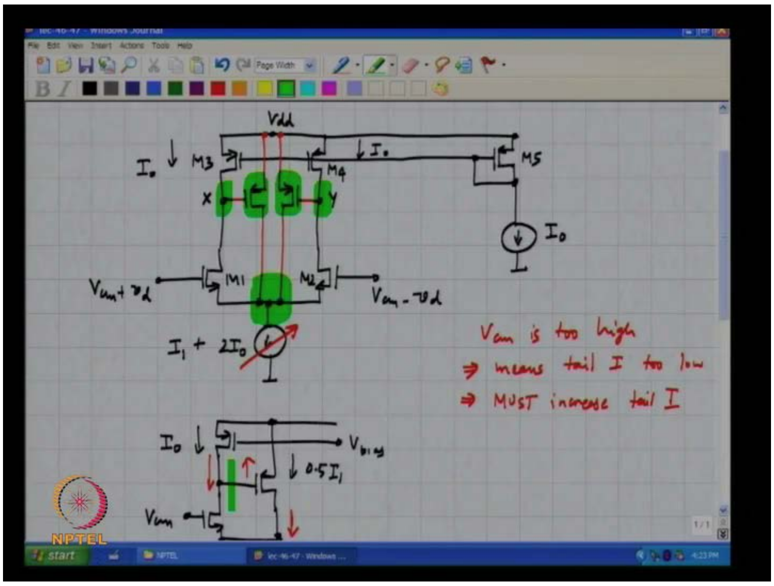
VLSI Data Conversion Circuits
Prof. Shanthi Pavan
Department of Electrical Engineering
Indian Institute of Technology, Madras

Lecture - 46
Autozeroing in flash ADCs

This is VLSI Data Conversion Circuits lecture 46. In the last class we discussed how one could make an amplifier with differential pair, while making sure that the output is fully differential and has got high differential mode gate. And the way we works thing out, it turned out that the output common mode impendence was ideally 0. In practice very small because, you have some kind of negative feedback loop making sure that, the output common mode is held either implicitly or explicitly constant against some reference.

(Refer Slide Time: 01:07)



So, now we will look at another way of implementing that feedback loop, where there is no explicit voltage or current reference. And as I said the number of common mode feedback, loops and circuits you can come up with are virtually infinite and one keep seeing, new flavours of these circuits from time to time, all the examples I am discussing are merely illustrative. So, as we discussed the last time around one could make either the load or the tail currents independent.

So, in this case let me decide to keep the load current independent meaning, it is gate potential is derived from a constant voltage. And constant voltage is simply that developed by a diode of same size carrying a current I naught. So, let us first label these transistor M 1, M 2, M 3, M 4 and M 5, M 3, 4 and 5 are the same size, so this current would nominally be equal to I naught.

So, what one would want to do is to measure the potentials of nodes x and y and tweak the tail current, such that the average potential of nodes x and y is maintained at a constant value. So, if V c m becomes too high what should one do, if V c m is too high it means what is the tail current too low or too high, means tail current too low must increase the tail current is this clear.

So, on a in other words if this potential is too high current must be pushed out of the tail current right, by the same token if V c m is too low in other words with the average potential of, nodes x and y is too low. What it means is that, the tail current is too high and must therefore be reduced. So, one way of doing this is to add two transistors like this all right if this potential, so what is the common mode equivalent.

The common mode equivalent circuit is, so this is some bias, so this is I naught this is V c m, what is and if this is half I 1, this current will be what must be the tail current, I naught plus half I 1. And how do you know there is negative feedback, if you break the loop here, if this node goes up, this node goes what happens to the drain, goes down if this node goes down what happens to this node, this node must go down, so there is indeed negative feedback.

And since this is the common mode half circuit this is common mode negative feedback right, for differential mode signals what is the equivalent. So, this for common mode signals the source couple node is grounded correct and these two transistors, drop out of the picture completely correct, because one is injecting a current in one direction the other one is injecting current the opposite direction, so the two of them cancel; and therefore, do not affect the differential mode loop at all correct. These 2 transistors do not load nodes x and y resistively, because that capacitance which cannot be avoided all right and stabilize the output common mode voltage to what value.

