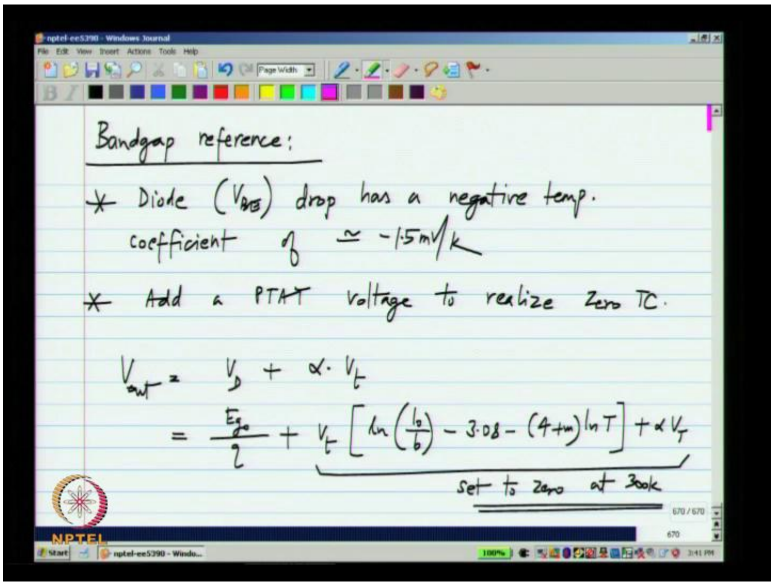
So, the diode drop turns have a negative temperature coefficient, and again you may have read about this that it has negative temperature coefficient of about 1.5 milli volts for Kelvin or 2 milli volts for Kelvin. It depends on the type of the diode that you have, and how much current you bios it with, but roughly that is a number. But, in this entire thing there is one quantity that is completely independent of temperature, and also anything related our circuit that is E g naught.

The extra polited band gap that a characteristics of silicon, and also not a dainty temperature, but extra polited 0 degrees. So, if we compensate for the remaining terms in this expression, we will get a voltage that is just E g naught by q, and that will be absolutely constant and independent of temperature. So, that is the sense behind this voltage reference, which is known as band gap reference it is called a band gap reference because, it is final output voltage will be the extra polited band gap voltage E g naught by q or 1.2 volts.

Now, you may have seen in catalogs that a number of band gap references are available with 1.2 volts as the output voltage. And the region is that the extra polited band gap silicon is 1.2 volts, so what we will do, is we will see how to cancel the remaining terms which our negative temperature coefficient. Now, how do we cancel in negative temperature coefficient by adding a term which has a positive temperature coefficient.

And that we already have, we have something that is proportional to absolute temperature, which has a positive temperature coefficient and that can cancel the negative temperature coefficient given here. In fact, if you look at the remaining term besides E g naught by q that is proportional to V t, except for the last term which also has l n T. So, it is blindly non-linear, so to that if you add something that is proportional T with a positive coefficient, then you can potentially cancel it at some particular temperature, and also get a rough cancelation at a wide range of temperatures. ((Refer Time: 29:15))

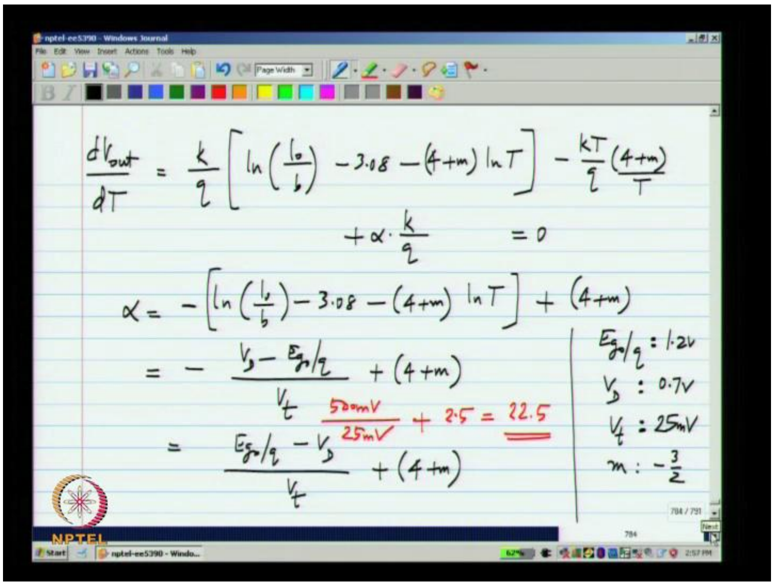
(Refer Slide Time: 29:15)



