

HIGH SPEED DATA CONVERSION

By Mike Koen (602) 746-7337

INTRODUCTION

The design considerations for high-speed data conversion are, in many ways, similar to those for data conversion in general. High speed circuits may sometimes seem different because device types can be limited and only certain design techniques and architectures can be used with success. But the basics are the same. High speed circuits or systems are really those that tend to press the limits of state-of-the-art dynamic performance.

This bulletin focuses on the more fundamental building blocks such as op amps, sample/holds, digital to analog and analog to digital converters (DACs and ADCs). It concludes with test techniques. Op amps, which tend to be the basic building blocks of systems, are considered first. Sample/holds which play an important role in data conversion are considered next followed by DACs and finally ADCs. ADCs are really a combination of the other three circuits. Emphasis is given to hybrid and monolithic design techniques since, in practice, the highest levels of performance are achieved using these processes. The material is presented from a design perspective. Theory and practical examples are offered so both the data conversion component designer and user will find the material useful. The concepts presented do not require extensive experience with data conversion. Fundamental concepts are discussed allowing the subject to be understood easily. The material emphasizes high speed circuit considerations—circuit theory is not treated in depth.

Topics Covered in this Bulletin

- A. Amplifier Architectures
 - 1. Buffer
 - 2. Operational
 - 3. Open Loop
 - 4. Comparator
- B. Amplifier Applications
 - 1. Sample/Hold
 - 2. Peak Detector
- C. Digital to Analog Converters
 - 1. Bipolar
 - 2. Deglitched DAC
- D. Analog to Digital Converters
 - 1. Successive Approximation
 - 2. Flash
 - 3. Sub-ranging
- E. Test Techniques
 - 1. Settling Time
 - 2. Aperture Jitter
 - 3. Beat Frequency Testing
 - 4. Servo Loop Test

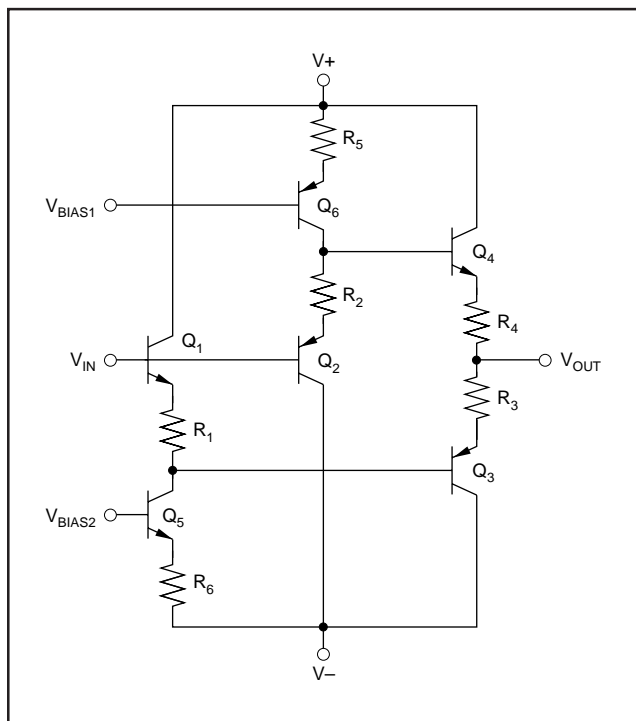


FIGURE 1. High Speed Bipolar Buffer.

AMPLIFIER ARCHITECTURES

Amplifiers of all types play an important role in data conversion systems. Since high speed amplifiers are both useful and difficult to design, an understanding of their operation is important. Four different types of amplifier architectures will be discussed. Buffers, op amps, open loop amplifiers, and comparators can be found in just about any signal processing application.

THE BUFFER

The open loop buffer is the ubiquitous modern form of the emitter follower. This circuit is popular because it is simple, low cost, wide band, and easy to apply. The open loop buffer is important in high speed systems. It serves the same purpose as the voltage follower in lower speed systems. It is often used as the output stage of wideband op amps and other types of broadband amplifiers. Consider the two buffer circuit diagrams, Figures 1 and 2. The output impedance of each buffer is about 5Ω and bandwidths of several hundred megahertz can be achieved. The FET buffer is usually implemented in hybrid form as very wideband FETs and transistors are usually not available on the same monolithic process. The all-bipolar form of the buffer is capable of

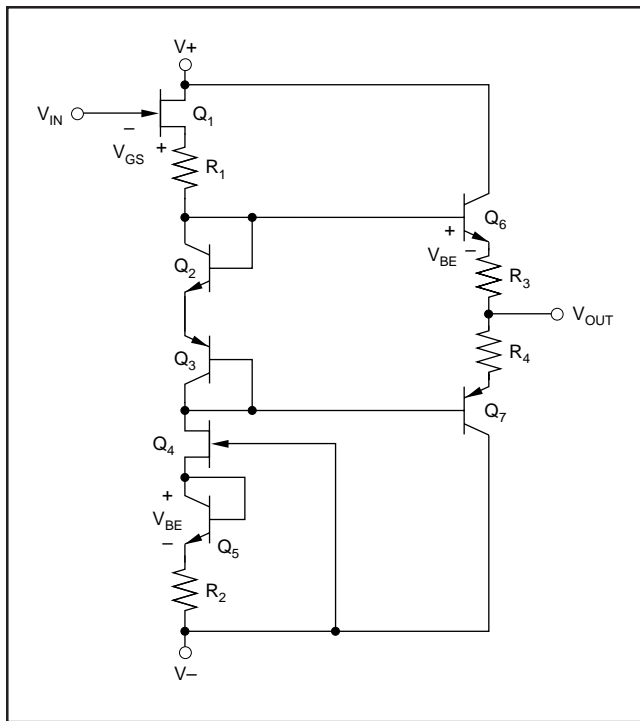


FIGURE 2. High Speed FET Buffer.

being produced on a complementary monolithic process where both the NPN and PNP transistors are high performance vertical structures. Figure 1 shows the buffer in its most basic form. The input to the buffer is connected to a pair of complementary transistors. Each transistor is biased by a separate current source. The input transistors Q_1 and Q_2 through resistors R_1 and R_2 are connected to the bases of output transistors Q_3 and Q_4 so that offset will be zero if the base to emitter voltage of the NPN and PNP are equal. Zero offset requires that transistor geometries are designed for equal V_{BE} s at the same bias current—achievable in a complementary process. This circuit is very useful as it has a moderately high input impedance and the ability to supply high current outputs. One important use of this buffer circuit is to amplify the output current of a monolithic op amp. Monolithic op amps usually do not have output currents that exceed 10mA to 50mA, while the buffer shown in Figure 1 is capable of putting out more than 100mA. Typically this type of a buffer has a bandwidth of 250MHz, allowing it to be used in the feedback loop of most monolithic op amps with minimal effect on stability. Figure 3 shows how the loop is closed around the buffer so that the DC performance of the amplifier is determined by the unbuffered amplifier and not the output buffer. An advantage of the connection shown in Figure 3 is that load-driving heat dissipating is in the buffer so that thermally induced distortion and offset drift is removed from the sensitive input op amp.

Figure 2 shows the FET version of the previously mentioned circuit. The FET buffer achieves zero offset by the mirror action of the NPN transistor Q_5 that is reflected as the gate

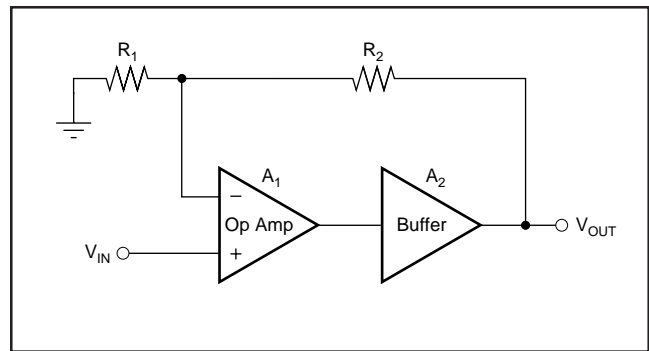


FIGURE 3. High Current Op Amp.

to source voltage of the input FET Q_1 . The V_{BE} of Q_5 determines the gate to source voltage of the FET current source Q_4 . Since the identical current flows in Q_4 and Q_1 the gate to source voltage of Q_1 will also be equal to V_{BE} . Since Q_5 and Q_6 are identical transistors the offset of the FET buffer circuit will be nominally zero. The circuit shown in Figure 2 is usually constructed in hybrid form so that it is usually necessary to adjust resistors R_1 and R_2 to set the offset of this circuit to zero. Setting the offset to zero is accomplished by laser-trimming resistors R_1 and R_2 with the buffer under power. (This is known as active trimming.) A common application of this circuit is to buffer the hold capacitor in a sample/hold. (See the section on sample/holds.) The high impedance of the FET buffer allows the capacitor to retain the sample voltage for a comparatively long time as the room temperature input current of a typical FET is in the vicinity of 50pA.

Another common application of either type of buffer circuit is to drive high capacitive loads without reducing the overall system bandwidth. Op amps, even though they have closed loop output impedances that are very low, can become unstable in the presence of high capacitive loads. The open loop buffer is usually more stable when driving capacitive loads, but this circuit will also develop a tendency to ring if the capacitive load becomes excessive. Figure 4 shows how the emitter follower can oscillate due to reactive output impedance. Figures 5 through 7 show calculated results for different conditions when a simple emitter follower is driving a capacitive load which illustrates this oscillatory tendency.

One very important application of the open loop buffer is to drive a “back matched” transmission cable. Back matching a cable is just as effective in preventing reflections as the more conventional method of terminating the cable at the receiving end. The advantage of the back matched cable is that the generating circuit does not have to supply steady-state current and there is no loss of accuracy due to the temperature dependent copper loss of the cable. Figure 8 shows circuit diagrams and explanations that describe the operation of the open loop buffer driving a “back matched” cable.

THE OPERATIONAL AMPLIFIER

Several examples will be shown that depict the architecture of wideband op amps. These amplifiers have settling times to $\pm 0.01\%$ in under 100ns and closed loop bandwidths in excess of 100MHz. The question is often asked, "How much loop gain is enough?" Wideband amplifiers generally do not achieve as much open loop gain as lower frequency amplifiers. This is the result of optimization of bandwidth and

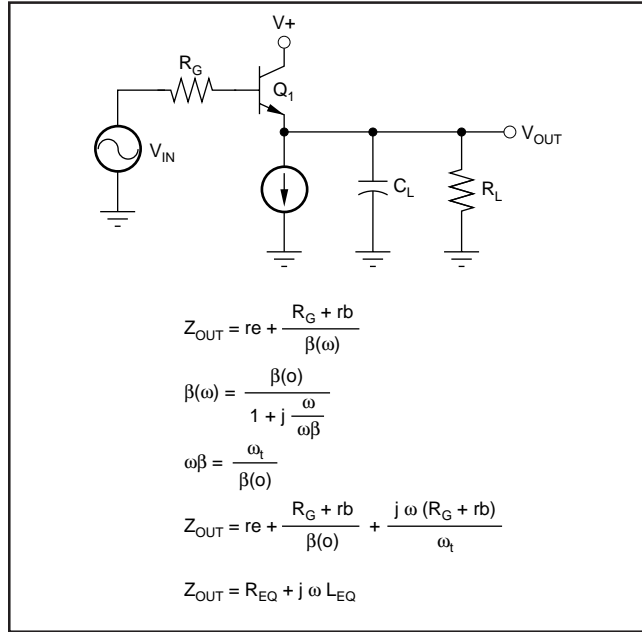


FIGURE 4. Output Impedance of Emitter Follower.

$$\frac{V_{OUT}}{V_{IN}} = 1 - \left[\frac{k}{(1 - k^2)^{1/2}} \sin 2\pi (1 - k^2)^{1/2} \frac{t}{T} + \cos 2\pi (1 - k^2)^{1/2} \frac{t}{T} \right] e^{-2\pi k(t/T)}$$

where: $T = 2\pi (L_{EQ} \cdot C_L)^{1/2}$

$$k = \left[\frac{R_{EQ}}{L_{EQ}} + \frac{1}{R_L C_L} \right] \frac{T}{4\pi}$$

$$R_{EQ} = r_e + \frac{R_G + r_b}{\beta(0)}$$

$$L_{EQ} = \frac{R_G + r_b}{\omega T}$$

FIGURE 5. Time Response.

$f_T = 1\text{GHz}$	$f_T = 5\text{GHz}$	$f_T = 5\text{GHz}$
$R_G = 50\Omega$	$R_G = 50\Omega$	$R_G = 50\Omega$
$r_b = 50$	$r_b = 50$	$r_b = 50$
$r_e = 5$	$r_e = 5$	$r_e = 5$
$C_L = 50\text{pF}$	$C_L = 50\text{pF}$	$C_L = 50\text{pF}$
$\beta(0) = 100$	$\beta(0) = 100$	$\beta(0) = 100$
$k = 0.35$	$k = 0.44$	$k = 0.51$
$T = 5.6\text{ns}$	$T = 4.7\text{ns}$	$T = 1.9\text{ns}$

FIGURE 6. Different Conditions.

phase margin. If open loop gain is stable over temperature and linearity with signal adequate, the requirement for high open loop gain is reduced. This is important since it is difficult to achieve high open loop gain for wideband amplifiers.

There are several ways to shape the open-loop-gain/phase characteristics, or Bode Plot, of an amplifier. The method chosen depends on whether high slew rate or fast settling is to be emphasized. The methods of stabilizing the closed-loop gain of these amplifiers will also result in different settling time characteristics. The benefits of each of these methods will be explained. The first amplifier has a FET input and the other has a bipolar input. High speed amplifiers should be designed to drive 50Ω loads to be most useful. 50Ω cable is commonly used in high speed systems to interconnect signals.

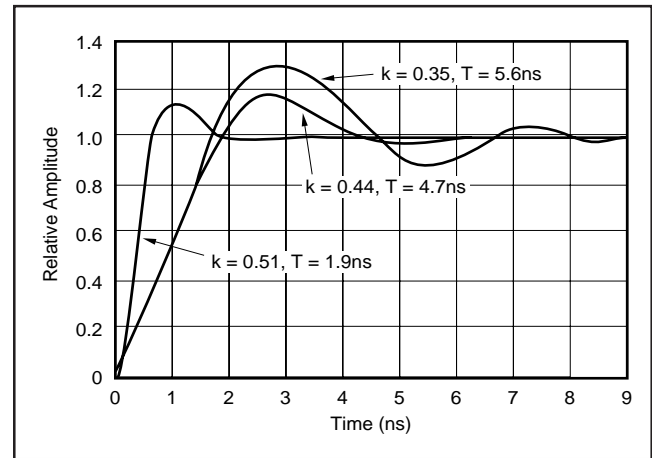


FIGURE 7. Results.

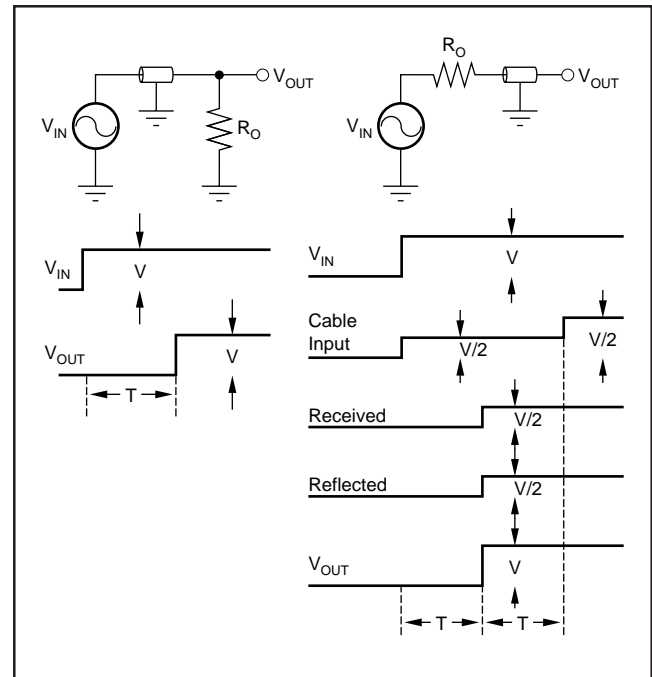


FIGURE 8. Back Matched Cable.

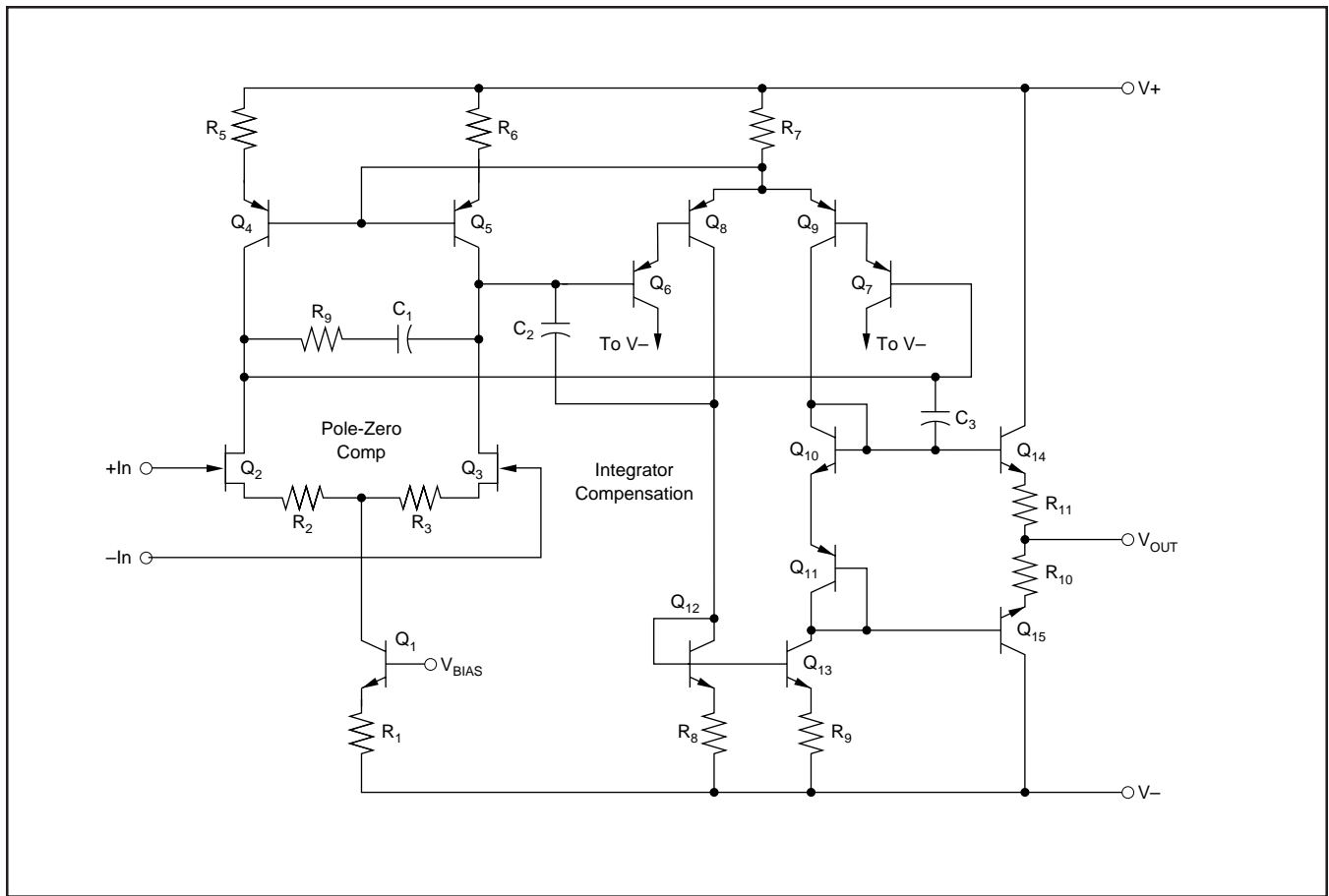


FIGURE 9. FET Operational Amplifier.