Student: ((Refer Time: 08:32))

V d d minus.

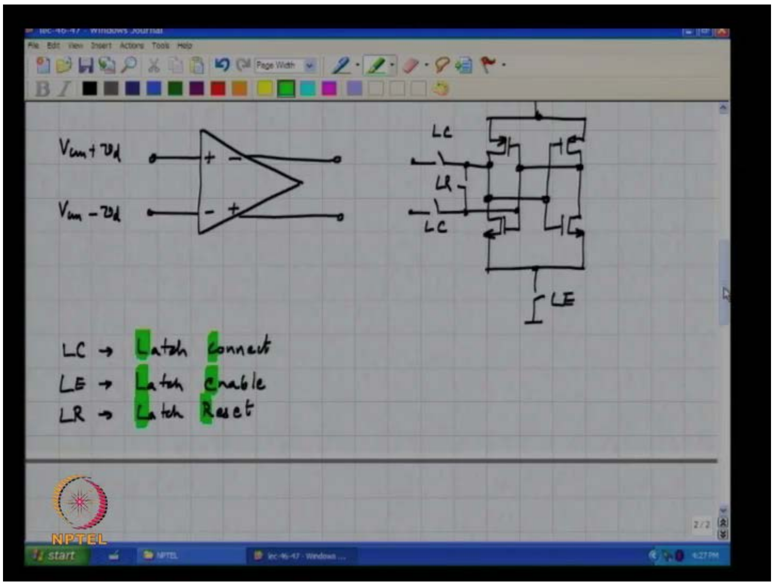
Student: ((Refer Time: 08:36))

of which p MOS transistor.

Student: ((Refer Time: 08:39))

The once which are shaded in green correct, again this is an example of situation where there is no explicit reference, like there is no V c m ref you can actually seen the circuit. The output voltages are converted into common mode components by sensing the voltages and converting them into currents right. And using, the current to correct the error again, there is no explicit current reference either right, everything is implicit in the... So, this is just one more way of doing things used as an illustration all right.

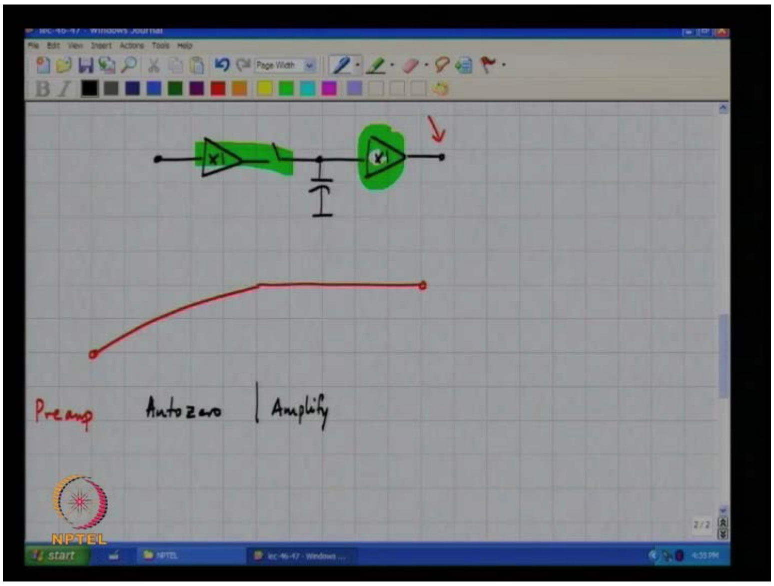
(Refer Slide Time: 09:37)



So, now that we know several ways of building a preamp, what we should do as the following, we need to interface it to our latch, which had two back to back inverters, somehow call these by representative names. So, L C stands for Latch Connect this is the phase during which the latch is connected to the preamp right, L E stands for Latch Enable and this is the phase which enables regeneration. And L C and L E must be complementary not only complementary, but must be non overlapping complementary is not necessary right you can have situations where both of them are low correct.

But, it must be they must definitely be non overlap and between, so after the latch is made its decision before it samples the next input again what have we suppose to do, we need to clear the latch of it is prior memory. So, we had a latch reset, so this L C E and R, so, the preamp on the other hand how will the preamp input look like, the preamp is being driven by sample and hold right or a track and hold for that matter.