So, what will do is said this part to 0 to 300 Kelvin, which will cancel of the temperature coefficient at that temperature. For other temperate, they will be it will not be exactly constant, but usually we can leave with that, and there also sophisticate by, so that correcting for that as well. Now, one of the things is this looks like it is dependent on this precise combers of I naught by B etcetera, etcetera, but that is not the case.

So, in simulation we can determine what is the value of alpha, and then as long as these things are modeled correctly, we will able to do the cancelation. We do not need to calculate from these values as well shortly show for illustration, but as long as the dependence of diode drop across temperature is modeled correctly in your device model you will be able to do this.

(Refer Slide Time: 31:41)



If derivative of the output voltage with temperature is given by this term of course, is constant with temperature, and this whole thing becomes V t which if you differentiate you had k by q, and you have this constant terms. And also this has to be differentiated which will give you and finally, for the last term which we are going to add will have alpha times k by q. And this entire thing has to be set to 0, so alpha terms are to be l n I not by b minus 3.08 minus 4 plus m l n T minus 4 plus m.

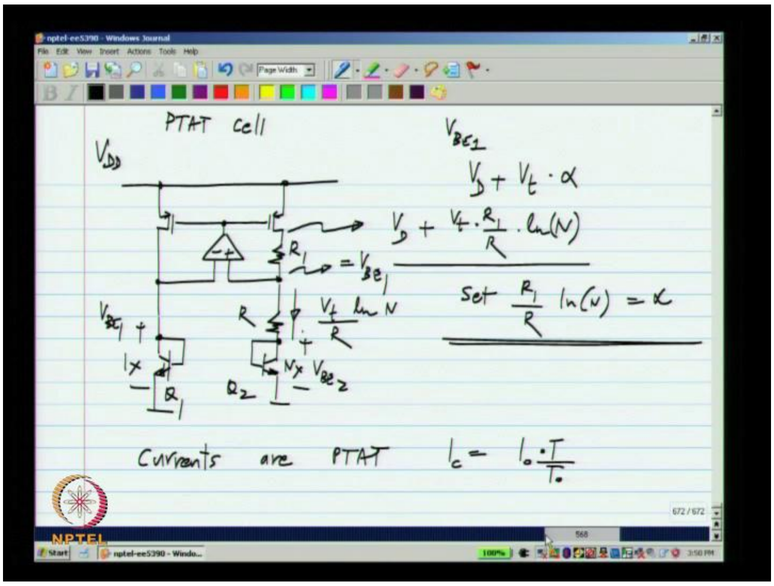
Again as I said, we will show the rough calculation assuming some parameters, but when you realize a band gap circuit, it is based on simulation and you adjust the parameter alpha. So, that the temperature coefficient becomes exactly 0, now you notice that this expression this part of it also appears in the expression for the diode voltage itself, the diode voltage is this much.

So, what we have there is what is inside these brackets, and that is equal to V D minus E g naught by q divided by V t, this is negative by the way this by substituting we will get minus V D minus E g naught by q divided by V t that is this entire thing in brackets and plus 4 plus m which is E g naught by q minus the diode voltage divided V t plus 4 plus m. And if you substitute E g naught by q to be 1.2 volts V D to be 0.7 volts, and V t just for hand calculation I will approximate it 25 milli volts and m we know is minus 3 by 2.