Consider a classic two stage hybrid amplifier as shown in Figure 9. It can be compensated either with integrator feedback or with pole-zero compensation. Hybrid amplifiers can achieve the highest possible dynamic performance because optimum input and output devices that can be used from widely differing technologies. Very often it is possible to achieve the combination of bandwidth, breakdown voltage, and current levels needed only with hybrid techniques. It is instructive to analyze the performance of this amplifier in detail as a way of demonstrating many pertinent considerations for a high speed amplifier. High speed amplifiers may be configured in other ways but the major design considerations are the same. FET input amplifiers are very useful as their high input impedance serves to buffer the hold capacitor in sample and hold circuits. Additionally, a FET can tolerate much larger differential input voltages during overload conditions than bipolar input stages and there is no error due to input current.

The input stage of the amplifier shown in Figure 9, draws 5mA per side and at 25°C the input current is typically 25pA. A bipolar input stage being operated at the same current would have an input current of approximately 50μA, which when transformed by the feedback resistor, would be an additional source of offset error and noise. To compensate for the low gain of the input stage ($G = 25$) it is desirable

to maintain a differential connection between the first and second stages. When a connection of this type is made it is necessary to establish the operating point of the input stage using “common mode” feedback. Assuming that FET pair Q_2 and Q_3 are well-matched, the current is split evenly and emerges as equal collector current for transistors Q_4 and Q_5 . The bases of Q_4 and Q_5 are connected together and applied to the common connection of the emitters of PNP transistors Q_8 and Q_9 . Therefore, in order to establish balance in the loop, a voltage is created across R_7 of such a magnitude to allow the current in transistors and Q_4 to be a value that will exactly balance the current needed by FETs Q_2 and Q_3 . Transistors Q_8 and Q_9 are driven from a pair of emitter followers to increase the overall loop gain. Emitter follower transistors Q_6 and Q_7 increase the gain of the first stage by preventing transistors Q_8 and Q_9 from loading the drains of the input FET pair. The differential output of transistors Q_8 and Q_9 are then connected to the output emitter followers directly and through the mirroring action of transistors Q_{12} and Q_{13} . The overall DC gain of this amplifier is 94dB. The current through the output emitter follower is established by the biasing action of the diode connected transistors Q_{10} and Q_{11} . The offset voltage of this amplifier is trimmed to under 1mV and the amplifier has a voltage offset drift coefficient of less than 10μV/°C.

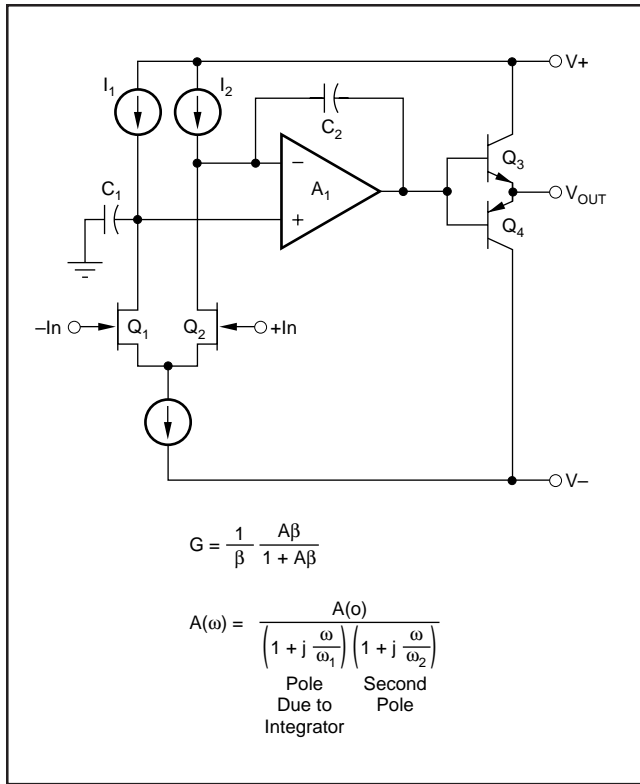


FIGURE 10. Integrator Compensation.

As previously mentioned, there are two methods for compensating the open loop frequency response of this amplifier. The first method to be discussed is called integrator feedback as a capacitor is connected from the output stage to the drain of the input stage. Figure 10 shows a block diagram of this connection which more clearly demonstrates why it is called integrator compensation as an integrator is formed around the output gain stage of the amplifier. The advantage of integrator feedback is that the closed loop frequency response has all the poles in the denominator which means that the transient response is tolerant to parameter variation. As will be shown, another type of frequency compensation is called “doublet” or “pole-zero cancellation” which can have poor transient response due to small parameter variations. Another benefit of integrator feedback is lower noise output as the integrator forms an output filter as contrasted to pole-zero cancellation which only forms an incomplete filter of the input stage. Figures 11 and 12 show the relationship between the frequency and time or transient response of a feedback amplifier that employs integrator feedback.

Figures 13 through 16 illustrate the effect of a pole-zero mismatch. A pole-zero mismatch creates a “tail” or a long time constant settling term in the transient response. Pole-zero compensation is not as effective as integrator feedback in stabilizing an amplifier but should be considered as there are times when the integrator itself can become unstable. Pole-zero compensated amplifiers often have higher slew rates.

$$G = \frac{1}{\beta} \frac{1}{1 - \frac{\omega^2}{A(0)\beta\omega_1\omega_2} + j\frac{\omega}{A(0)\beta}\left(\frac{1}{\omega_1} + \frac{1}{\omega_2}\right)}$$

$$G = \frac{1}{\beta} \frac{1}{1 - \frac{\omega^2}{\omega_n^2} + 2\zeta\left(\frac{\omega}{\omega_n}\right)}$$

where $\omega_n = \sqrt{A(0)\beta\omega_1\omega_2}$

$$\zeta = \frac{\omega_1 + \omega_2}{2\sqrt{A(0)\beta\omega_1\omega_2}}$$

Step Response:

$$e_o(t) = \frac{1}{\beta} e_i(t) \left[1 - \frac{t - \zeta\omega_n t}{\sqrt{1 - \zeta^2}} \sin(\omega_n \sqrt{1 - \zeta^2} t + \cos^{-1} \zeta) \right]$$

FIGURE 11. Transient Response Integrator Compensation.

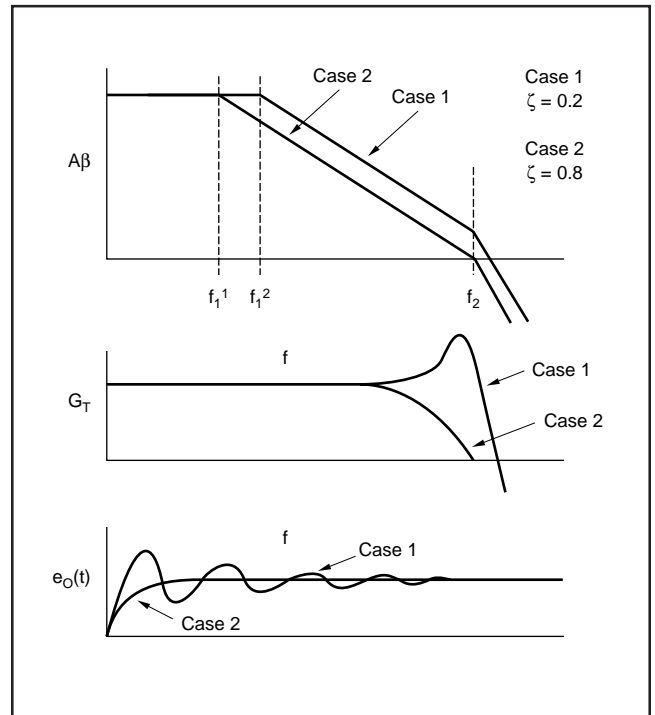


FIGURE 12. Open Loop Gain, Closed Loop Gain, and Transient Response Integrator Compensation.

The second architecture that will be discussed is known as the folded cascode operational amplifier. This circuit arrangement is very useful as all the open loop gain is achieved in a single stage. Since all of the gain is developed in a single stage, higher usable gain bandwidth product will result as the Bode Plot will tend to look more like a single pole response which implies greater stability.

Figure 17 shows a simplified schematic of this type of amplifier. The input terminals of this amplifier are the bases of transistors Q1 and Q2. The output of transistors Q1 and Q2 are taken from their respective collectors and applied to the emitters of the common base PNPs Q4 and Q5. Transistors Q4 and Q5 act as cascode devices reducing the impedance at

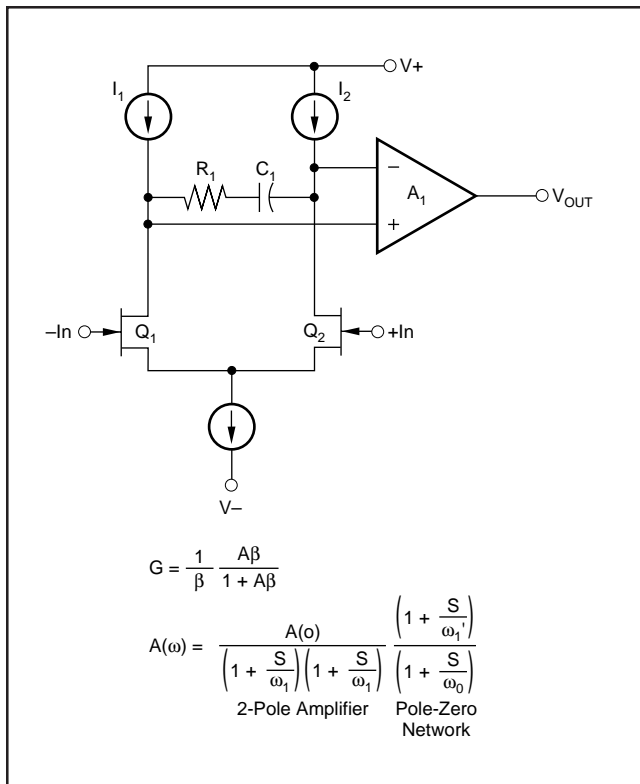


FIGURE 13. Pole-Zero Compensation in Op Amp.

$$G = \frac{1}{\beta} \frac{A\beta}{1 + A\beta}$$

$$A(\omega) = \frac{A(o)}{\left(1 + \frac{S}{\omega_1}\right)\left(1 + \frac{S}{\omega_2}\right)} \frac{\left(1 + \frac{S}{\omega_1'}\right)}{\left(1 + \frac{S}{\omega_0}\right)}$$

For simplicity assume $A(o) \omega_0 \gg \omega_2$.

$$A(\omega) = \frac{A(o)}{\left(1 + \frac{S}{\omega_1}\right)} \frac{\left(1 + \frac{S}{\omega_1'}\right)}{\left(1 + \frac{S}{\omega_0}\right)}$$

If $\omega_1 \approx \omega_1'$:

$$G = \frac{1}{\beta} \frac{A\beta}{1 + A\beta} \frac{\left(1 + \frac{S}{\omega_1'}\right)}{\left(1 + \frac{S}{\omega_1'}\right)\left(1 + \frac{S}{(1 + A\beta)\omega_0}\right)}$$

Step Response:

$$\frac{e_{OUT}}{e_{IN}}(t) = \frac{1}{\beta} \frac{A\beta}{1 + A\beta} \left[1 - \frac{\omega_1' - \omega_1}{\omega_1} e^{-\omega_1 t} - e^{-(1 + A\beta)\omega_0 t} \right]$$

FIGURE 15. Pole-Zero Transient Response.

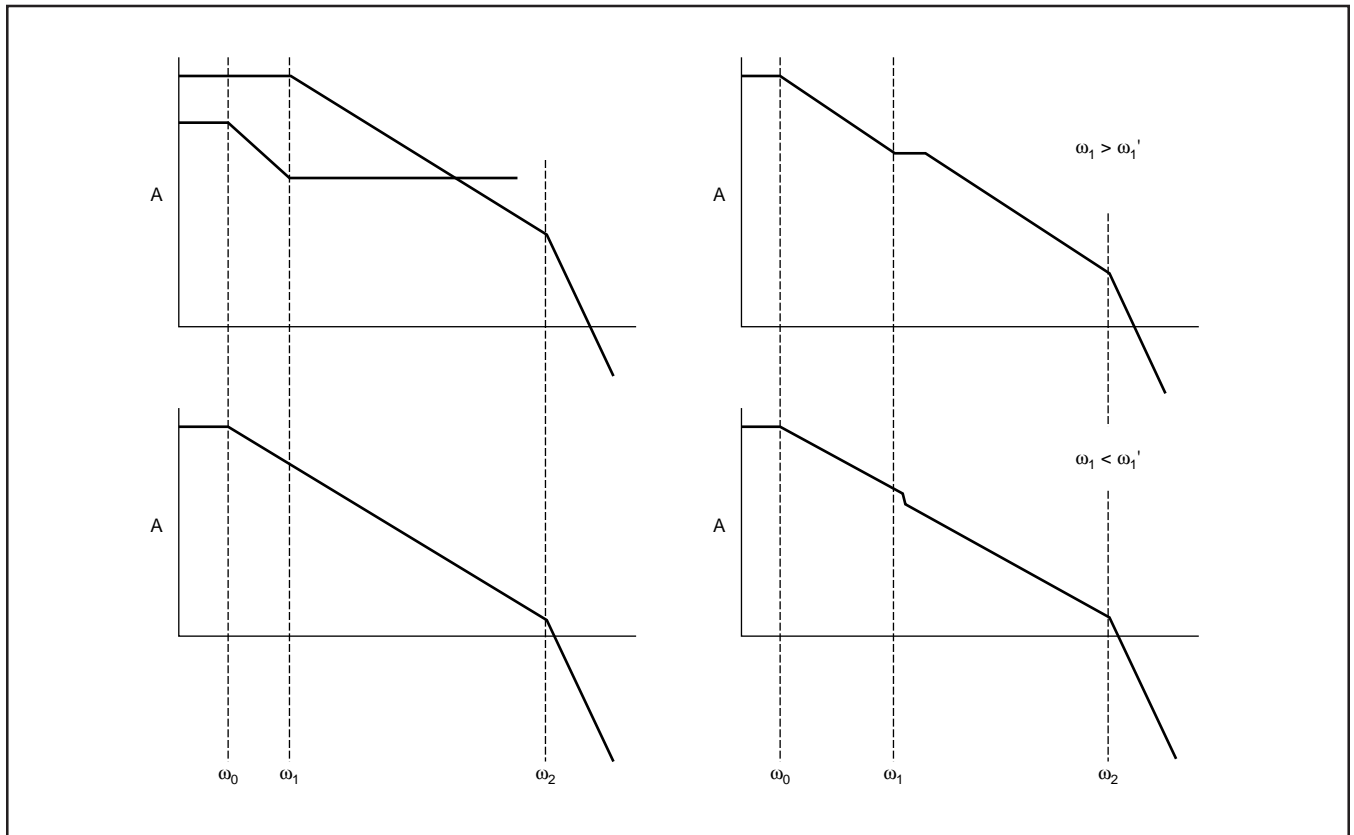


FIGURE 14. Pole-Zero Compensation Bode Plots.

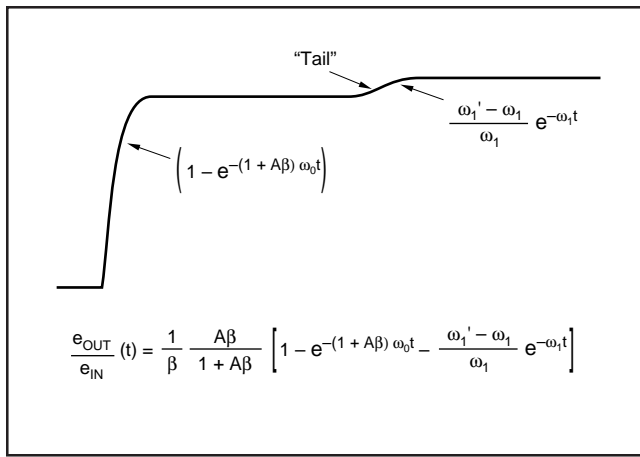


FIGURE 16. Pole-Zero Transient Response and Pole-Zero Mismatch.

the collectors of Q_1 and Q_2 while allowing the signal current to pass through transistors Q_4 and Q_5 with little attenuation. The term “folded cascode” refers to the fact that the PNP transistors not only serve as cascoding devices but also “fold” the signal down to a load connected to the negative power supply. Transistors Q_8 and Q_9 act as current source loads for transistors Q_4 and Q_5 thereby enabling the amplifier to achieve gains of up to 80 in a single stage. Emitter followers drive the output load in a similar manner to the method described for the FET operational amplifier. An additional benefit of this architecture is that the amplifier

can be stabilized with a single capacitor thereby approximating a single pole response without a settling “tail.”

COMPARATOR

The comparator is a common element in a signal processing system and it is used to sense a level and then generate a digital signal, either a “1” or a “0,” to report the result of that comparison to the rest of the system. Comparators can be implemented two different ways, either using a high gain amplifier or by using the latching type approach. Each type of comparator has advantages as will now be explained.