(Refer Slide Time: 13:41)



So, what will be the output of the track and hold look like let us take a simple track and hold or argument sake, as we saw right, you need a buffer at the input to make sure that the switch in the capacitor do not load. The circuit which is going to drive the a to d converter, this is the track phase in and when the switch is off the voltage is held on the capacitor, the buffer after the capacitor is necessary to drive the comparator array.

The comparator consist of the preamp and the latch all right, so the continuous time output waveform here will look like this during some part it is tracking and during some other part it is hold, again for simplicity sake I have kept the same time intervals for tracking and for holding this is a reasonable thing. Because, if I you know make the tracking period too small and increase the holding period at it is at the expenses the tracking period, what do I lose and what do I gain, you understand the question what I am saying is, this total period is fixed.

This total period is fixed and why one might argue that I would chose say only 20 percent of this time for tracking and use 80 percent for holding, which is perfectly reasonable thing to do; the question is what am I gaining in this process and which blocks are getting hit by this choice.

Student ((Refer Time: 15:46))

Actually we will state.

Pardon.

Student: ((Refer Time: 15:52))

not really.

Student ((Refer Time: 15:55))

Means last to first.

Pardon.

Student ((Refer Time: 15:57))

Bringing is last to first.

Where.

Student ((Refer Time: 16:00))

When holding.

No.

Student ((Refer Time: 16:05))

If n is more necessary.

The if the hold time is more which blocks benefits.

Student ((Refer Time: 16:15))

Will be.

Comparative.

Student ((Refer Time: 16:18))

No if the hold time is long you will I mean you have lot of time to regenerate.

So, if the hold time is long the comparator is what is benefitting if the hold time is long right, because the comparators working off of the held output, the longer that held output is the I mean the more the comparative benefits, I mean how 1 apportions that hold time within the comparator itself, is you know is another discussion. I mean some of you may argue that I would like to give more time for it to regenerate, some of you may argue that you may be I lead the preamp work for a longer time, but overall a longer hold time benefits the comparator, on the other hand what becomes more difficult if I make the hold time very large.

Student ((Refer Time: 17:24))

The track and hold operates very fast. So, that time constant and.

I mean can you be a little more specific, what do you mean by the track and hold is working at a fixed rate.

Student ((Refer Time: 17:37))

this fixed resistance has to be smaller.

If the if the tracking period is small it means that the tracking bandwidth has to be much higher you understand, so retrieving the track phase it will become very difficult for the circuitry with a very simply ((Refer Time: 18:00)) put a unity gain amplifier and a switch and a capacitor, as well as this unity gain buffer here making the track and; hold track with sufficiently low distortion becomes a challenge when the tracking period is very small.

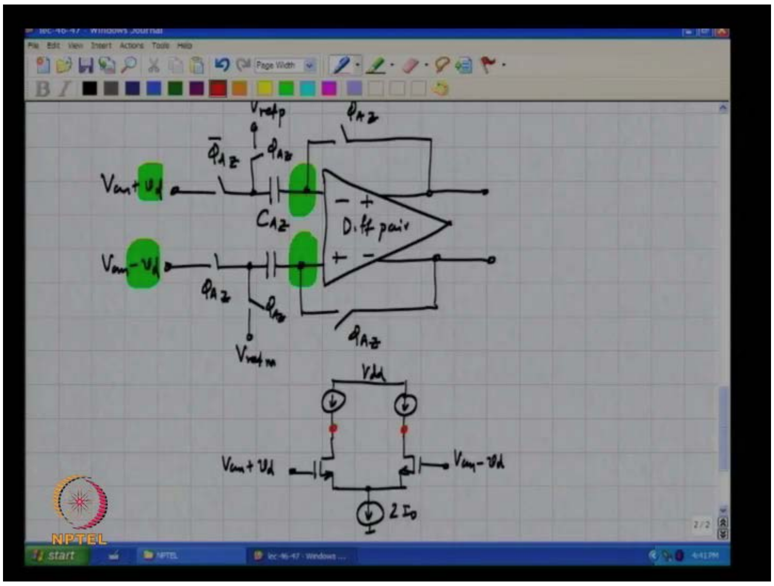
So, the time you apportion the track and the hold periods respectively is a careful trade off between tracking bandwidth and distortion on one hand and comparator gain and degenerate etcetera on the other. So, most often you will end up with a split which is close to fifty, fifty, so that you burden I mean you put the burden on both blocks rather than make one very difficult to do and the one very easy to do is this clear.