So, what we will get is these differences 500 milli volts divided by 25 milli volts, and 4 plus m is 4 minus 1.5 which is 2.5. So, that comes out to be about 22.5, so if alpha is

22.5 that is you add the diode voltage at a constant current plus 22.5 times the thermal voltage V t, you will get a voltage that is independent of temperature and that voltage will be equal to the extra polited band gap of silicon, which is 1.2 volts. So, will see how to realize this in circuit, we have to refine these results are a little bit we are going to do that. And again this alpha I showed it by hand calculation, but in reality you will just do it by simulation.

(Refer Slide Time: 35:45)



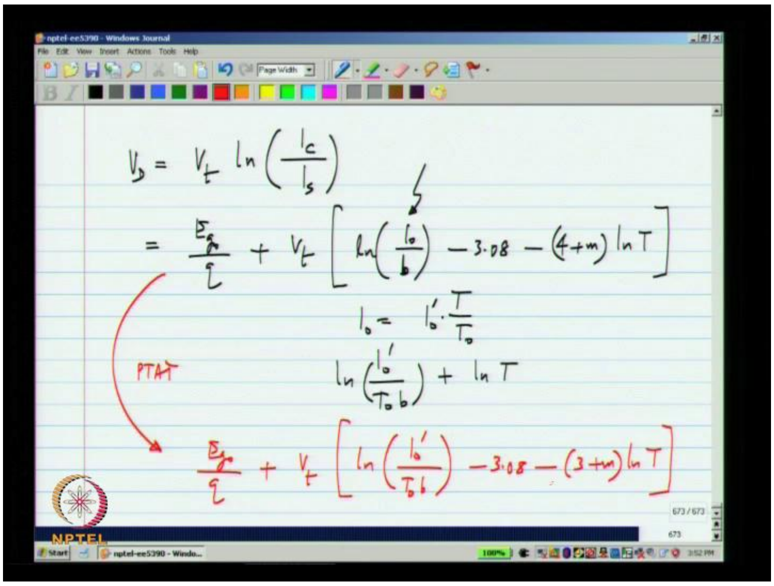
So, how do we go about doing this we know will start with the P TAT cell because, we know that we will need a voltage that is proportional to absolute temperature and I will show bipolar transistors here. So, although I said diode drops and, so on even when you want to diode, it is a better to realize it as a diode connected bipolar transistor, you can consult devices about to say why that is the case. You can have a diode that is single p n junction or you can have a bipolar transistor which is lets a n p n junction, which the collector connected to the base it becomes a diode.

The transistor connected as a diode follows the exponential lot more closely, than the single junction diode. So, again you can consult devices looks to see why this is the case, but that is what is frequently done, so what we have here this is the P TAT cell, we have V B E 1 and V B E 2. And when it is operating correctly this voltage will be equal to V B E 1, and this current will be equal to V t log N divided by R, where this Q 1 has an area of 1 and Q 2 has an area of N.

Now, what we needed was V D plus V t times alpha V D is nothing, but V B E 1, so how do we do that it is extremely simple all I have to do is to add a voltage to this branch, I will call this R 1. So, this voltage will be V D plus V t times R 1 by R log N, so all we have to do is to set R 1 by R log N to the appropriate value alpha, this gives a band gap voltage and that is about 1.2 volts. Now, there is one set will be here earlier while calculating the temperature coefficient of the diode voltage and, so on, I assume that the diode was biased at a constant current that is current add the 0 temperature coefficient.

Now, we have found a simple way of adding the diode voltage or V B E 2 a proportional absolute temperature voltage that is across R 1. But, here in this cell the current through the diode is not temperature independent, but it has a proportional to absolute temperature dependence. Now, this changes things only slightly, so I am not going to go through the details I will just outline the difference, to the current in the cell or proportional to absolute temperature I C will be it will be something proportional to absolute temperature I will call it I naught times T by T naught.

(Refer Slide Time: 39:31)



So, all that happens is that earlier we had V t log I C by I s, we took care of the temperature dependence everywhere except this I C was assumed to be constant. And we when we finally, wrote down the expression, we had got E g naught by q plus V t times log I naught b which was assumed to be constant minus 3.08 minus 4 plus m log t. Now, the only thing that happens is that, this will have a temperature dependence.