When a high gain amplifier is used as a comparator, many low gain stages are cascaded to achieve high gain bandwidth product. A simplified example of a 20ns comparator is shown in Figure 18. This is in contrast to the way a wideband operational amplifier would be designed. A design objective for a wideband operational amplifier would be to achieve high gain in a single stage to avoid accumulating an excessive amount of phase shift. Feedback will be applied around an operational amplifier. It is important to achieve a phase characteristic approaching single pole response. Phase shift through a comparator is usually not important although high bandwidth and low propagation delay is desirable. The design of an open loop amplifier and a comparator are similar. The main differences are that comparators do not have to have stable, or linear, gain characteristics and the output is designed to be logic compatible such as TTL or ECL. Unlike a linear open loop ampli-

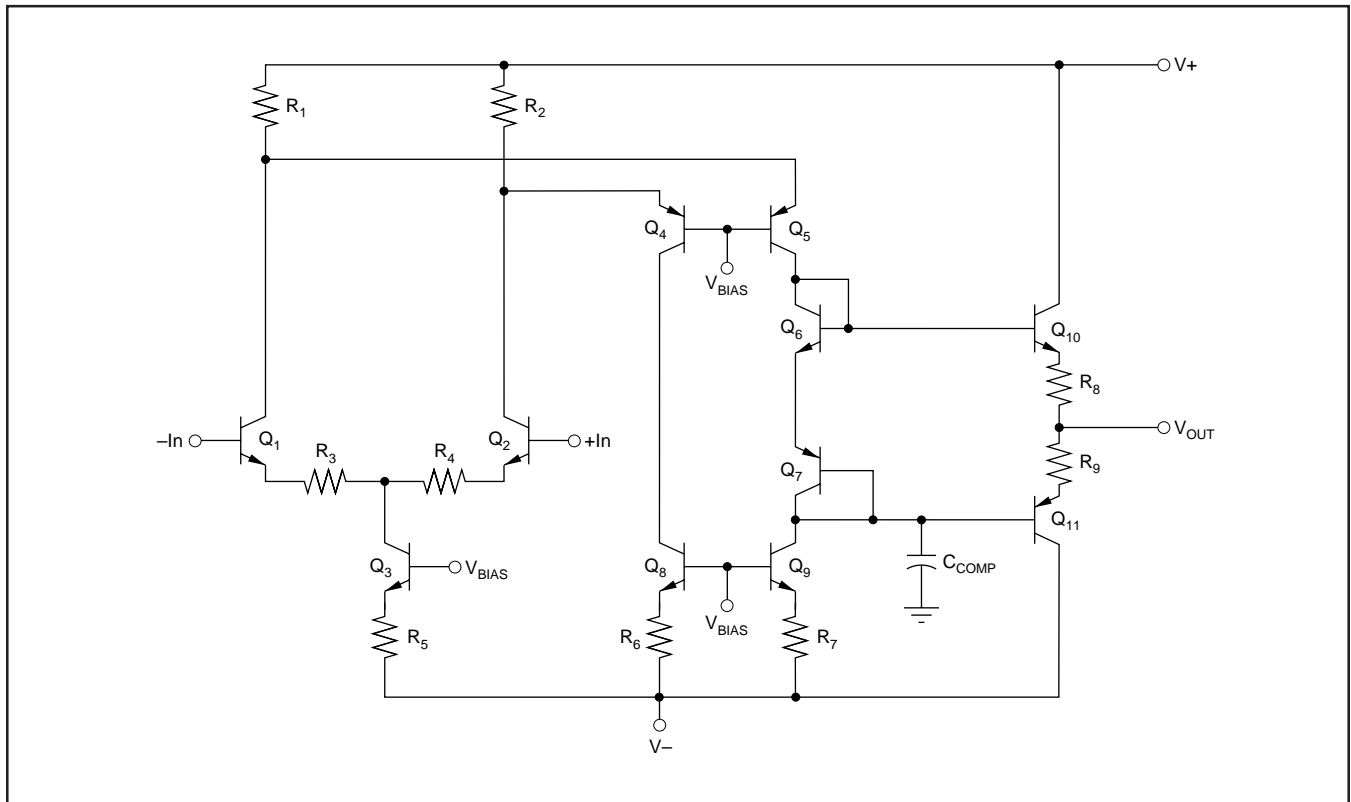


FIGURE 17. Folded Cascode.

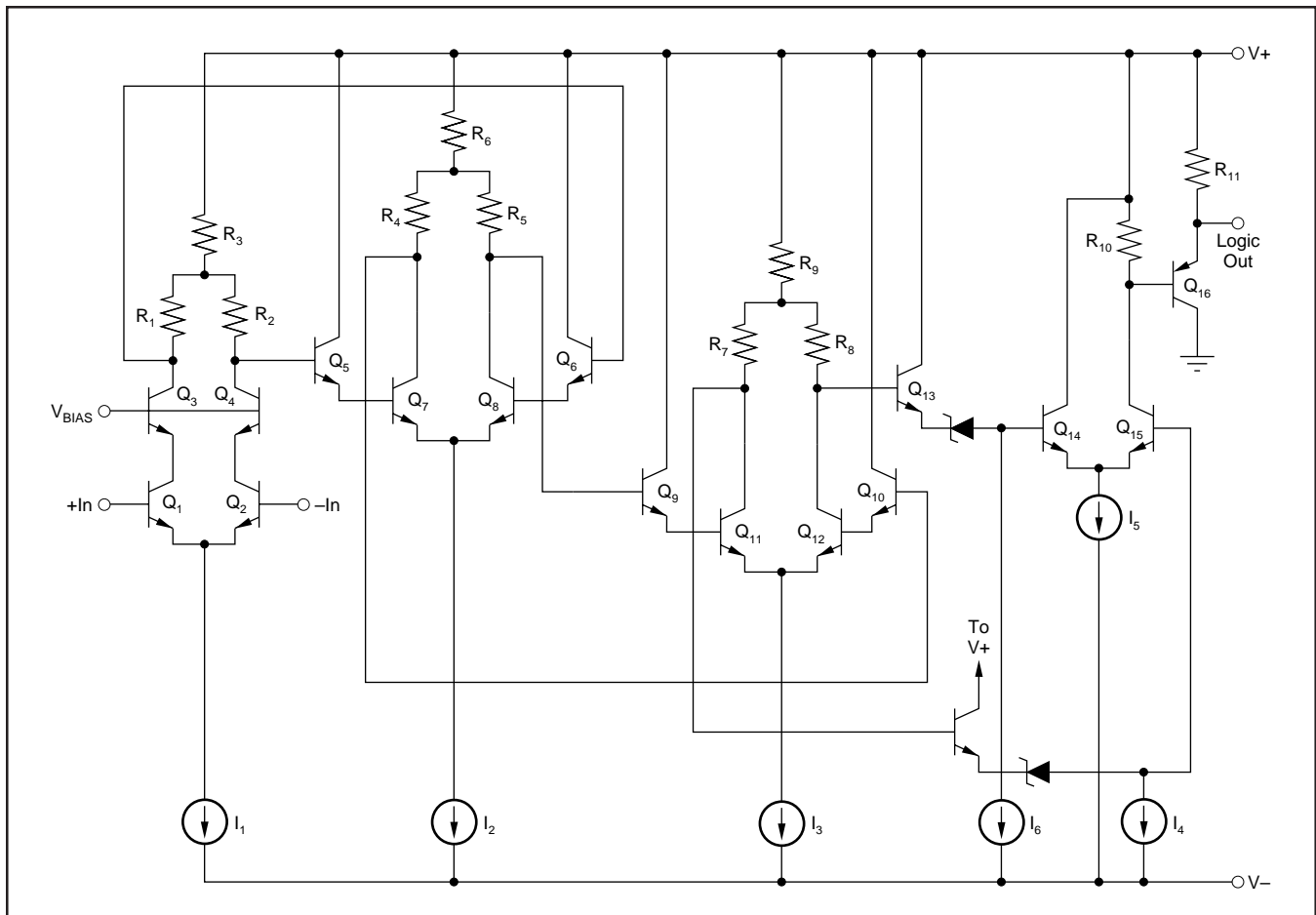


FIGURE 18. High Speed Comparator.

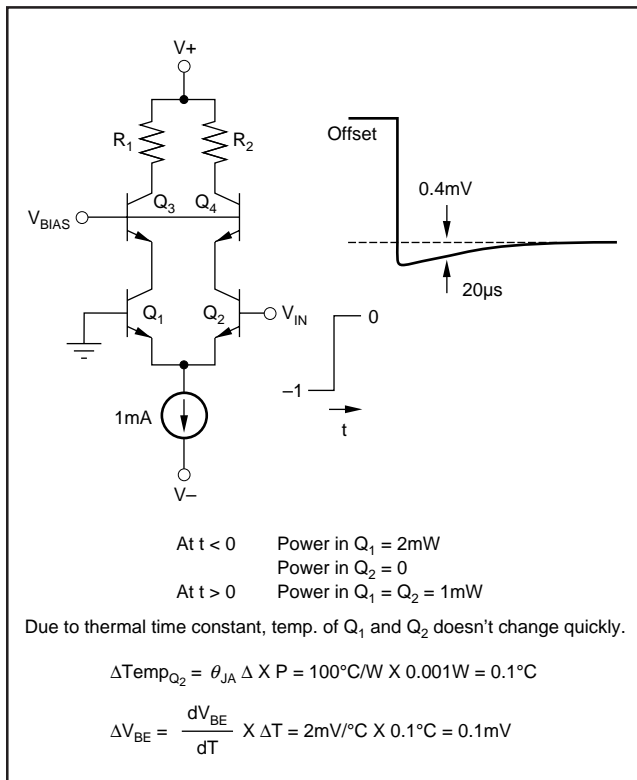


FIGURE 19. Thermal Offset.

fier, a comparator is designed to operate in a non-linear mode with the output saturating at either logic extreme, depending upon whether the input signal exceeds the input reference. Additionally, care is taken when the intermediate stages are designed to ensure excellent overload recovery and minimize propagation delay. Each stage of a well-designed comparator is designed in much the same fashion as an ECL logic stage in the sense that saturation is avoided and maximum interstage bandwidth is preserved by using emitter followers to couple the signal from stage to stage.

Comparator oscillation problems can be solved using a latching comparator, but both architectures are sensitive to the “thermally” induced offset. The thermally induced offset is created when the input signal biases the input differential amplifier to either being entirely “on” or “off,” thereby changing the power dissipation of one side of the differential pair with respect to the other side. This effect should be minimized by reducing the power dissipation of the input differential pair. The limit is determined by bandwidth and slew rate requirements of the inputs. Figure 19 shows a calculation which estimates the amount of thermally induced offset in a comparator. This calculation shows that the comparator offset will initially be 0.1mV before coming into balance at the rate determined by the thermal time constant of the system. The thermal time constant of the system is in

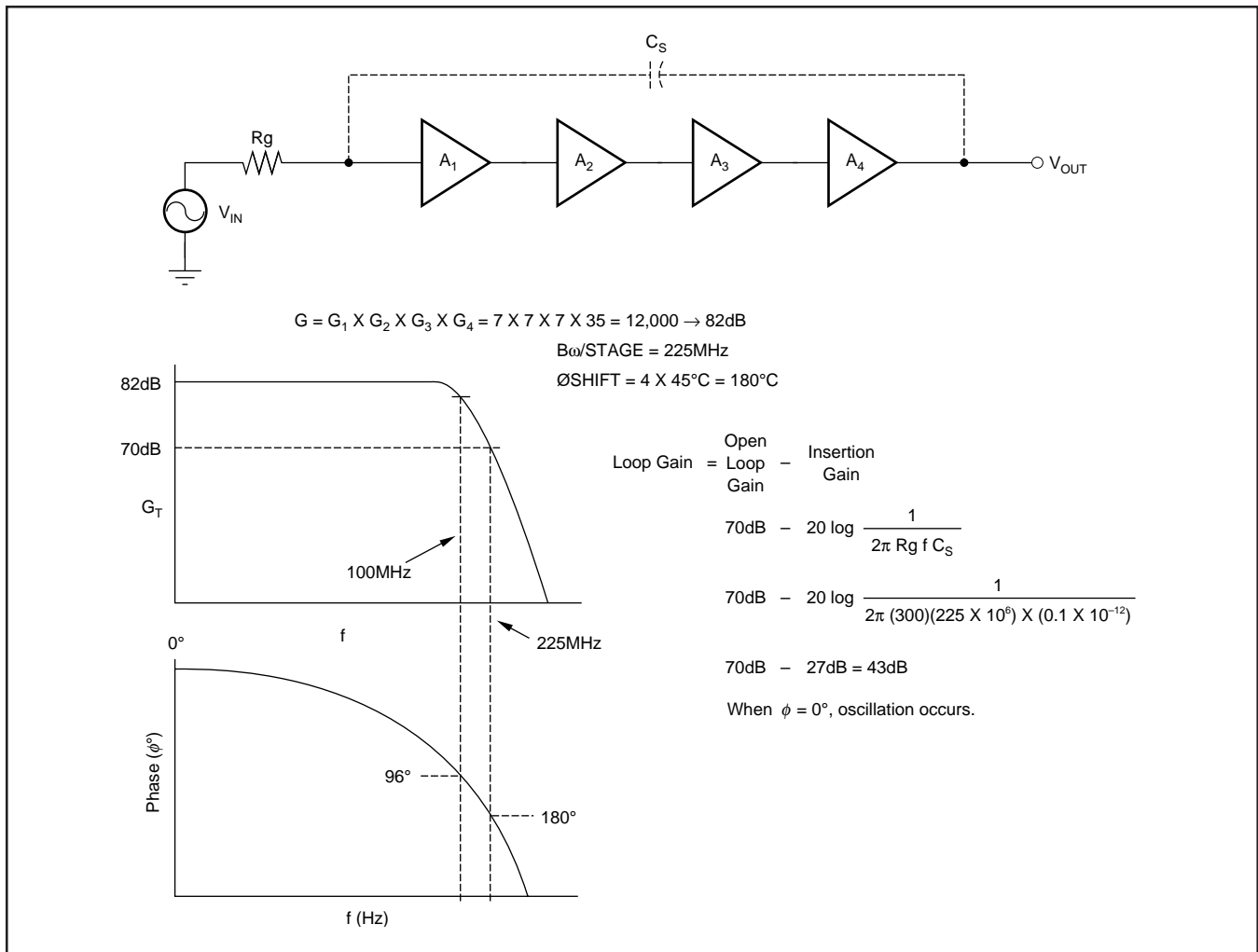


FIGURE 20. Comparator Oscillation.

the order of 10 μ s to 100 μ s and is affected by factors such as the physical size of the transistor as well as the method by which the transistor die is attached to the header. Thermally induced offset can become a serious problem in high speed, high accuracy systems and can often be the limiting factor that determines performance.

The other effect that limits the accuracy of the non-latched comparator is the tendency for oscillations. Comparator oscillations are usually due to parasitic feedback from the output signal being capacitively coupled back to the comparator's input. In discrete form this problem is very difficult to solve while still trying to maintain high sensitivity and low propagation delay. Figures 20 and 21 show a diagram which describes how the parasitic feedback between the pins of the package comparator can create enough feedback to stimulate an oscillation. Even in hybrid form, comparator oscillation is a serious problem. Integrating the comparator onto a monolithic chip is beneficial as the planar nature of this means of construction will reduce the amount of parasitic capacitance.

As previously mentioned, the other type of comparator that is employed is known as the "latching type." Figure 22 shows a simplified schematic of the front end of this type of

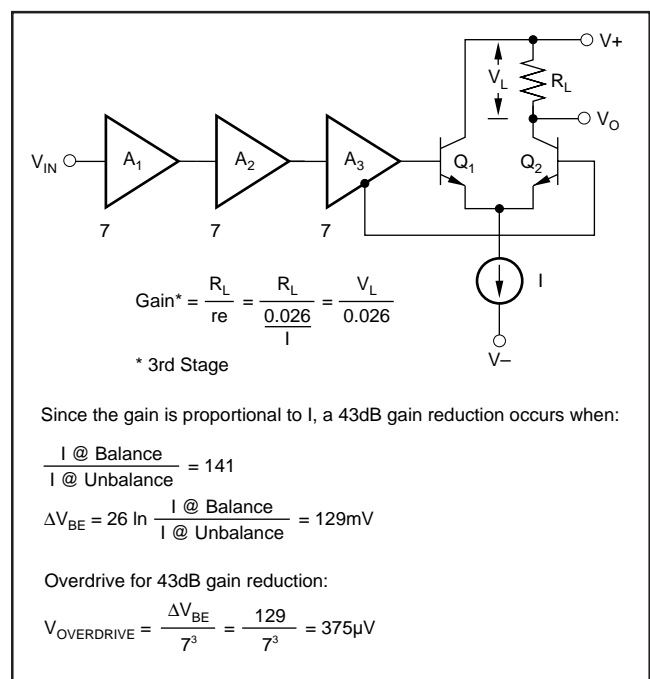


FIGURE 21. Gain Reduction to Stop Comparator Oscillation.

comparator. The latching comparator develops “high gain” by going into a regenerative state when being strobed by the latch enable signal. Typically the input pair, Q_1 and Q_2 , will have a gain of at least 10 when the comparator is in the “tracking” mode. At the instant of latching the “tail current,” I_3 , is then switched from the input “linear pair” to input “latching pair” Q_3 and Q_4 . The state of the latch will then be determined by the state of the signals on the input bases of Q_1 and Q_2 with respect to each other. The latching pair receives its feedback through the two emitter followers, Q_7 and Q_8 . The emitter followers also feed the appropriate logic level. While the comparator is held in the latch state it is impossible for oscillations to occur as the comparator is permanently held in the previous state. If the comparator is placed in the low gain state for a short amount of time the tendency for oscillation is reduced as: 1) the loop gain is too low to support oscillation, and 2) if the strobe time is short the state of the latch is already determined before the parasitic oscillation is permitted to build up.

The fastest analog to digital converter is composed of an individual latch comparator for each quantizing level. The design of this type of converter would not be practical by any other technique as only the latching comparator offers sufficient simplicity to allow for the necessary amount of integration. More detail about “flash converters” will be

given later. GaAs is an emerging, exceptionally high speed technology; while being able to achieve high speed, it does so at the expense of low gain. The latching comparator is extremely useful here, as it would not be practical to achieve the comparator function in this technology any other way.

TRACK AND HOLDS

One of the most important elements of a data conversion system is the track and hold. Track and holds and sample and holds are very similar and for all practical purposes are identical. If the track and hold or sample and hold command is direct coupled, the two types of samplers are identical. Some types of extremely fast samplers have their sample command AC coupled and for a short period of time the signal is “sampled” and then held; hence the name “sample and hold” was coined. Before a description of many track and hold architectures are given it will be appropriate to explain why track and holds generally precede an analog to digital converter.