So, get me now exaggerate this greatly and then draw way from which looks like this is a track phase, this is the hold phase. So, this is where the preamp should start looking at the output of the sample and hold, is in it. And we also saw that the preamp will automatically be able to subtract references through the out of 0 cycle, I mean the same capacitors, which I used to auto 0 the preamp can also be used to subtract the reference.

So, the preamp during this phase is doing, what is suppose to be doing do we need needed to be doing anything, with regard to the signal at all. The preamp in principle is not needed in this phase, because the real action starts only after the track and hold circuits right or holds. So, if the preamp is not technically needed during this phase one could use it for instance to out of 0 the preamp.

So, 1 possible way of managing time is the preamp is auto zero during this time, then the preamp amplifies during a begins to amplify during this phase, so what do you think the output of the preamp will look like during the hold phase. What is the input to the preamp, the input to the preamp is basically the differential input minus the differential what is the actual input, the input to the preamp here is simply the differential output of the sample and hold, internally if we assume that the coupling.

(Refer Slide Time: 22:38)



The auto 0 capacitors are inside the actual voltage between the two inputs of the differential pair or.

Student ((Refer Time: 22:19))

V s C held signal.

Pardon.

Student ((Refer Time: 22:24))

Held signals.

But, what is the value of the held signal, let me, so this is the differential pair, this is the auto 0 capacitors, this is pi A z I am sorry V ref p V ref M phi A Z. So, when the preamp is auto zeroing, what is the sample and hold doing, it is tracking is this clear. When the preamp is brought out of auto 0 what is happening, what is the differential input between these two nodes.

Student ((Refer Time: 24:42))

Let say v d d v d plus v d plus v f.

Student ((Refer Time: 25:03))

V d that difference between.

No the this is plus v d this is minus v d.

Student ((Refer Time: 25:11))

2 v d..

To the difference between the inputs of the diff pair in the preamp are is basically 2 v d minus V ref p minus V ref M minus whatever offset there is been, which is being stored during the auto zero phase. And the only comparator that is really of interest is the one where the input is closest to one of the reference syllabus, if the input is very far away from one of reference levels we do not need to worry at all, because this difference will then be, so large that the complete tail current of the differential pair will get steered to one of the sides causing it to go very high and the other side to go very low, which the latch will happily resolve because the difference is, so large anyway, so that there is no problem.

So, the only comparators we really need to be bothered about is when are those comparators or that particular comparator which is, whose reference is the closest to the differential input technically for that comparator the difference occurring at, the input of the differential amplifier is very small does make sense. Now, you have a differential pair the input is very small how will the output look like in the time domain.

Let us assume that the some kind of common mode feedback network, forcing the output common mode to be some reference can you comment on the drain potentials of these two transistors.

Student ((Refer Time: 27:49))

Pardon.

But the if the all the transistors have infinite output impedance what happens.

Student ((Refer Time: 28:09))

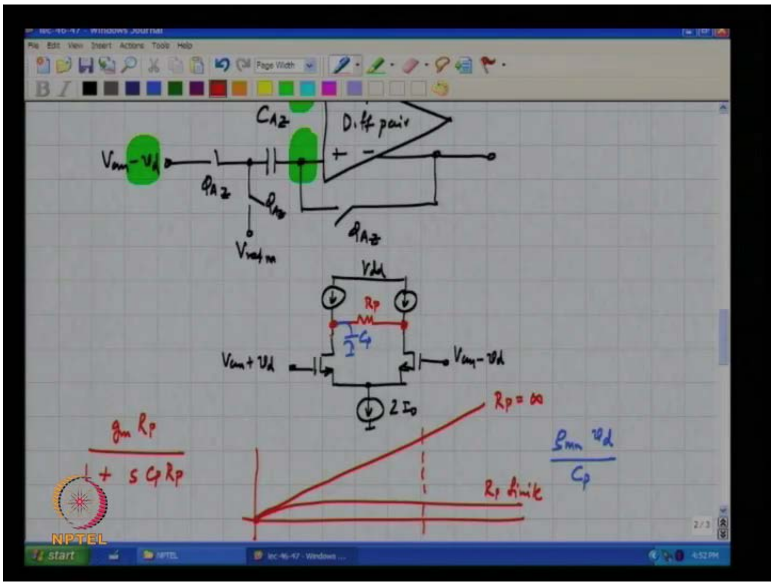
I mean 1 into 2 2 most.