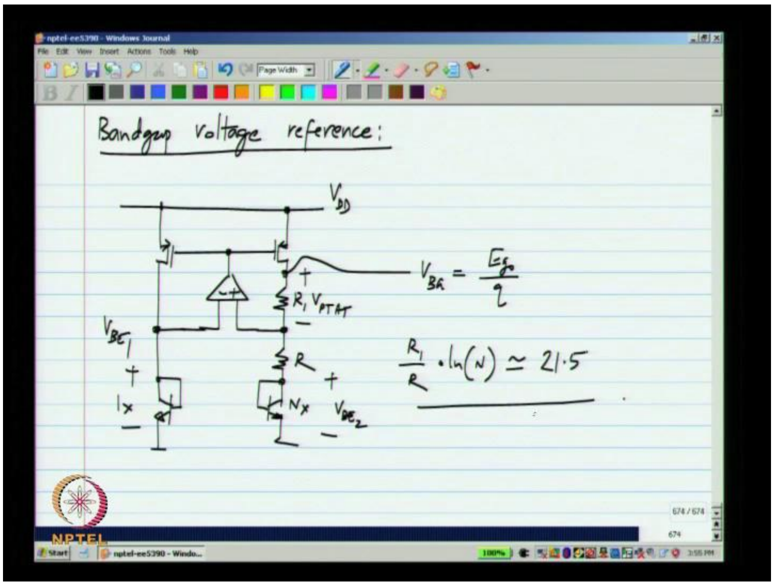
So, like I said if I naught is I naught times T by T naught, so let me naught use the same symbol, some I naught prime time's T by T naught. What we get here would be log I naught prime by T naught times b plus log T, so we will get this extra dependence. So, all that happens is that we have 4 log T here and a log T there, so this will get modified to when the current is P TAT E g naught by q plus V t log I naught prime by T naught b, the exact thing is not important it is a constant minus 3.08 minus 3 plus m log T.

So, wherever you add 4 plus m earlier now we have 3 plus m, so we can again approximately calculate the ratio. So, this becomes 3 plus m, so instead of 22.5 you will have 21.5, again a small difference and like I said earlier, you would calculate this by simulation. This is because, here in the first term I have taken E g naught by q minus V D to be 1.2 volts minus 0.7 volts, E g naught by q is 1.2 volts that is pretty accurate, but V D can be anything it depend on the kind of diodes you have.

So, to get this alpha value exactly you determinate by simulation, and like many of the other reference circuits, this circuit also sensitive to how accurately your transistors are modeled. Now, this is basically the band gap reference circuit, which is used very widely which used in every integrated circuit, sometimes you may have more than one that generates a constant voltage source, it is also commercially available as a standalone reference for many applications.

You will want constant voltage source is everywhere for voltage regulators and many other cases. So, this band gap reference because, it is finally, dependent on this single voltage which is independent of everything right that is it is the extrapolated band gap of silicon, it can be very accurately realize with as sufficiently elaborate circuit, and it is also used extremely widely.

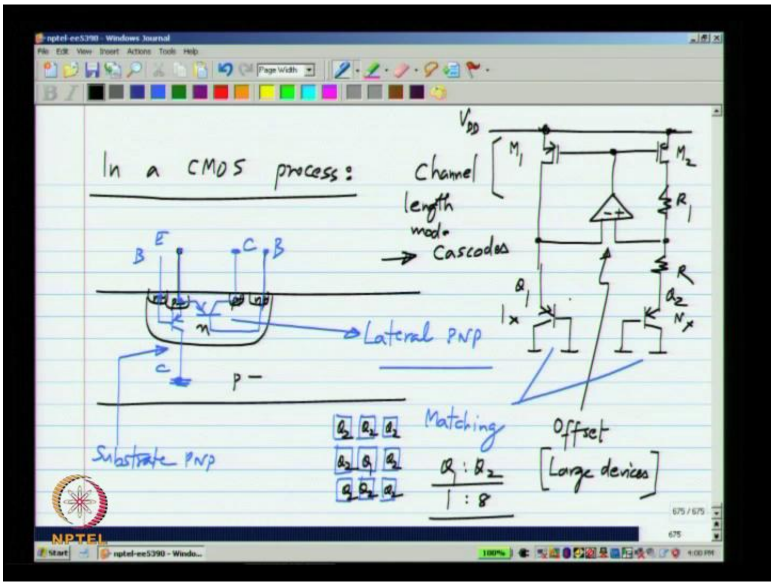
(Refer Slide Time: 43:04)