A track and hold is used to reduce the aperture time of the sampling system. (A sampling system would be a track and hold driving an analog to digital converter.) In general it is necessary for the input signal to the analog to digital converter to be constant during the conversion process to avoid

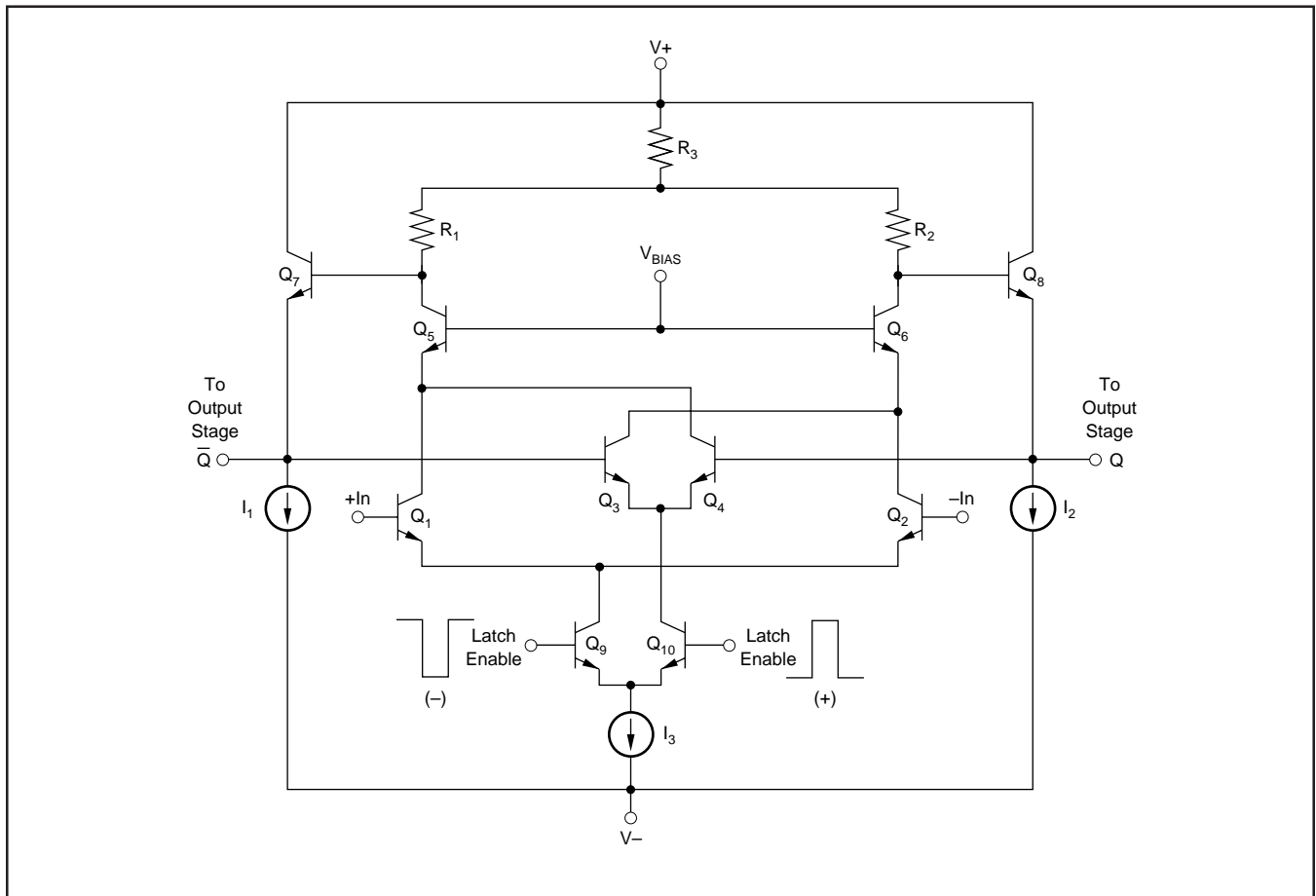


FIGURE 22. Latch Comparator Front End.

$$e(t) = \frac{EFS}{2} \sin 2\pi ft$$

EFS = Full Scale ADC Range

$$\frac{d e(t)}{dt} = f EFS \pi \cos 2\pi ft, \quad \frac{d e(t)}{dt} = f EFS \pi$$

$$f = \frac{\frac{d e(t)}{dt}}{EFS \pi}$$

Assume maximum allowable change during ADC conversion time.

$T = 1/2LSB$ and $EFS = 2^N LSB$ where N is the number of bit ADC.

$$f = \frac{1/2LSB}{T 2^N LSB \pi} = \frac{1}{\pi 2^{(N+1)} T}$$

As an example, let $N = 12$ and $T = 1\mu s$:

$$f_{MAX} = \frac{1/2LSB}{T 2^N LSB \pi} = 38.9Hz$$

With a sample/hold, the maximum frequency would be 500kHz.

FIGURE 23. Maximum Input Frequency for ADC Without a Sample/Hold.

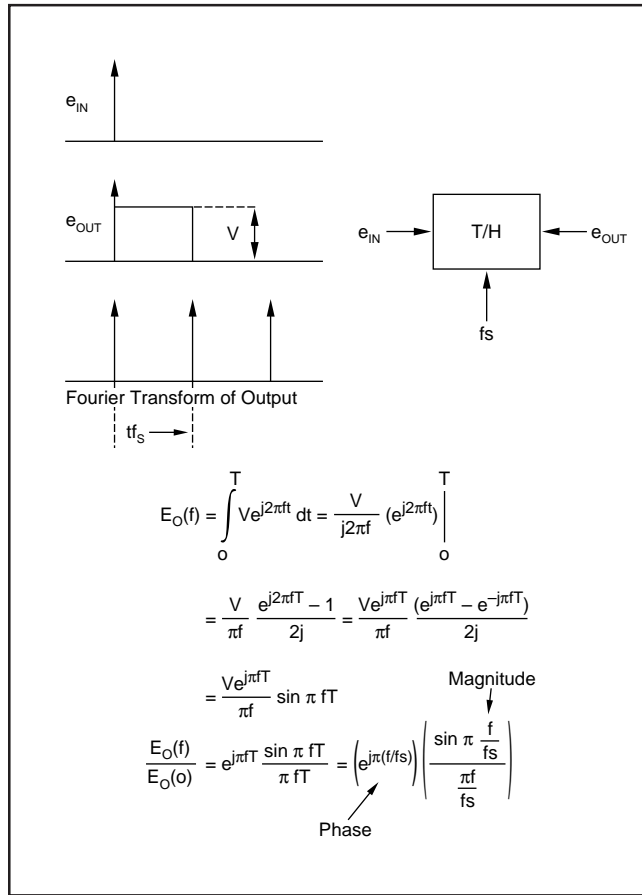


FIGURE 24. Track/Hold Bandwidth.

error. A successive approximation ADC uses an N-step algorithm when forming the conversion and if the signal varies during the conversion process the wrong approximation would take place. Even flash converters can benefit from being driven from a track and hold since the time delays of all the comparators are not identical. Figure 23 calculates the improvement in ADC performance that can be obtained when a track and hold precedes an ADC. Figure 23 shows that the maximum frequency that can be processed by a $1\mu s$ ADC would be only 38.9Hz. When a sample and hold drives the ADC the maximum frequency would rise to the Nyquist rate of 500kHz. Additionally, applications will be shown of how track and holds can be used to “deglitch” DACs and how a peak detector can be formed. Many track and hold architectures will be presented with a discussion of the strengths and weaknesses of each type. The discussion will show how the characteristics of the sample and hold interact to gain an understanding of how to optimize the design for particular applications. This will also be a useful way of understanding the increasing level of complexity. Before the comparison of different types of architectures begins, Figure 24 calculates the bandwidth of a track and hold and Figure 25 shows a plot of the frequency response.

As a way of introduction, the most elementary track and hold is shown in Figure 26. A FET switch is connected to a capacitor which in turn is isolated from the output by a high input-impedance buffer. When the sampling signal, which is connected to the gate of the N-channel enhancement mode FET, is in the high state, the FET series resistance is at its lowest which is R_{ON} . During this condition the output of the buffer is the input signal modified by the low pass filter action of R_{ON} and the holding capacitor C . The voltage across the holding capacitor will follow the input voltage until the gating signal is returned to the low state and the FET is turned off. At that point the holding capacitor retains the input voltage at the instant of sampling. Figure 27 shows waveforms that depict the dynamic characteristics of the track and hold. When the track and hold is driving an analog to digital converter, the held voltage is then converted to its digital equivalent. The circuit previously described has limited capability. To determine the nature of the limitations, a design example will show how the performance is determined. For this design example a typical N-channel D-MOS FET will be used along with a FET op amp connected as a voltage follower. The FET has the following characteristics:

1. $R_{ON} = 50\Omega$
2. $V_T = 2.5V$
3. $C_{GD} = 0.5pF$
4. $C_{DS} = 0.1pF$
5. $I_{DSS} = 25mA$
6. $I_{OFF} = 50pA$

Assume for this example that the input signal range is 10V peak to peak and it is tolerable that each error source can contribute $\pm 0.01\%$ of V_{IN} to the overall error. Particular

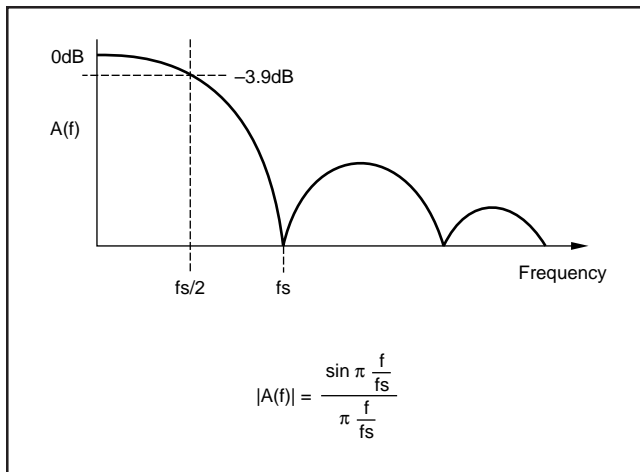


FIGURE 25. Frequency Response of Sample/Hold.

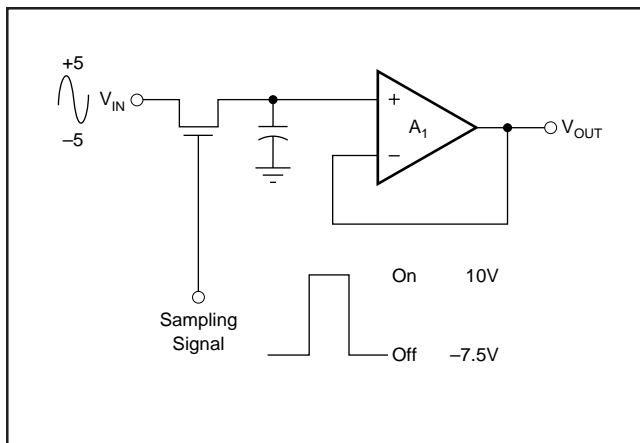


FIGURE 26. Basic Sample/Hold.

applications may assign the value of individual error sources differently. The sources of error that will be considered are:

1. Change induced offset error
2. Aperture non-linearity
3. Signal feedthrough
4. Aperture jitter
5. Aperture delay
6. Droop
7. Acquisition time
8. Track to hold settling
9. Full power bandwidth

CHARGE INDUCED OFFSET OR PEDESTAL ERROR

To ensure that the FET is turned on with a low resistance it is necessary to exceed the peak input signal by 5V. Therefore the voltage applied to the gate of the FET is

$$V_{ON} + V_{PEAK} = 5 + 5 = 10V$$

To ensure that the FET is off it is necessary that the FET is reverse biased under the worst case conditions. The minimum voltage that the sample and hold has to process is -5V and it is desirable to reverse bias the gate to source under

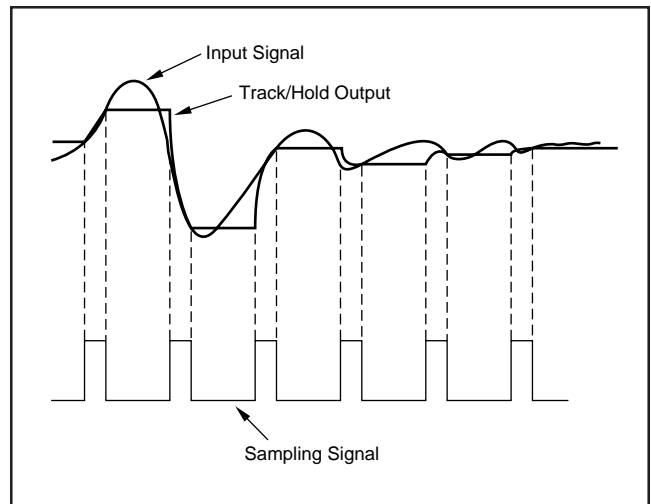


FIGURE 27. Track/Holds Wave Forms.

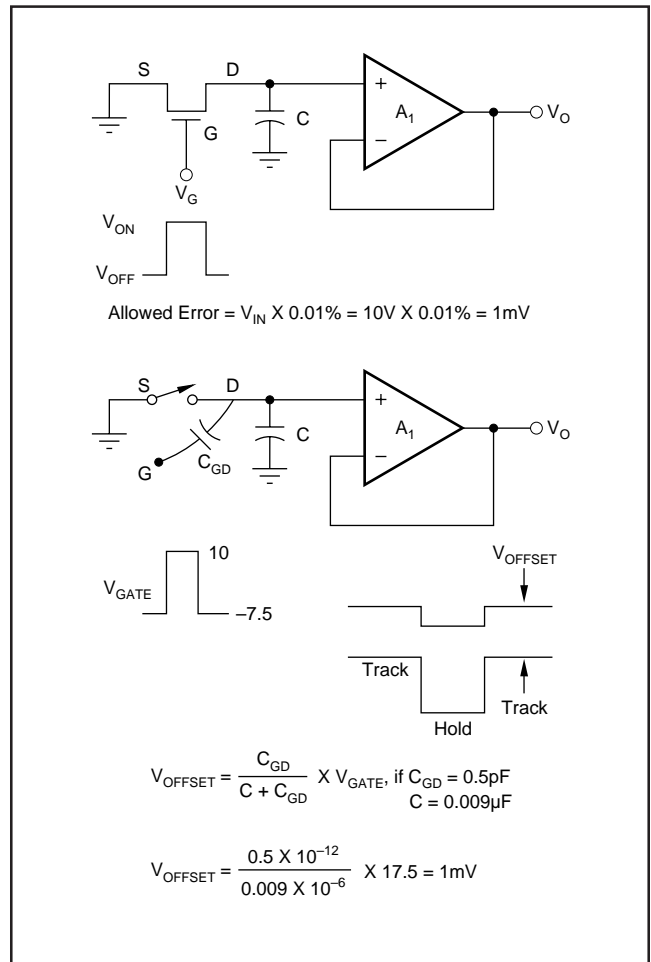


FIGURE 28. Charge Induced Error.

these conditions so the off signal that is applied to the gate of the FET is -7.5V. See Figure 26. The total signal swing that is applied to the gate of the FET is therefore 17.5V, the sum of the on and off signals. Figure 28 shows how a voltage divider is formed by the gate to drain capacitance C_{GD} and the holding capacitor C . A charge induced offset error is

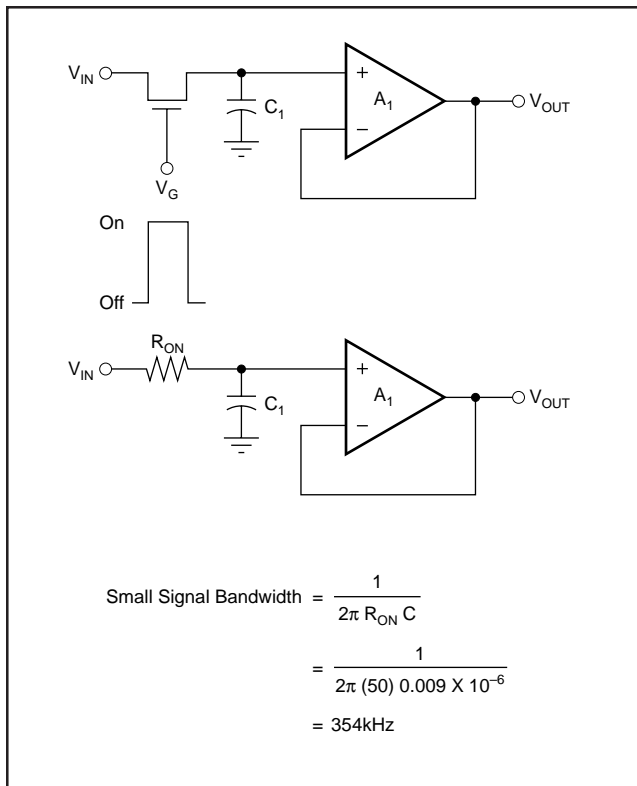


FIGURE 29. S/H Frequency Response.

then created by this voltage divider action and its value is given by:

$$V_{\text{OFFSET}} = V_{\text{GATE}} \left(\frac{C_{\text{GD}}}{C + C_{\text{GD}}} \right)$$

Therefore, to reduce the charge induced offset error to:

$$0.01\% \times 10 = 1\text{mV}$$

requires a holding capacitor of:

$$C = \frac{C_{\text{GD}} \cdot V_{\text{GATE}} - C_{\text{GD}} \cdot V_{\text{OFFSET}}}{V_{\text{OFF}}}$$

$$C = \frac{0.5\text{pF} \cdot 17.5\text{V} - 0.5\text{pF} \cdot 1\text{mV}}{1\text{mV}} = 8.75\text{nF}$$

Now since the value of the holding capacitor ($C_H = C$) is determined, the track and hold bandwidth would be (see Figure 29):

$$\text{BW} = 1/2\pi(R_{\text{ON}})(C_H) = 1/2\pi(50)(9 \times 10^{-9})$$

$$= 354\text{kHz}$$

APERTURE INDUCED NON-LINEARITY

In the previous discussion on charge induced offset error it was assumed that the gate turn off signal was always 17.5V. If the input signal were sampled at its peak of 5V and the FET threshold voltage were 2.5V, the FET would stop conducting when the voltage on the gate was 7.5V. The effective gate signal swing would be reduced to 15V and the amount of charge induced offset would also be reduced.