Pardon.

Student ((Refer Time: 28:14))

Usually it be.

(Refer Slide Time: 28:18)



So, the outputs at t equal to 0 or just the beginning of the amplifier phase will both be sitting at some common mode, which is been set by the common mode feedback loop then, what will happen they move in opposite direction they move in opposite directions one will go up and the other 1 will go down, what will be the slope of these waveforms. So, the slope will be g m of the n MOS transistors times v d by whatever parasitic capacitance that is here including loading on the next stage switches etcetera. eventually if you give it enough time this voltage will go and saturate to, whatever some voltage close to v d d other 1 will go low.

But, in the finite amount of time you have most often what will happen is that the output has not reached, sufficiently large value to cause.

Student ((Refer Time: 29:57))

The latching will.

The no not latching, the output to cause the current source or the input devices to get into the tryout regions only, if you give it enough time that will happen. But, in the beginning they will ramp up simply, because constant current is being pushed into a. into a capacitor. So, the time constant technically is infinite, because the R is infinity and you have some C p.

So, the question before I put to you now is does it makes sense to leave the time constant between infinity or should I make some efforts to reduce the time constant. So, this is the time I have for amplify for instance, a question I am asking is does it makes sense for me realising, the fact that the time constant is very large to deliberately reduce the time constant.

Student ((Refer Time: 31:27))

No now if you reduce the time constant the effect to high frequency being its I mean its low for this amplifier right.

I mean how do I mean that depends on I guess, from question I would ask is then how do you propose to reduce the time constant in the first place.

Student ((Refer Time: 31:50))

Then if you do not have gain then how will the offset of the next stage in gate I mean purpose of this amplifier is to cut off the.

Correct

Student ((Refer Time: 31:57))

Remove the offset of the next latch the regenerative back.

Correct.

Student ((Refer Time: 32:00))

If we if the time constant of this is too high.

Yes.

Student ((Refer Time: 32:0.3))

Then your high frequency gain is low.

The

Student ((Refer Time: 32:06))

I mean if your capacitance is too high or your g M is too low.

No, g M is something time constant here is simply, the resistance across the parasitic capacitance which is infinite times C p which is infinite correct.

Student ((Refer Time: 32:26))

But the gain is proportional to g m.

The..

Student ((Refer Time: 32:28))

How much the output will strain that is proportional to g M and inversely proportional to C p.

Correct.

Student ((Refer Time: 32:34))

If you try to reduce power at the states the only way you can.

No, I am saying I am trying to reduce the I mean we all agree that this output waveform will not settle correct unless given infinite time, assuming none of the devices go and you try out and. So, on. The question I am asking therefore, is that is this ok when normally we expect to see things to be you know to be settle, before we do anything is in it. So, the question is this ok or should we make any deliberate attempt to try and reduce, the time constant and one way of reducing time constant would be to simply add a resistor in parallel.

So, can I have any comments on that.

Student ((Refer Time: 33:25))

Sir ideally resistor will be amplitude cause the that to interface to mutually go down,, but if you reduce I mean if you could somehow reduce capacitor that would.

No, they I mean we cannot play with capacitance because it is a fundamental thing.