The very basic band gap voltage reference is what we just outline, we have the proportional to absolute temperature cell, which gives as a current that is proportional to absolute temperature. And in this circuit we also have the diode drop or V B E, so we can stack the two together, this is V B E 1 and here we have V P TAT, so the sum of the two is the output V B G, which should be equal to E g naught by q if you done the cancelation correctly.

Now, there are number of other things, so first off all we have to realize lesson CMOS technologies as well, I have assumed that both bipolar transistors, and CMOS transistors available. If you have a purely bipolar technology, we can do that the OPAMP band the current mirror on top has to be realize choosing bipolar transistors that is e g. But, if you have a purely CMOS technology, you still have to make these diodes or diode connected transistors. And they can be made using the parasitic PNP transistors that are available in any CMOS process.

(Refer Slide Time: 45:11)



You have the n well, in which you would be making the PNP transistor, so and sub state is p as well. So, you see that you have PNP over here and also PNP over there, so the two kinds of PNP bipolar transistors available, one type where this is the emitter because, it is highly doped this is the base and this is the collector. A collector is straight to the substrate, so it has to be connected to ground or the lowest potential, this is the base and the collector, this is known as the substrate PNP.

And you can also use the other set of PNP junctions, in which case all three terminal will be available. This is the annual junction and that is the base, and this is known as the lateral PNP and this also can be used one of the problems is that in this case the PNP transistor symmetrical. The drain and source are the same, so; that means, the collector and emitter are the same this is not a good feature in a bipolar transistor, you would like the emitter to the most highly doped base to have intermediate doping and a collector to be lightly doped.

But, this lateral PNP high symmetrical, but sometimes it can be used, now for the particular case of a band gap circuit, you do not need a transistor with P terminals, you need a diode connected transistor. So, the substrate PNP can be used very effectively, so all you would is that the diodes that you had would be realize like that, the fact that the collector is constraints to be ground is because, here also we need the collector to be grounded.

So, this is all we have to do 1 x and N x, and these two transistor are identical, as usual the sources of error include the channel length modulation of these for which, you will use cascades to overcome if you necessary. If the sensitivity to the supply voltage is to high, then you should use cascades here, here alternatives which one of which we will discuss later. And the other important thing is the offset of this OPAMP, which means that you have to use large devices.

And finally, you also need good matching between these two that is if this device is one that N. And you can use all the other techniques that you know for good matching like common centroid to make sure that they are match and dummy devices and, so on. So, one popular way of laying out transistors for the band gap is to do this, this is that why I call this Q 1 and Q 2, this would be Q 1 and all these would be Q 2. And this will realize a ratio of 1 to 8 in the area.

So, this is how a band gap reference would be realized in a CMOS process, as I mentioned earlier it is used very widely to generate voltage references. This gives you 0 temperature coefficient at a particular temperature, assuming that your model is accurate. Let say you said the temperature coefficient to 0 add 300 Kelvin or room temperature, so at that point it gives you 0 temperature coefficient, but the temperature coefficient will not be 0 at other temperatures, there will be certain variation with temperature.

Now, there are more sophisticated corrections that you can apply, if you know the kind of temperature dependency you have, just like we added a P TAT correction to the diode voltage. We can add higher order corrections, they are known as curvature corrections they are known as curvature connection and the circuits are known as curvature corrected band gaps. So, in the next lecture we will briefly discuss those things, and also another way of combating the errors due to variations in the power supply voltage.

Thank you I will you in the next lecture.