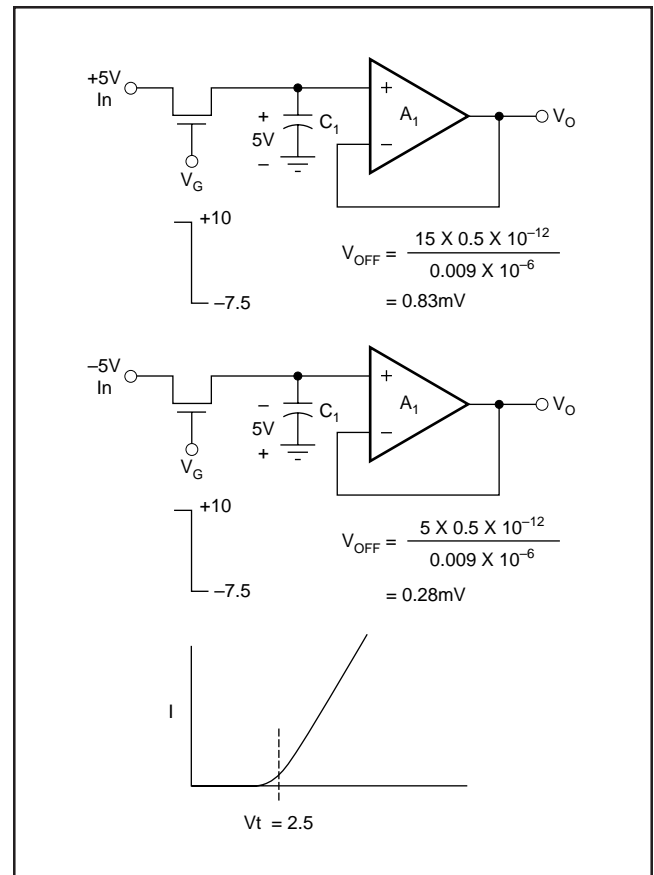


FIGURE 30. FET Threshold Characteristics and Aperture Non-linearity.

Similarly, if the input signal being sampled is at the minimum level of -5V, the effective gate swing would be 5V. In the previous section it was calculated that if the gate swing were 17.5V, the charge induced offset would be 1mV. See Figure 30. Actually the charge induced offset is modulated by the signal and varies between 0.83mV for the positive extreme and 0.28mV for the negative extreme. Since both offsets are less than the allowable error this is not a problem; the holding capacitor is relatively large. As will be seen later, this will not always be the case when it is desirable to achieve wider band operation. This effect will be considered again for wider band designs when it could become a serious source of error.

SIGNAL FEEDTHROUGH

Signal feedthrough occurs because of the presence of a capacitor that is connected from the drain to the source of the FET. This is a parasitic capacitor that is either due to layout or other stray effects. Referring to Figure 31, it is seen that the input signal will be coupled to the hold capacitor and is given by:

$$V_{\text{FEEDTHROUGH}} = (C_{\text{DS}}/C_H)(V_{\text{IN}}) = (0.1\text{pF}/0.009\mu\text{F})(10)$$

$$= 111\mu\text{Vp-p}$$

which is a tolerable error. Again it will be seen that wider

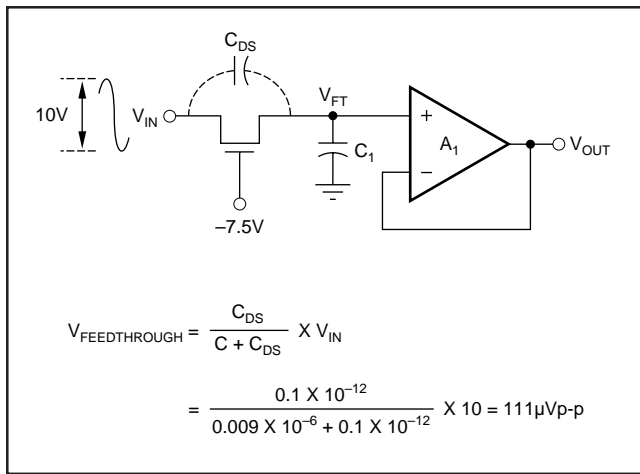


FIGURE 31. Signal Feedthrough.

bandwidth designs that must resort to smaller holding capacitors will not be able to meet this specification as easily.

APERTURE JITTER

Aperture jitter or uncertainty is the variation in the time when the sample and hold switch opens after a sample and hold transition occurs, or the time variation in the aperture delay. (Aperture delay is the elapsed time from the activation of the sample to hold command to the opening of the switch in the hold mode.) There are two sources of aperture jitter: power supply induced noise and threshold variation due to thermal noise. If attention is paid to filtering the power supply properly, as well as using a well-regulated power supply, this will not be a source of aperture jitter. As a practical matter, because of measurement difficulty, determining the amount of aperture jitter that is present in the system is often more of a problem than limiting it to an acceptable level. Techniques for measuring aperture jitter will be shown in the measurement section. Assume that the noise associated with the logic threshold is $50\text{nV}/\sqrt{\text{Hz}}$. This would be ten times greater than the noise of a typical linear amplifier. Further assume that the bandwidth of the logic circuit that develops the gate signal is 300MHz . The noise variation of the logic level would then be (see Figure 32):

$$\text{Threshold variation due to logic noise} = (50\text{nV})(\sqrt{300\text{E}6}) = 0.87\text{mV}$$

If the logic signal rate of transition were 0.4V/ns , the aperture jitter would be (see Figure 33):

$$\text{Aperture jitter} = (\text{threshold noise})/(\text{logic slew rate}) = t_A$$

$$= (0.87\text{mV})/(0.4\text{V/ns}) = 2.2\text{ps}$$

which will be seen to be negligible for all but the highest sampling rate data conversion applications. Aperture jitter can create amplitude noise by causing a variation of the sampling point of dynamic signals. The noise can be predicted by:

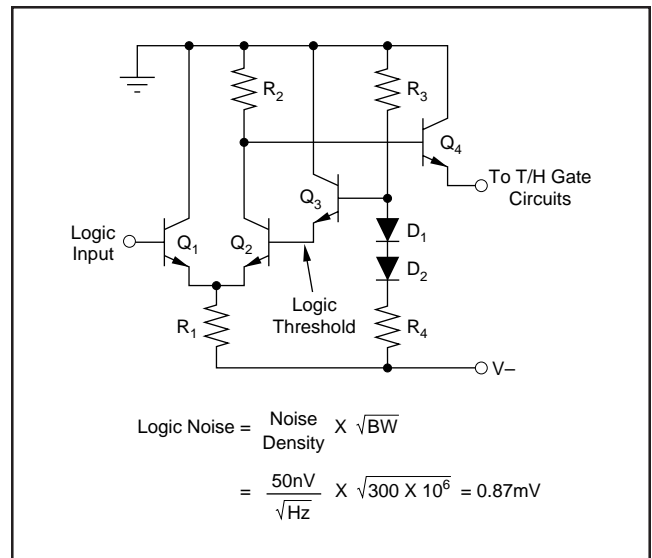


FIGURE 32. Logic Noise.

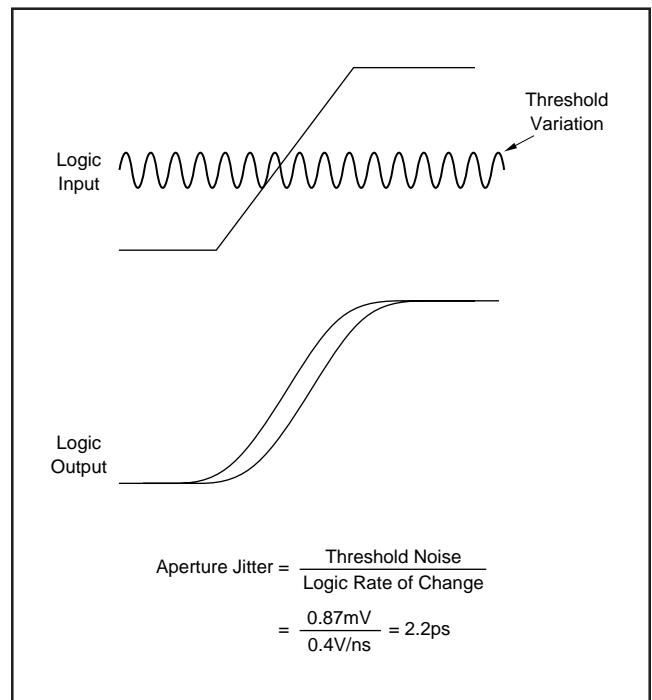


FIGURE 33. Aperture Jitter.

$$\text{Aperture induced noise} = (\text{signal rate of change})(\text{aperture jitter})$$

$$= (de/dt)(t_A)$$

$$= (\text{FS})(\pi)(f)(t_A)$$

Assume a 12-bit ADC with a sampling rate of 20MHz . $\text{FS} = 4096\text{LSB}$, $f = 10\text{MHz}$, $t_A = 2.2\text{ps}$.

$$\text{Aperture induced noise} = (4096)(\pi)(t_A)$$

$$= (4096)(\pi)(10\text{E}6)(2.2\text{E}-12) = 0.28\text{LSB}$$

0.28LSB aperture induced noise would be acceptable for a 12-bit ADC with a Nyquist rate of 10MHz. Figures 34 and 35 illustrate this effect.

APERTURE DELAY

The aperture delay is the elapsed time from the activation of the sample to hold command to the opening of the switch in the hold mode. See Figure 36. Controlling aperture delay is important when multiple channels need to be matched to

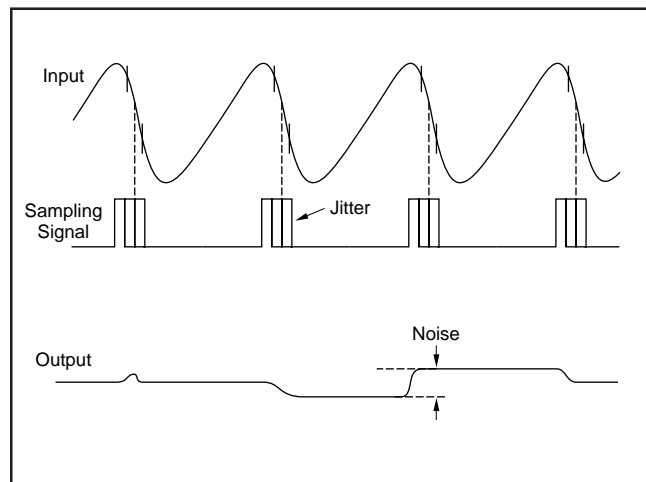


FIGURE 34. Aperture Induced Noise.

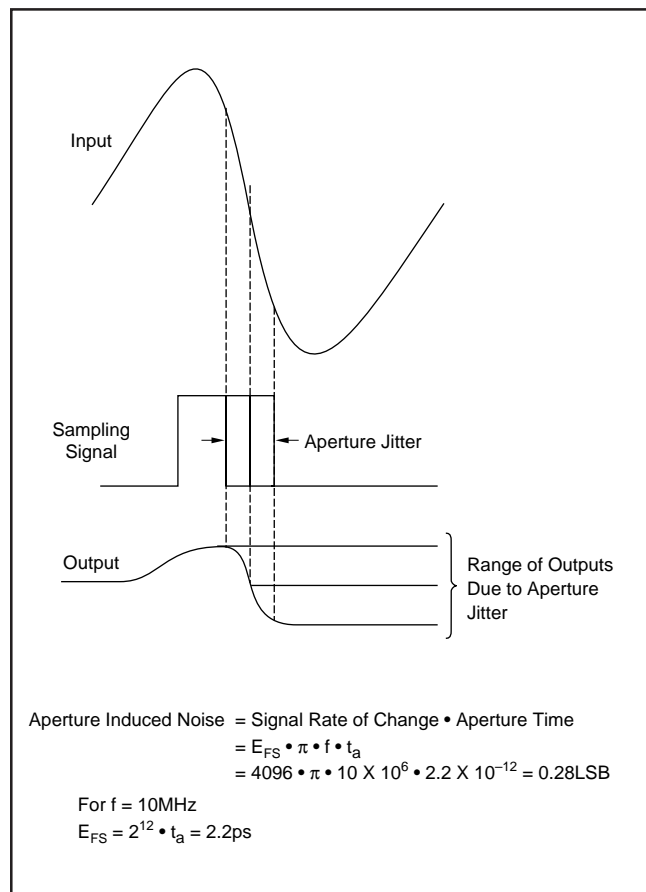


FIGURE 35. Aperture Induced Noise.

each other. Figure 37 shows a circuit diagram of a FET driver circuit that is TTL compatible and is suitable for driving the sample switch.

DROOP

While the sample and hold is in the hold mode the leakage current that flows through the FET and the input bias current of the operational amplifier will tend to discharge (or charge) the holding capacitor. Both sources of current are about 50pA at 25°C so the capacitor will change at a rate of (see Figure 38):

$$I/C = 100\text{pA}/0.009\mu\text{F} = 0.011\text{V/s}$$

If the sample and hold were driving an ADC with a 10μs conversion time, the held value would change by 0.11μV during the conversion process. Since the allowable error from each source is 1mV, this is not a source of error at room temperature. Since leakage current doubles every 10°C, when the operating temperature increases to 125°C, the voltage change, due to droop, during the conversion would increase to 0.11mV, which is still below the allowable value. Wider band designs that use smaller holding capacitors will not meet this specification as easily and other methods will be shown that can reduce the droop to acceptable levels.

ACQUISITION TIME

The calculation of acquisition time of a sample and hold is identical to the way settling time is determined for an operational amplifier. (It is really the same phenomenon.) The sample and hold will slew in response to a large signal change until the output rate of change of the sample and hold

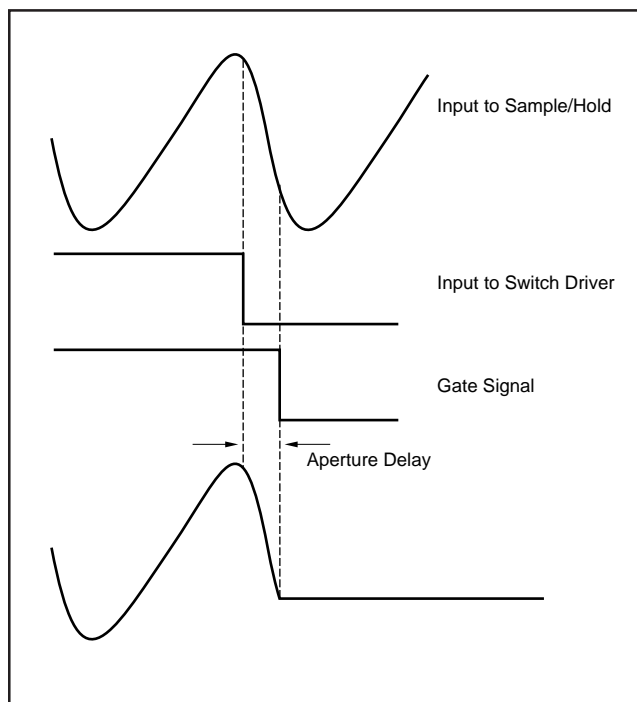


FIGURE 36. Aperture Delay.

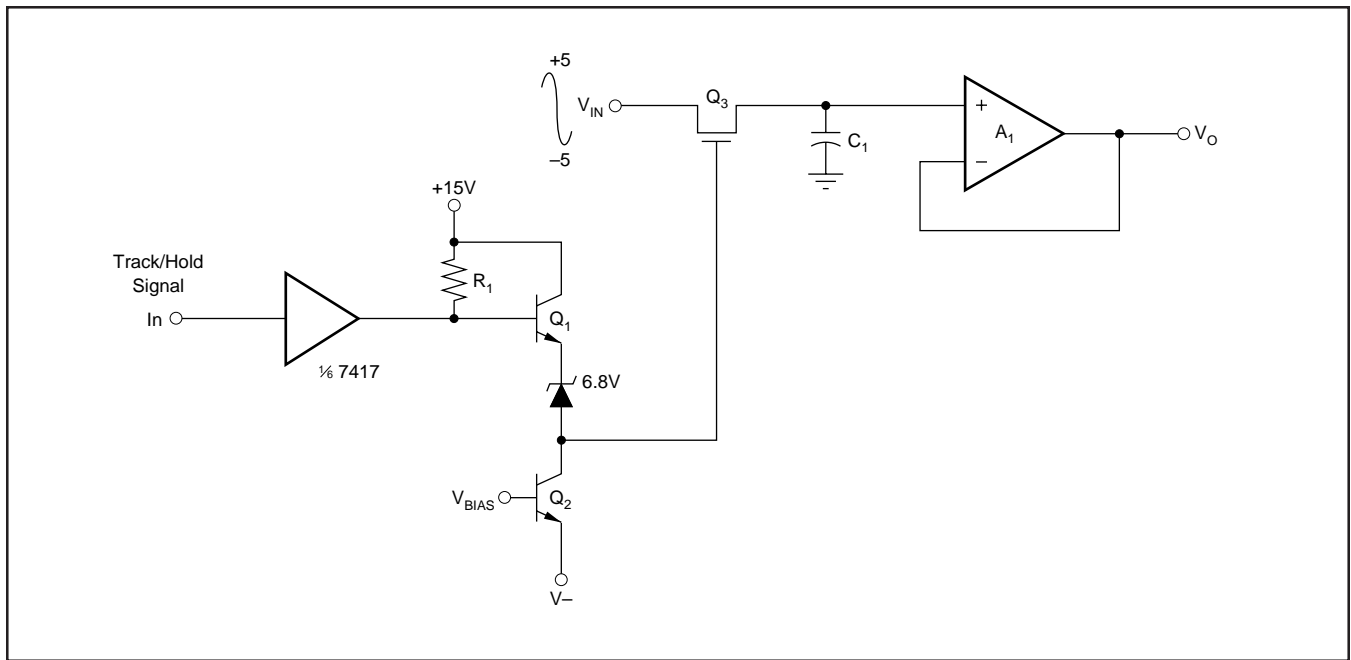


FIGURE 37. FET Switch Driver.