Student ((Refer Time: 33:38))

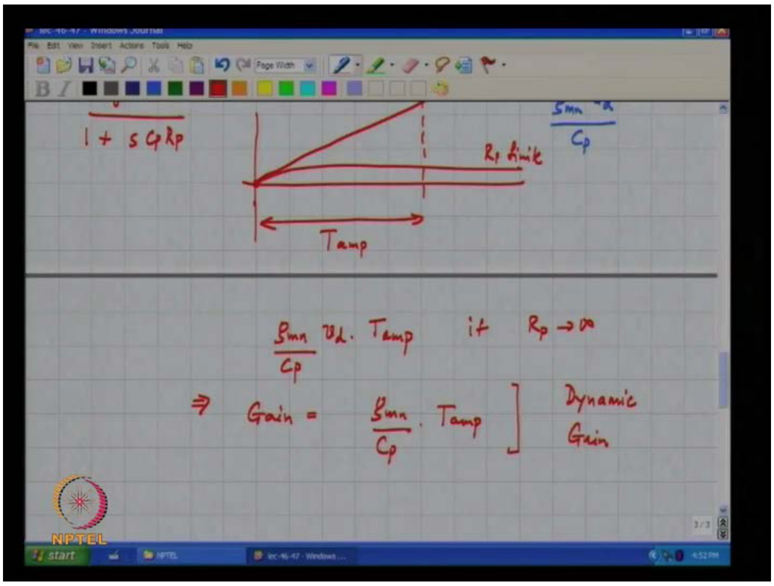
So, if resistor is non-linear then adding resistor in parallel would not help, because the output is will be even further lower so offset reduces.

So, well while couple of things to observe one is that the input to this differential pair within this time period is d C correct, it is a constant voltage. Now, if one says no I want the output to settle before I can do anything, the only way to do that is to deliberately put a small resistance in parallel with C p at the output perhaps like this, this will make sure that the output settles however the gain will become lower.

So, the gain eventual gain is g M times R p whereas, the pole at the output node will be of the form 1 plus s C p times R p. So, if 1 is interested in the voltage developed across the outputs of the preamp, after a fixed amount of time which is a time we have for amplification, then if you chose R p to be very small it will settle very quickly. But, the gain is too small on the other hand as you go on increasing R p what happens, the settling time the time constant tends to infinity.

But, the output of the preamp will keep doing will keep going up as a ramp, if we gave it enough time it would go to infinity right, but we have a finite amount of time. So, after a certain amount of time let me just draw this diagram like this, so after a for a certain of time even dough the output waveform does not settle with a large R p we see that the differential output developed when R p is infinite is much larger than what would be developed the if R p was finite, in other words this preamplifier for all practical purposes is behaving like a , the input is at d C is constant and the output is a ramp is behaving like a integrator.

(Refer Slide Time: 36:56)



So, the output differential output voltage developed, after it certain time T amp is simply g M n by C p times v d times T amp, if R p tends to infinity. So, the Gain is nothing but g m n by Cp times T amp, so this is not gain in the conventional sense of the word. The d C gain of this amplifier is infinite, but if you have a certain time t amp, the output voltage is only multiplied by a finite factor, clearly that finite factor depends on the amount of time you have.

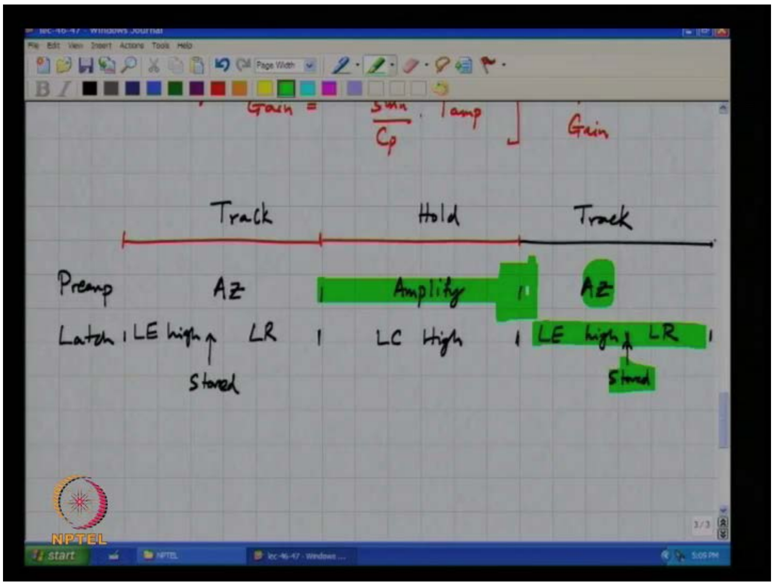
So, this is called this is gain, but it is not the static gain this is called the dynamic gain of the preamp, as you give more time the output of the preamp will become larger and this helps the latch, because the voltage it has to resolve, becomes I mean it has become larger which makes it much easier to resolve you understand. So, this family of preamps where one deliberately you know has high output impedance and increases the time constant, I mean why I specifically mentioning time constant is that one usually thinks of fast circuits is having time constants which are very small.