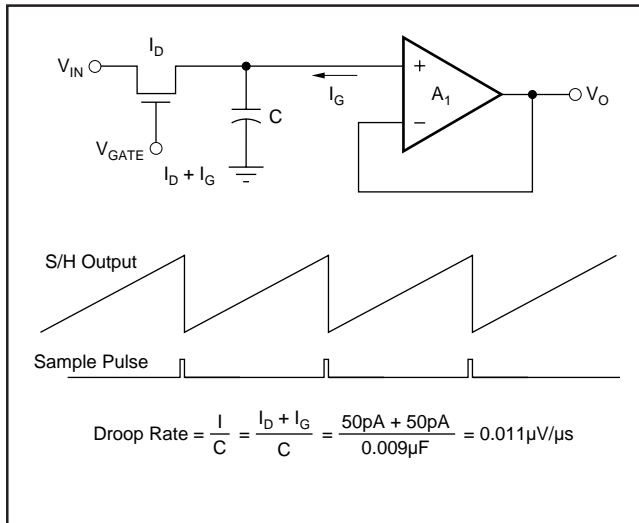


FIGURE 38. Droop.

is within its linear capability. When V_{IN} is large enough to pinch off the FET, the slew rate of the sample switch is given by (see Figure 39):

$$\text{Slew rate} = I_{DSS}/C_H = 25\text{mA}/0.009\mu\text{F} = 2.8\text{V}/\mu\text{s}$$

The sample and hold will slew until the remaining output change is within the linear capability of the sample and hold (see Figure 40):

$$V = (R_{ON})(I_{DSS}) = 1.25\text{V}$$

For a 10V input change the sample and hold will slew until the output is within $10\text{V} - 1.25\text{V} = 8.75\text{V}$ of its final value which will take $(8.75\text{V})/(2.8\text{V}/\mu\text{s}) = 3.1\mu\text{s}$. The remainder of the acquisition time occurs as the remaining 1.25V has to settle to within 1mV of the final value. Assuming single pole

settling, the linear part of the settling is given by the formula (see Figure 41):

$$V = V(1 - e^{-(t/R_{ON} \cdot C_H)})$$

Rearranging:

$$t = (R_{ON})(C_H) \ln(V/0.001) \\ t = (0.45\mu\text{s}) \ln(1.25/0.001) = 3.2\mu\text{s}$$

The acquisition time of the sample and hold is then the sum of the time spent during slewing and the time spent during linear settling, or:

$$\text{Acquisition time} = 3.1\mu\text{s} + 3.2\mu\text{s} = 6.3\mu\text{s}$$

It has been assumed that the settling time of the buffer or operational amplifier that the holding capacitor drives is much smaller than the above number of $6.3\mu\text{s}$. In fact, it is possible to get monolithic FET amplifiers with settling times under $1\mu\text{s}$, which will increase the above acquisition time by perhaps 3%.

TRACK TO HOLD SETTLING

Track to hold settling is the time that the sample and hold takes to recover from the gate transient that is coupled on to the hold capacitor and the settling time of the buffer that isolates the hold capacitor from the output. The sample and hold that is being designed has a large enough hold capacitor so that track to hold settling is not an important factor for this sample and hold. Another example will better illustrate track to hold settling.

FULL POWER BANDWIDTH

The full power bandwidth of a sample and hold is calculated in the same manner as it is for an operational amplifier.

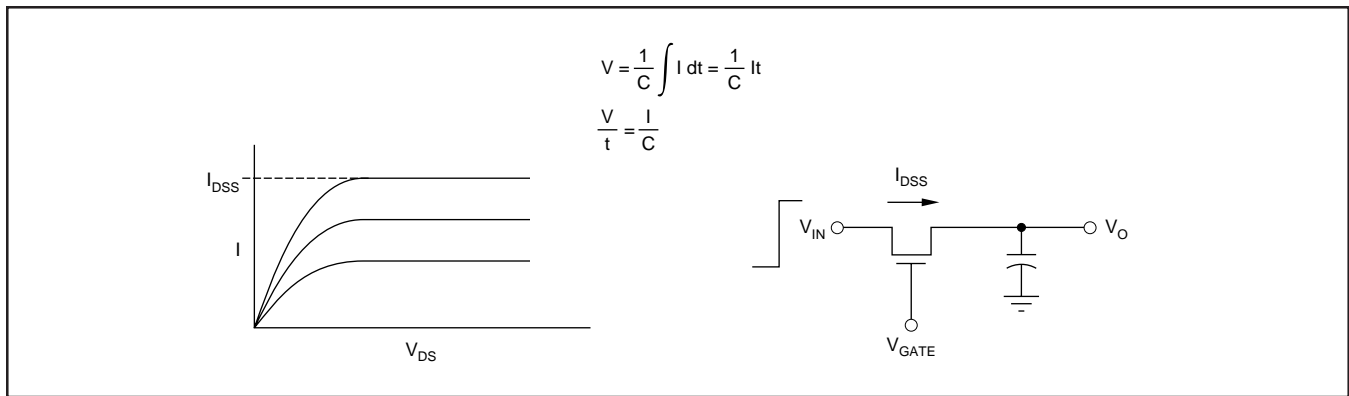


FIGURE 39. Capacitor Charged from a Constant Current.

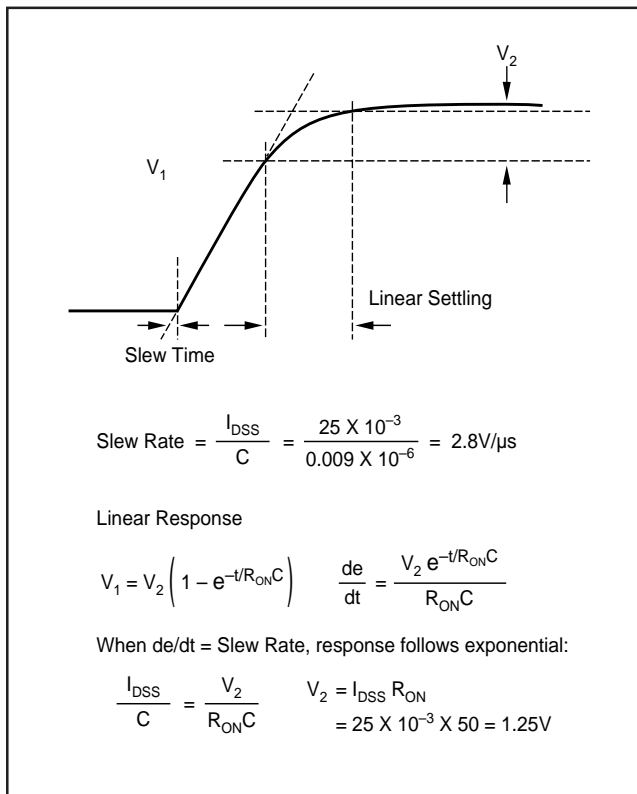


FIGURE 40. Acquisition Time.

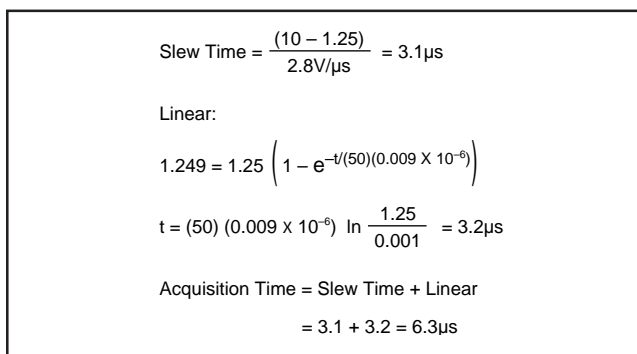


FIGURE 41. Acquisition Time.

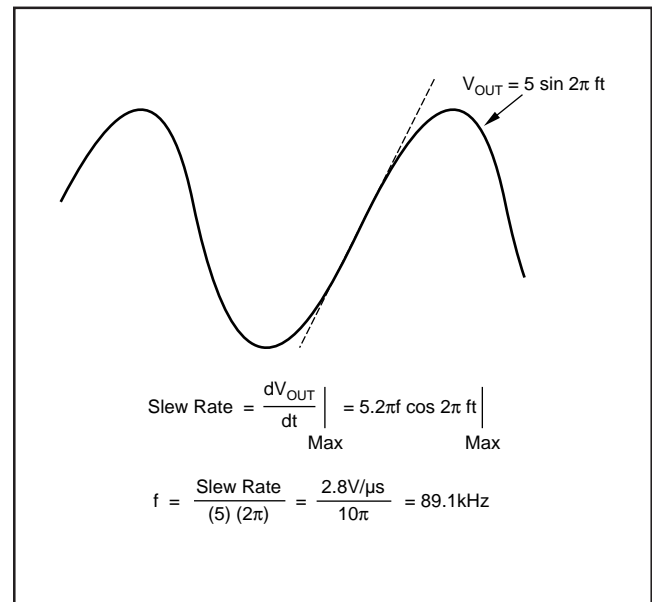


FIGURE 42. Full Power Bandwidth.

Knowing the full power bandwidth is important as it is necessary to operate at less than that frequency to maintain low levels of distortion. For the design example in question (see Figure 42):

$$V_{OUT} = (5) \sin(2\pi fT) \text{ and:} \\ dV_{OUT}/dT = \text{Max Slew Rate} = 10(\pi)f$$

Rearranging terms:

$$\text{Full Power Bandwidth} = (\text{Slew Rate})/10(\pi) = 89.1\text{kHz}$$

The above example demonstrates how to approach the design of the simplest type of track and hold. Even though it is simple, it would be very useful as the full power bandwidth of 89.1kHz would be adequate for processing audio signals. A sample and hold with an acquisition time of 6.2μs driving an ADC with a 10μs conversion time would have an adequate sampling rate to process an audio signal. Furthermore, this circuit could be built for a cost in the \$5-10 range. While the design of this circuit is relatively straightforward, it does have limited bandwidth. Several

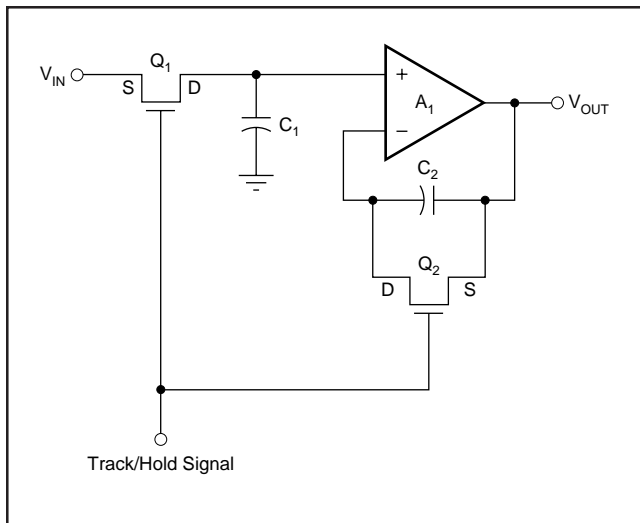


FIGURE 43. Balanced Track/Hold.

more examples of track and hold designs will be given showing how substantial increases in bandwidth can be made without sacrificing much in the way of linearity.

Assume for the purposes of this example that it is necessary to reduce the acquisition time of the sample and hold to 300ns. This would be about a twenty-fold decrease compared to the previous sample and hold that was designed. From the reference point of the previous example the holding capacitor would have to be reduced by a factor of twenty to reduce the acquisition time from 6.2s to 0.3s. If that were done, both the aperture induced linearity and the droop would increase by that same factor. Neither is desirable, as the goal of this new design is to achieve $\pm 0.01\%$ linearity. Figure 43 shows a simplified circuit diagram of a sample and

hold which is capable of faster operation while still maintaining good linearity. The track and hold shown in Figure 43 is faster due to the balanced connection of matched FETs Q_1 and Q_2 . When this track and hold makes the transition from the track to the hold mode, the gate to drain capacitance and the hold capacitor associated with each FET form a differential connection, thereby eliminating the charge induced offset or pedestal error. Figure 44 illustrates this fact. This circuit then becomes sensitive to how well C_1 and C_2 are matched as well as the C_{RSS} of the FETs. Analysis shows that the charge induced offset error is given by:

$$V_{OFF} = V_G \cdot \frac{C_{RSS}}{C} \left(\frac{\Delta C_{RSS}}{C_{RSS}} + \frac{\Delta C}{C} \right)$$

$$= 17.5 (0.5/450) (0.05 + 0.05) = 1.9\text{mV}$$

The sample and hold with the balanced hold capacitor arrangement does not quite meet the goal of a 1.0mV error. Now it remains to be seen what the settling time of this configuration will be. As previously shown, the linearity goal of $\pm 0.01\%$ will be met as the effective voltage swing is not as large as shown in the calculation of V_{OFF} . Since the acquisition time is directly proportional to the holding capacitor, the acquisition time of this faster sample and hold will be 300ns. Once the settling time becomes that fast, other factors that we previously neglected must now be taken into consideration. Monolithic FET op amps are just now becoming available with 300ns settling times so that the calculation of the acquisition time of this architecture must be increased. A good approximation to estimate the combined acquisition time would be to "RMS" the individual settling times. The acquisition time when measured at the output of the operational amplifier would be 424ns. When performing this

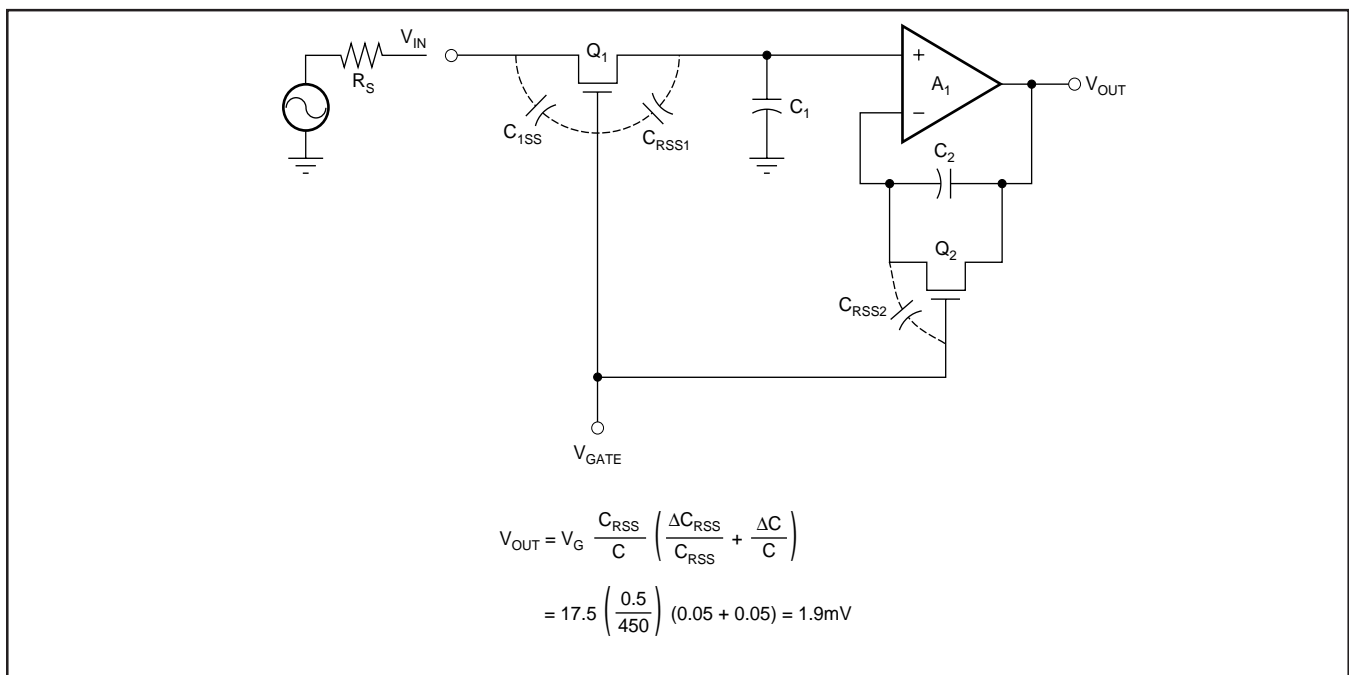


FIGURE 44. Capacitance Mismatch.

calculation for the example in question, a subtlety about acquisition time should be pointed out. The previous two architectures that have been discussed are ones where the sample and hold function is separate from the associated buffer or op amp. This is not true for a feedback architecture that will be discussed later. In some systems applications the distinction between the voltage developed across the hold capacitor and the output of the buffer are important. One important application that will be discussed in detail in the section on analog to digital converters is where the sample and hold drives a sub-ranging ADC. Even though the input signal must be acquired accurately, the voltage does not need to be accurately known to begin the conversion process. Figure 45 shows how the hold capacitor acquires the signal in 300ns while the output of the buffer reaches the same point in 424ns.

One of the drawbacks of this and the previously described circuit is the charge injection of the gate signal through the source to gate capacitance as shown in Figure 44. The nature of the driving impedance can create a great deal of uncertainty as to the nature of the pedestal during the time when the FET is being switched from on to off. The source may be ringing or settling in some unfavorable manner and the track and hold will store the results of the driving source not settling. The track and hold would then benefit from being driven from a buffer to eliminate this problem.

Another problem with the circuit shown in Figure 43 is the poor feedthrough performance. If the feedthrough capacitance is 0.1pF and the hold capacitance is 450pF, the feedthrough voltage could be as much as $(10)(0.1/450) = 2.2\text{mV}$ which exceeds the goal of 1mV. To reduce the feedthrough voltage, the holding capacitor would have to be increased to 990pF. Increasing the hold capacitor to 990pF to reduce the feedthrough voltage would increase the acquisition time across the hold capacitor to 600ns.

This previous calculation shows how the various design parameters can interact and even though one specification is met other specifications must be re-evaluated before the design is complete. It has been shown that even though the arbitrary design goal of 300ns could not be met, this circuit should not be discarded, as improved performance has been achieved. The cost of this sample and hold is relatively modest and could be produced for about \$10-15. As a practical matter, a sub-500ns sample and hold can be very useful when interfacing with a 5 μs ADC and it is desirable to minimize the overall conversion time.