But, here one see is that I mean making deliberately trying to make the time constant small in order to make thing settle is actually making things worse, you understand while it certainly settles, we see that the output voltage developed is not as large as it would be if R p was not around at all. So, you indeed want to make the preamp, so called integrating preamp, where they are differential output simply keeps ramping up.

Now, of course, you cannot have the entire you know presumably you cannot have the entire amplify phase, for the Preamp they you would like to give as much as possible. So, let us see what all needs to get done, before we get ready for the next sample. So, the preamp must amplify, the maximum amplification time as simply this much correct.

So, this is the track and this is the hold phase, the maximum time you can have for amplification is the same as the hold phase of the sample and hold.

(Refer Slide Time: 41:44)



Now, what must the latch do the latch must sample at the.

((Refer Time: 41:33))

Let me, we go to the next page, so this is Track, This is hold preamp is amplify during this phase what is the Latch doing,

Student ((Refer Time: 42:18))

Latch you can come on a latch best observe in that case.

No, after the hold phase is over. So, the lamp the Latch must sample, the output of the must be connected to the output of the Preamp, when the preamp is amplifying correct. So, L C must be high during this phase, as soon as the amplification is done the.

Student ((Refer Time: 42:50))

The latch must be ah.

Latch must, so latch enable, so let me call this is the next track phase, so L E must be high, during the next track phase see during the track phase, the comparator is basically not doing anything. So, it can basically use that time to advantage, so that it gets as large a total time is possible. So, the if L C is high during the hold phase and the L E is high. During the next track phase and what should we do next.

Student ((Refer Time: 44:01))

And then there you have to short.

So, then after L E is been hyper sometime one needs to one needs to latch store the decisions and this is simply a delayed version of, when do you store he output of the latch.

Student ((Refer Time: 44:32))

After it has been.

After it is regenerated to a sufficient degree, so that will simply be a delayed version of ((Refer Time: 44:43)) and then at this point you can, so L E at this point the decision of the Latch stored as soon as the decision is stored one can ((Refer Time: 45:08)) what do you do next, you reset the Latch and then you get back to Latch connect you understand. So, in steady state of course, what will happen, is that what happens in this phase will wrap around.

So, I am going to say L E high latch output is stored, then you have latch reset correct, so what do you think will happen to the preamp, now after the preamp is amplified what should 1 do.

Student ((Refer Time: 46:05))

Looking out of the amplification.

In principle 1 could say I will auto zero during the track phase, in this track phase is the same, so this is what you can do you understand. So, this is a as I said a very rudimentary way of doing things one point to notice is that, the preamp at the end of the hold phase has its output parasitic capacitors charged to v to v d minus V ref p minus V ref n multiplied by the gain of the, dynamic gain of the Preamp.

Now, one when one these act as initial conditions, before for the auto zero phase, so these two capacitors which were charged to some very different voltage have to eventually come to, I mean during the auto zero phase what happens.

Student ((Refer Time: 47:26))

I mean essentially they come to a form of a phase plus mod offset.

So, the whatever they were before the track phase at the end of the track phase, which was also the proposed auto zero phase for the preamp, they must come down to the output of the preamp at the end of the auto zero phase, will be approximately the offset. So, this I mean the two initial conditions on the output will act as you know, as potentials which I mean please realise the that difference could be very large, especially if the comparator is very far away from the references.

So, there are some I mean as far as gain is concerned you really worried about that comparator which is very close to the reference, as far as this effect is concerned where the preamp is suppose to come back to v off set at the end of the track phase, we are worried about the comparator, which is very far away from the reference, because that is when the difference at the output of the preamps is the largest, in which case you know you make sure that the auto zero loop is sufficiently fast.