Let us return to the original design challenge, which was to design a sub-300ns sample and hold. Another architecture that is worth considering is shown in Figure 46. This architecture employs the switching FET in the summing junction of an inverting feedback amplifier. The advantage of this connection is that it is possible to drive the FET with a much smaller gate signal which allows the holding capacitor to become smaller. This architecture also has the compensating FET connected in a differential fashion so the circuit is only sensitive to the match of the FETs and is not

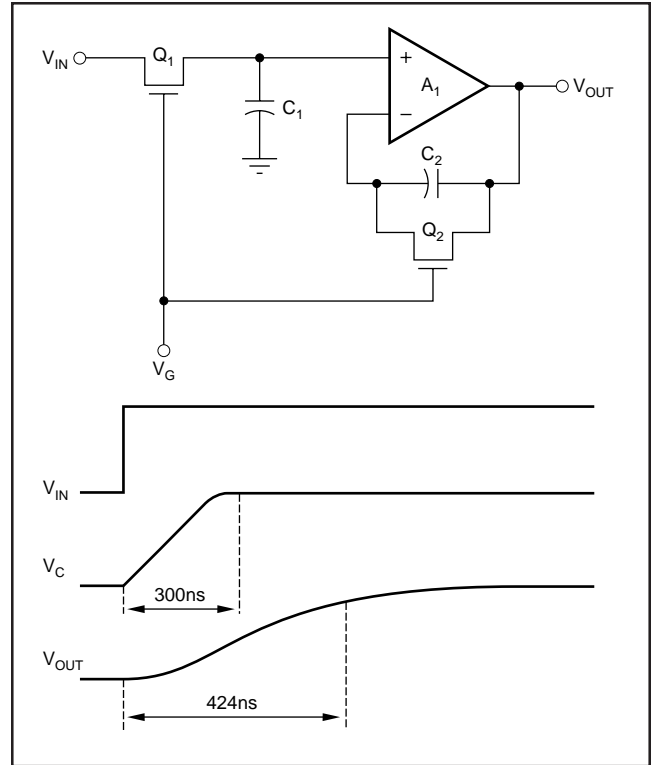


FIGURE 45. Difference in Acquisition Time Between Buffer and Hold Cap.

sensitive to the absolute value of C_{RSS} . Notice the clamping diodes that are placed at the summing junction. These diodes reduce the signal that the FETs have to hold off when the sample and hold is in the hold mode, thereby minimizing the magnitude of the drive signal that needs to be applied to the FET. Circuitry to minimize the feedthrough problem could have also been applied to the previous two design examples, although it would have been much more complicated than the two diodes connected to the feedback track and hold. The gate drive signal can be made smaller since the switch is located in the summing junction of the operational amplifier and the feedback action of the amplifier tends to drive the signal at the summing junction to zero. This also has the effect of linearizing the operation of the track and hold since the charge induced offset pedestal is not signal dependent as it is in the original designs. To ensure that the FET is on, a 5V signal is applied during the sample or track mode and to ensure that the FET is off during the hold mode, a -2.5V level needs to be applied. Therefore the total gate swing will be 7.5V.

Since diodes have been placed at the summing junction, the maximum voltage that can be developed at the input to the FETs is 0.6V peak. From the previous example it was found that it was necessary to have a 500pF holding capacitor to reduce the feedthrough voltage to an acceptable level. Since the feedback track and hold reduces the effective input voltage to the FET to 0.6V, the holding capacitor can be reduced to 60pF. The effective small-signal time constant for this track and hold connection is $T = (2R_{ON} + R_F)C$.

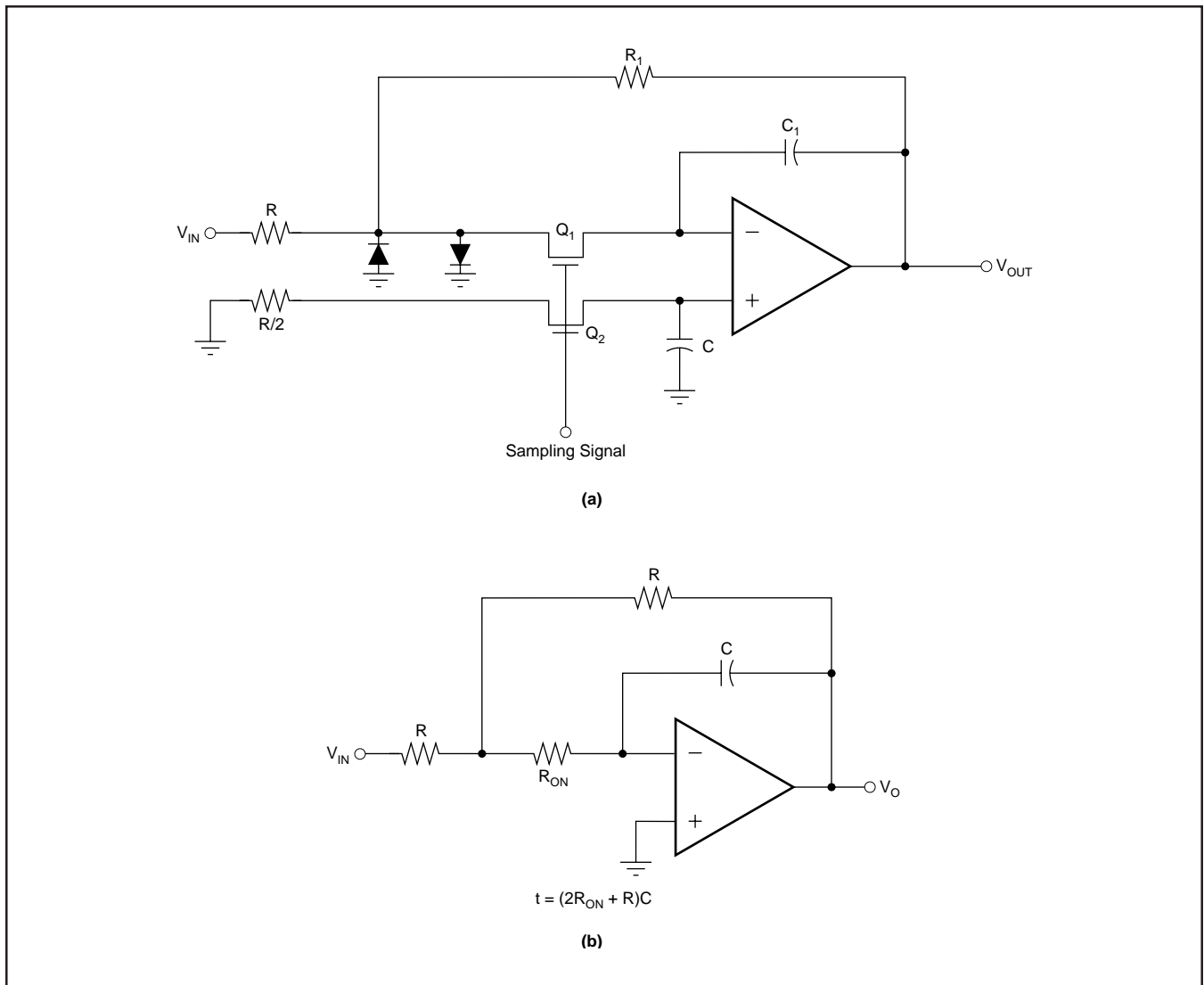


FIGURE 46. Inverting Sample/Hold.

Amplifier will slew until slew rate = $\frac{E}{T}$

$$T = (2 R_{ON} + R_F) C = (2 \times 50 + 300) 60\text{pF} = 24\text{ns}$$

$$E = T \times \text{Slew Rate} = 24 \times 10^{-9} \times \frac{200\text{V}}{10^{-6}} = 4.8\text{V}$$

$$\text{Acquisition Time} = \frac{\text{Input} - E}{\text{Slew Rate}} + T \ln \frac{E}{\text{Error}}$$

$$= \frac{10 - 4.8}{200\text{V}/\mu\text{s}} + 24\text{ns} \ln \frac{4\text{V}}{0.001\text{V}}$$

$$= 26\text{ns} + 203\text{ns} = 229\text{ns}$$

Feedthrough = $0.6\text{Vp-p} \times \frac{0.1\text{pF}}{60\text{pF}} = 1\text{mVp-p}$

Input Clamped by Diode

Feedthrough Capacitance

Feedback Capacitance

FIGURE 47. Performance of Feedback Track/Hold.

Assume that the previously mentioned FET was used along with a monolithic amplifier with a $200\text{V}/\mu\text{s}$ slew rate and a 30MHz small signal bandwidth. Let R_F be 300Ω . The amplifier will then slew until the remaining voltage change is within the linear slew rate capability of the op amp. The small signal time constant of this track and hold is then:

$$T = [(2)50 + 300][60] = 24\text{ns}$$

This corresponds to a small-signal bandwidth of 6.6MHz so the small-signal bandwidth of the track and hold will be determined by the external components rather than by the op amp. Therefore, when the remaining voltage that the track and hold has to change is 4.8V , the track and hold will cease to slew. The time consumed in slewing is then $(10 - 4.8)/(200\text{V}/\mu\text{s}) = 26\text{ns}$. (See Figures 40 and 41 for a view of the acquisition time calculation.) The remaining time is given by $(24) \ln(4.8/0.001\%) = 203\text{ns}$; therefore the acquisition of the track and hold is 229ns and the goal of 300ns can be met with the architecture shown in Figure 46. Figure 47 summarizes the performance of the feedback track and hold. While this track and hold configuration is able to achieve a lower

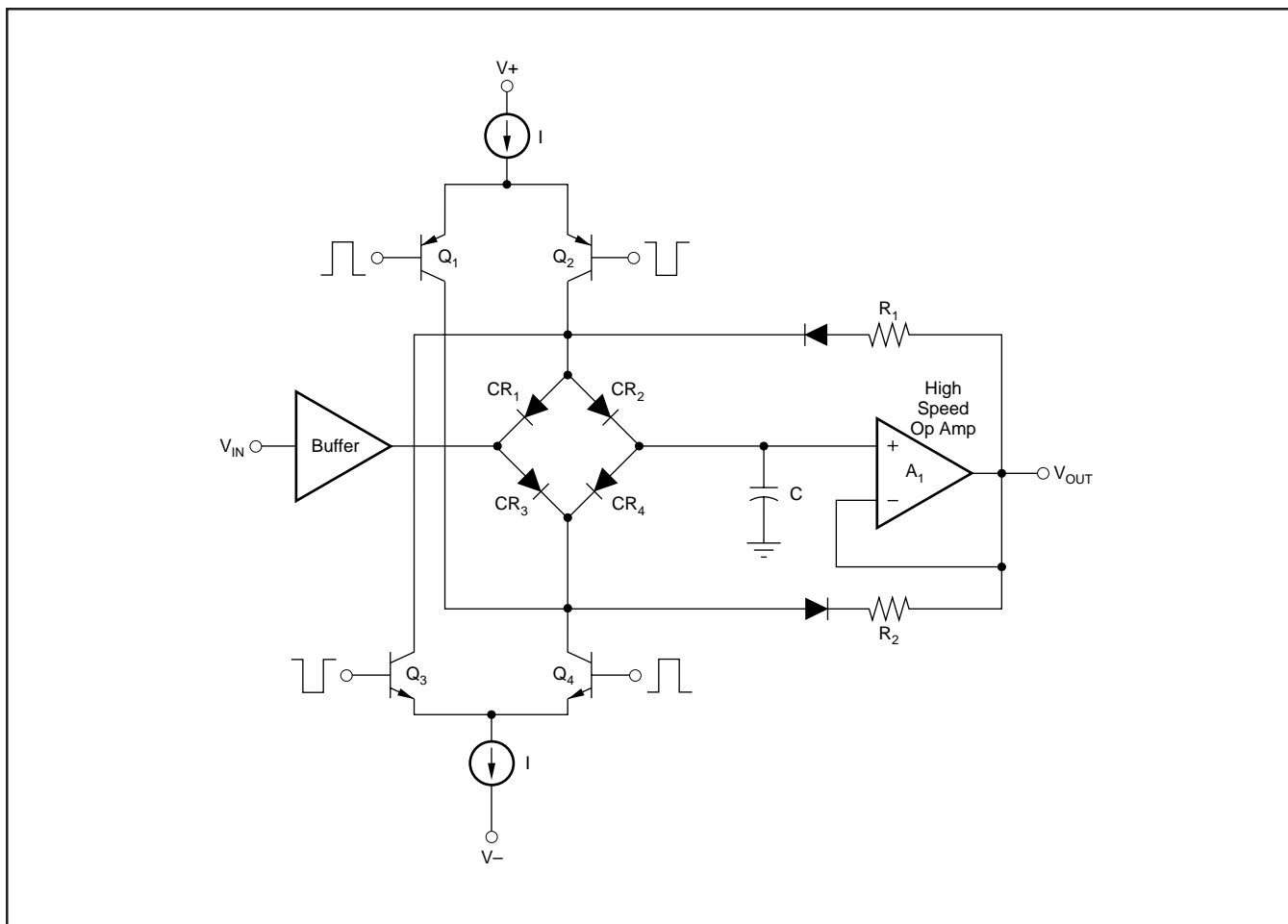


FIGURE 48. Very High Speed Sample/Hold.

acquisition time, it does so at the expense of a lower input impedance. This may not be much of a penalty as the input impedance of 300Ω is within the capability of many op amps to drive with $\pm 5V$ input.

The last track and hold that will be described is capable of acquisition times that are about an order of magnitude faster than the last one that was described. This track and hold, shown in Figure 48, shares some of the architectural features of the previously described ones, although the sampling element is different. This higher speed sample and hold uses hot carrier diodes in a bridge configuration to form the sampling element. Diodes, while more complex to form a sample and hold, achieve high sampling speed due to the lower time constant compared to a FET and lower threshold voltages. As an example, a hot carrier diode operated at $5mA$ has a resistance of 5Ω , V_D of $0.6V$ and a capacitance of $5pF$. Figure 48 shows a diagram of a sample and hold that has an acquisition time of $40ns$ to $\pm 0.02\%$ for a $2V$ step input. This sample and hold has a measured aperture time of under $3ps$. (A technique to measure aperture time is shown in the measurement section.) The sampling function is performed by switching the bridge of hot carrier diodes CR_1 through

CR_4 from the “on” to the “off” state. During the sample mode the current I is steered through the diode bridge by turning on transistors Q_2 through Q_4 . The bridge is returned to the hold mode by turning Q_3 and Q_4 off and turning Q_1 and Q_3 on. The action of turning Q_1 and Q_3 on creates a negative bias on CR_1 and CR_4 . Since these bias voltages are referenced to the output, creating “bootstrap effect,” the reverse bias voltage that diodes CR_1 through CR_4 experience becomes independent of signal level. This is an important aspect of the design as this action prevents the charge offset pedestal from becoming a non-linear function of signal level. An ECL signal is coupled to switching transistors Q_1 through Q_4 . The hold capacitor is isolated from the output by the type of high speed buffers and op amps described in the amplifier section. The sampling bridge is isolated from the analog input signal by a high speed open loop buffer.

As a means of comparison, calculations will demonstrate the different performance parameters of this track and hold. As will be seen from the calculations below, the diode bridge will not achieve as accurate performance as compared to the FET designs.

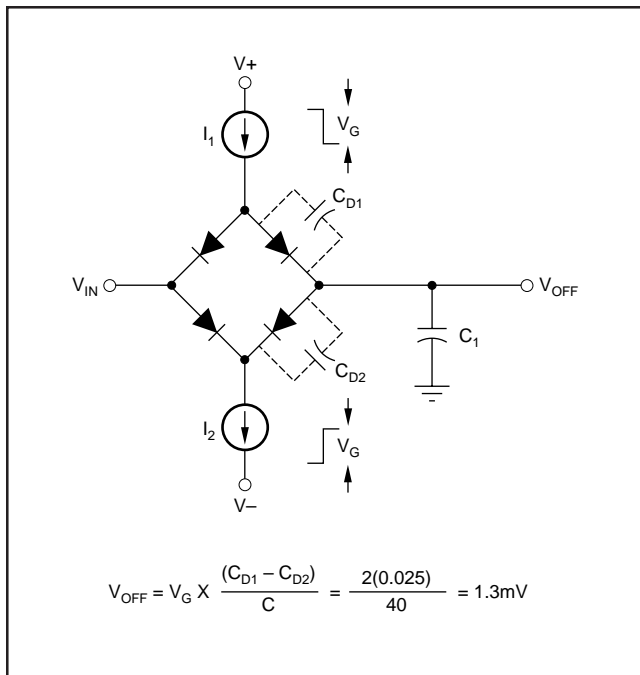


FIGURE 49. Pedestal for Diode Bridge.

CHARGE INDUCED PEDESTAL ERROR

To ensure that the diode bridge is always off, the diode bridge is driven by a complementary signal at the top and bottom of the bridge. The charge induced pedestal error is principally due to the diode capacitance mismatch and is driven by (see Figure 49):

$$V_{OFF} = (V_{OFF})(C_{D1} - C_{D2})/C$$

With care diodes can be matched to 0.025pF. (In actual practice means are provided to adjust the capacitance to this level and once the adjustment is performed the difference in diode capacitance can be held to 0.025pF.) For proper dynamic operation $V_G = 2\text{V}$ and the holding capacitor is 40pF. Substituting into the above equation yields:

$$V_{OFF} = (2)(0.025/40) = 1.3\text{mV}$$

Unlike the FET designs, this is strictly an offset error as the bootstrapping action renders this offset voltage independent of signal level.

SWITCH DELAY PEDESTAL ERROR

The diode bridge switching arrangement has an additional source of error that is not possessed by the FET switch. If the current sources that bias the bridge are not symmetrically switched, the hold capacitor will start to discharge until the other current source is switched. This error manifests itself at the system level as if it were an offset voltage. To a first approximation the cross-coupling eliminates the time mismatching that exists between the NPN and PNP switching pairs. However, due to second order effects as a result of different levels of parasitic capacitances, there is typically as

much as a 50ps mismatch between the complementary signals that switch the bridge. This translates to an offset voltage of (see Figure 50, assume a bridge current of 1 or 5mA):

$$V_{OFF} = I(T/C) = 5\text{mA} (50\text{ps}/40\text{pF}) = 6.3\text{mV}$$

This effect is also largely independent of signal level due to the bootstrapping.

TRACK TO HOLD SETTLING

Track to hold settling is a measure of the recovery time of the track and hold in response to the track and hold going into the hold mode. The previous three architectures are dependent upon the sampling element, whether it is a FET or a diode, to switch exactly with identical waveform shape and turn-off characteristics. In general this will not happen and a small signal will be injected into the buffer op amp. Track to hold settling is a complex calculation and representative waveshapes are shown in Figure 51.

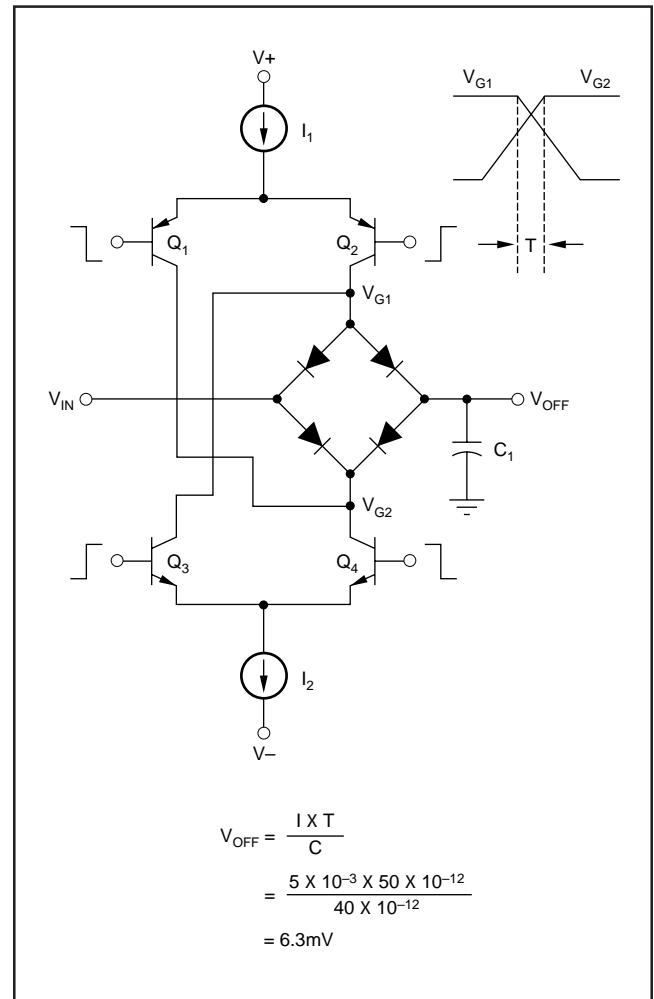


FIGURE 50. Switch Delay.