So, that it can settle within that track instant that is one point and please note that the auto zero, during the auto zero phase the effective capacitance at the that the preamp has to drive is not the parasitic, but C a zee, during the auto zero phase this is shorted. So, C a zee has to be driven by the preamp, we just which is difficult to do given the high speeds of operation you understand.

So, can you think of you know I mean can you do you have any comment on how I can improve the situation.

Student ((Refer Time: 50:00))

could have no that will not let you get the outputs seen too much.

Well, but then that will reduce the dynamic gain is in it.

Student ((Refer Time: 50:20))

No, then we can put a switch I mean after the I mean after the hold phase when we know that the offset is voltage is small and both of them are going to be close. So, put a switch,, but we have to.

So, one it is very good, so one approach which you want to use to reduce memory from cycle to cycle is to put a you know, that the at the end of the auto zero phase anyway the two outputs of the preamp will be very close to each other. So, one might insert a small reset phase after the amplifier, where the two outputs are momentarily zap together, this way two of them will start off with a differential output voltage of 0 and we will only need to settle to V offset, yet another thing to observe is to note that there is really no need to do this auto zeroing every cycle.

Because, the offset remains constant and the references also remain constant and there is no point in refreshing this auto zero capacitor every cycle, if you now say that let us say I decide to do this only once in, so many cycles where this many cycles is a function of how big an auto zero capacitor, you use how much leakage you have and. So, on that is a big advantage as you can see, because the time which is used for auto zeroing can now be used for.

Student ((Refer Time: 52:10))

One of the t m.

Student ((Refer Time: 52:15))

When only being on amplifying and.

When you can opportunate to do other things like amplification and I mean and more importantly you do not need to I mean, if you want the preamp to settle to the offset in this track phase, you have to dry a drive that C auto zero. During the auto 0 phase which means that the g M of the preamp must be very large. Because, the time constant is finally, the output impedance of the preamp when during the auto zero phase the input and output of the preamp are shorted together, which means that the output impedance is nothing, but g M n that is driving a capacitance C a zee.

So, the time constant is nothing but one over, I mean the time constant is C by G m, so if you want it to settle quickly right you want the g M to be very large. Because, C a zee has to be large, because you wanted to be much larger than the input parasitic capacitance of the preamp. So, unnecessarily you will be burning a lot of power and. So, on. So, in the next class we will see strategies which can, where we can get rid of the auto zero phase every cycle.

Because, enough point in writing the same information onto those capacitors every cycle and burning, a lot of power to do it and you spending a whole half cycle on this you understand I mean, but you must still understand that the amplifier phase cannot be larger than the hold phase, but at least we do not have to sit and settle to, the offset in that small half cycle time and as the speeds you the flash convertor, become higher and higher the time we have for tracks slash hold all start to become Smaller and smaller.

So, it becomes more and more power inefficient to try and I mean to settle to an adequate accuracy within this track phase, becomes a charge. Please note that unlike in the amplifier phase where we do not need settling, they are only interested getting an gained up version of the differential input during the amplifier phase. The more time you have the more gain you get in the track phase however, or during the auto zero phase you want the outputs to settle.

Because, that is what is going to be stored on your auto zero capacitors correct, so that becomes a problem in the next class, we will see how what we can do to make sure that if we postulate that the auto zero can be done once in very long while. Which in some systems is possible for example, it turns out that in disk drives which are one big user of high speed flash convertors, there is some interval during which the a d C is not being used at all and that happens once in you know, so many 1000 cycle.

So, if you know that you’re not going to be using the a d C during this time you can use that time interval to do lot of housekeeping, so it turns out that many practical convertors will simply auto zeroes all there comparators during, that time and they are already for action and therefore, during the next burst of data. Which can last for you know as I said several 1000 of cycle there is no auto zeroing at all.

Because the references and the offset have been stored already during that long interval similarly things, which where you need to digitize video in some systems there is a blanking interval. So, you scan the line and you come back right and during that blanking interval you do not need things to be functional, so during that time you can do house keeping all right for instance cancel offsets do all sorts of stuffs. So, the next class we will see how we can actually exploit these to simplify designing all right.