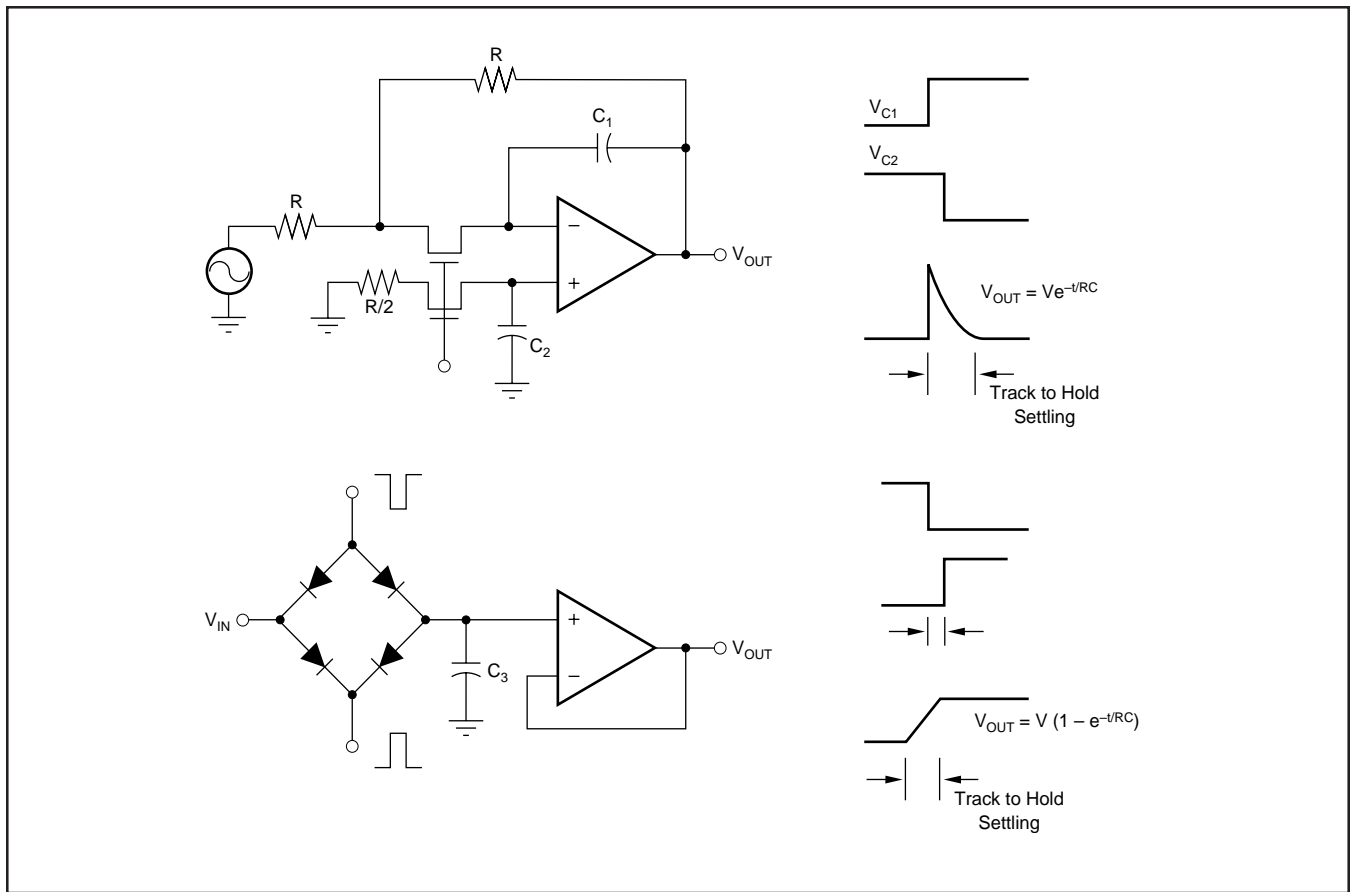


FIGURE 51. Track to Hold Settling.

SIGNAL FEEDTHROUGH

When in the off state, the top and bottom of the bridge are clamped by a low impedance, thereby preventing any signal coupling through that path. Signal feedthrough does occur due to layout and with care a coupling capacitance of 0.01pF can be achieved between the input and output of the bridge. This would yield a feedthrough level of (Figure 52 shows the bridge in the off state):

$$V_{\text{FEEDTHROUGH}} = V_{\text{IN}}(C_c/C) = 2(0.01/40) = 0.5\text{mV}$$

APERTURE JITTER AND DELAY

Aperture jitter of less than 3ps can be achieved and aperture delay of 3ns is also achievable. The lower aperture delay is due to the interface circuitry being wideband ECL.

DROOP

The leakage current that can be achieved with a pair of matched hot carrier diodes is much higher compared to the current levels that can be attained with FETs. Leakage current of 1nA can be achieved with proper thermal level layout. The droop will then be:

$$\text{Droop} = 1\text{nA}/40\text{pF} = 25\mu\text{V}/\mu\text{s} \text{ at } 25^\circ\text{C}, \\ \text{or about } 25\text{mV}/\mu\text{s} \text{ at } 125^\circ\text{C}$$

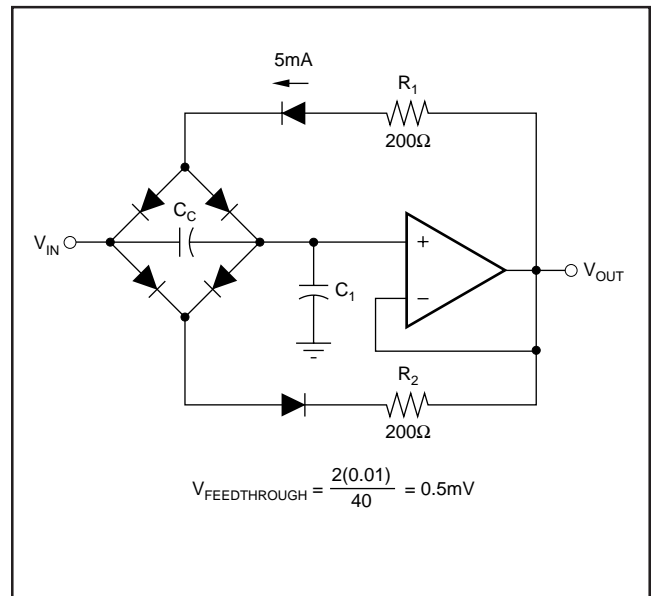


FIGURE 52. Bridge in Off State.

ACQUISITION TIME AND FULL POWER BANDWIDTH

To complete the comparison a calculation will be made of the acquisition time and the full power bandwidth using methods previously demonstrated. The fastest sample and hold is designed to handle only a 2V waveform so the equivalent 0.01% error is 0.2mV. Assume that the amplifier bandwidth is 80MHz with a slew rate of 300V/μs. Figure 53 shows this calculation.

As previously mentioned, one of the most common applications for a track and hold is to precede an analog to digital converter for purposes of reducing the aperture time. Towards the end of the section on digital to analog converters another application will be shown on how a track and hold can be used to “deglitch” a DAC. A third application is how a track and hold can be used to make a precise peak detector. Figure 54 shows the block diagram of a peak detector. The

$$\begin{aligned}
 \text{Amplifier will slew until slew rate} &= \frac{E}{T} \\
 T &= \frac{1}{2\pi B\omega} = \frac{1}{2\pi 80 \times 10^6} = 1.99\text{ns} \\
 E &= T \cdot \text{Slew Rate} = 1.99 \times 10^{-9} \cdot 300\text{V}/\mu\text{s} = 0.6\text{V} \\
 \text{Acquisition Time} &= \frac{\text{Input} - E}{\text{Slew Rate}} + T \ln \frac{E}{\text{Error}} \\
 &= \frac{2 - 0.6}{300\text{V}/\mu\text{s}} + 1.99\text{ns} \ln \frac{0.6}{0.0002} \\
 &= 4.7\text{ns} + 15.9\text{ns} = 20.6\text{ns} \\
 \text{Full Power Bandwidth} &= \frac{\text{Slew Rate}}{(V_{\text{PEAK}})(2\pi)} \\
 &= \frac{300\text{V}/\mu\text{s}}{(1)(2\pi)} = 47.7\text{MHz}
 \end{aligned}$$

FIGURE 53. Acquisition Time and Full Power Bandwidth Calculation for High Speed Sample/Hold.

delay line and comparator serves to form a digital means for locating the point in time where the peak occurs. The output of the comparator allows the track and hold to track the signal until the peak is located. Once the peak occurs the comparator reverses state thereby placing the track and hold in the hold mode, which stores the peak amplitude for further processing.

DIGITAL-TO-ANALOG CONVERTERS

The schematic shown in Figure 56 is typical of the architecture of a high speed digital to analog converter. The digital to analog converter shown in Figure 56 is ECL compatible but shares many of the same elements of TTL compatible DACs as the core current steering mechanism is similar. Most recently, CMOS technology has been used to design high performance digital to analog converters. CMOS DACs have been designed with 12-bit resolution but have not been able to achieve the speeds that can be achieved with bipolar technology. Recently GaAs technology has been used to design exceptionally high speed DACs, with settling times in the 1ns vicinity, and in some ways are similar in topology to the way a bipolar design would be approached. Therefore, describing the design considerations for a high speed digital to analog converter implemented with bipolar technology will serve as a means to understand the design considerations for a high speed, high resolution DAC. Along with the high speed switch, other elements such as the “servo amp” and reference circuitry are also representative of other high precision digital to analog converters ranging in settling times down to 5ns and resolutions to 16 bits. The particular DAC that will be described has 12 bits of resolution with a settling time to ±0.01% accuracy in 25ns and is capable of operating over the temperature range from –55°C to +125°C. This converter is representative of what can be achieved with modern monolithic processing. The DAC is built on a 20V process that contains 1GHz NPNs along with compatible thin film resistors. As will be described later, the thin film resistors are laser-trimmed to achieve true 12-bit linear-

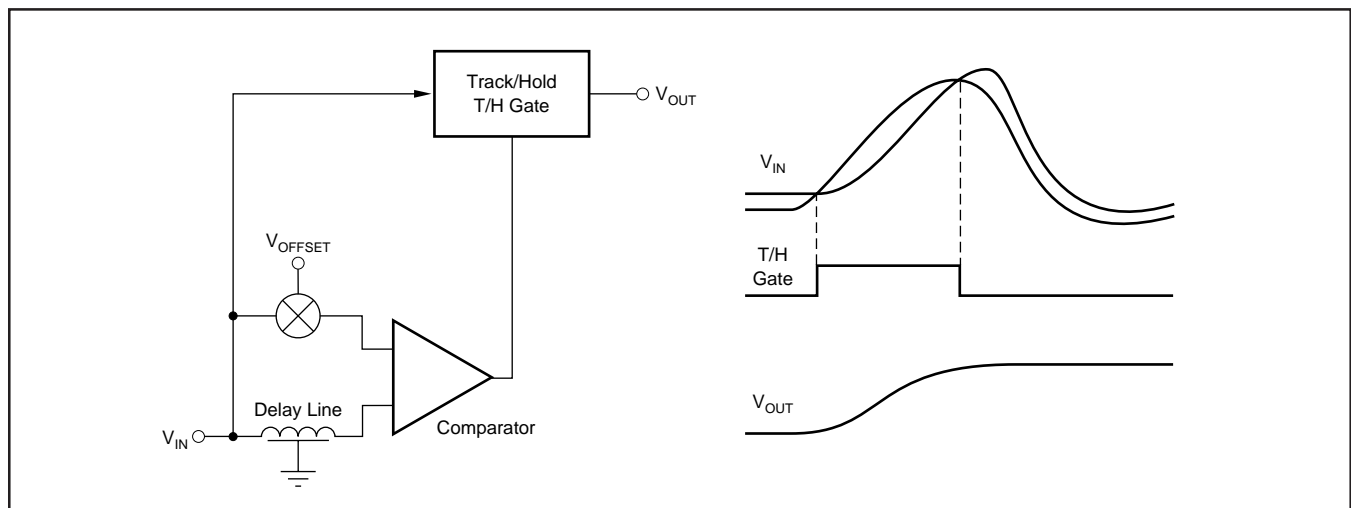


FIGURE 54. Peak Detector.

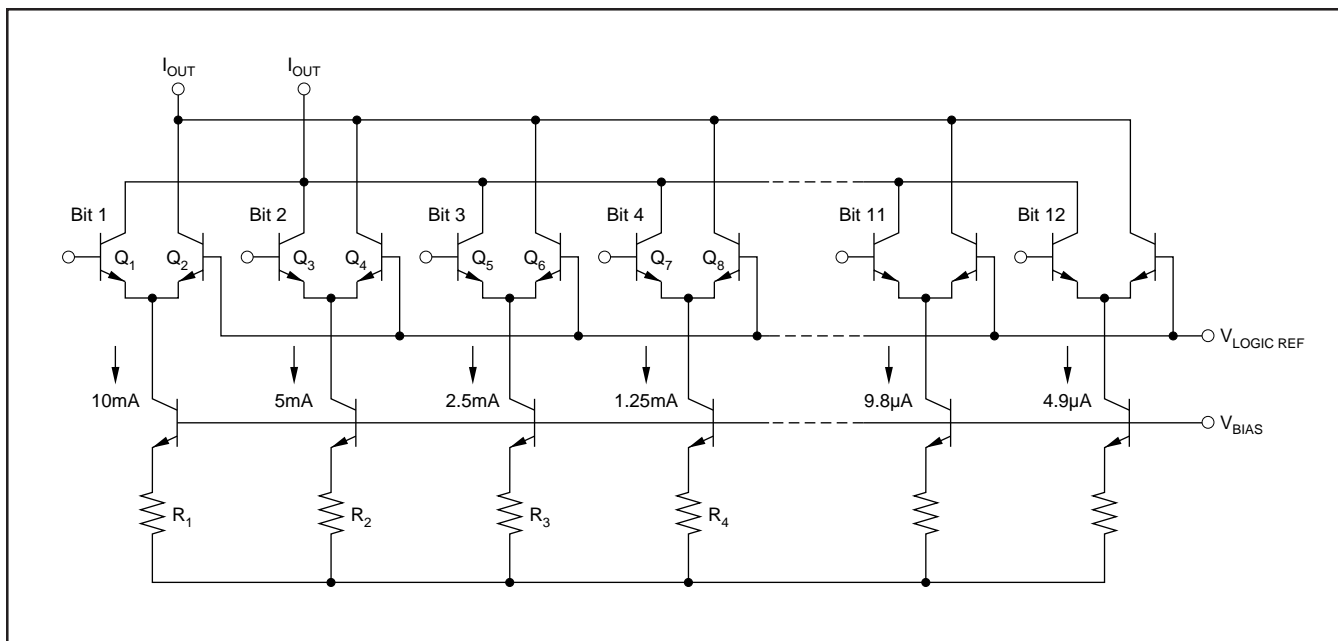


FIGURE 55. Binarily Weighted Current Source DAC.

ity over a very wide temperature range. Furthermore, the thin film resistors are capable of maintaining their accuracy over long periods of time and represent a reliable technique for producing a high speed, high resolution, low cost digital to analog converter. The converter that will be described is entirely monolithic as it contains the precision current switches, servo amp, and low drift references. Only a few capacitors that are too large to be integrated and are needed for filtering and bypassing are left off of the chip.

The converter consists of twelve switches that are driven in a non-saturating manner. In order to steer the current as fast as possible through the output switch, it is very important to pay careful attention to avoid saturation; once a transistor saturates, the recovery time can easily increase by a factor of twenty or more. There are many ways to approach the design of this kind of a DAC. The detailed design considerations will be described, but before that explanation will be given, an overview of different DAC architectures is offered. One method would be to binarily weight the individual bit switches and then sum the outputs as shown in Figure 55. High accuracy can not be achieved using this method as it is difficult to accurately match the separate current sources and switches over such a wide range of currents. If the full scale output current of the 12-bit DAC were 10mA, the weight of the LSB would be 4.9μA which would be too low to achieve high speed switching. Additionally, with all those switches in parallel, the output capacitance would become quite high. The only redeeming feature of a binarily weighted DAC is that there would be no wasted current and the net power dissipation for this type of digital to analog converter would be the lowest as compared to other design approaches.

Another way to approach this design would be to have twelve equally weighted current switches. The twelve equally

weighted current switches would then be binarily weighted by passing their currents through an R-2R ladder as shown in Figure 56. Twelve equally weighted current sources could then be precisely matched using a “servo mechanism” control loop as shown in Figure 57. The servo loop is able to cause the value of the output current to be exactly (within circuit tolerances) the same as the reference current. A reference current is connected to the positive input of the op amp and the collector of transistor Q_1 . The same reference current then passes through Q_1 and emerges as emitter current by the addition of base current. The emitter current of Q_1 then becomes the collector current of Q_2 . The voltage current developed across the base to emitter junction of Q_2 and the voltage drop across R_1 create an identical current through Q_3 . The collector current of Q_3 becomes the emitter current of Q_5 which in turn emerges from Q_5 as the output current. Examination of the analysis shown in Figure 57 shows that if all the transistors and resistors are well-matched, the output current will be equal to the reference current. This is an ideal technique to be implemented in a monolithic process, as it is very practical to make transistors and resistors identical. A more detailed analysis of error sources will be shown later. A digital to analog converter designed in this manner would have the lowest glitch performance but at the expense of the highest power dissipation. “Glitch” refers to the uncertain DAC output that occurs when the digital input changes and the DAC switches do not change simultaneously. More will be given on the design of low glitch DACs toward the end of the section on high speed digital to analog converters.

Practical digital to analog converters are a mixture of the two previously described examples as shown in Figure 60. Starting with the MSB (most significant bit), the currents are

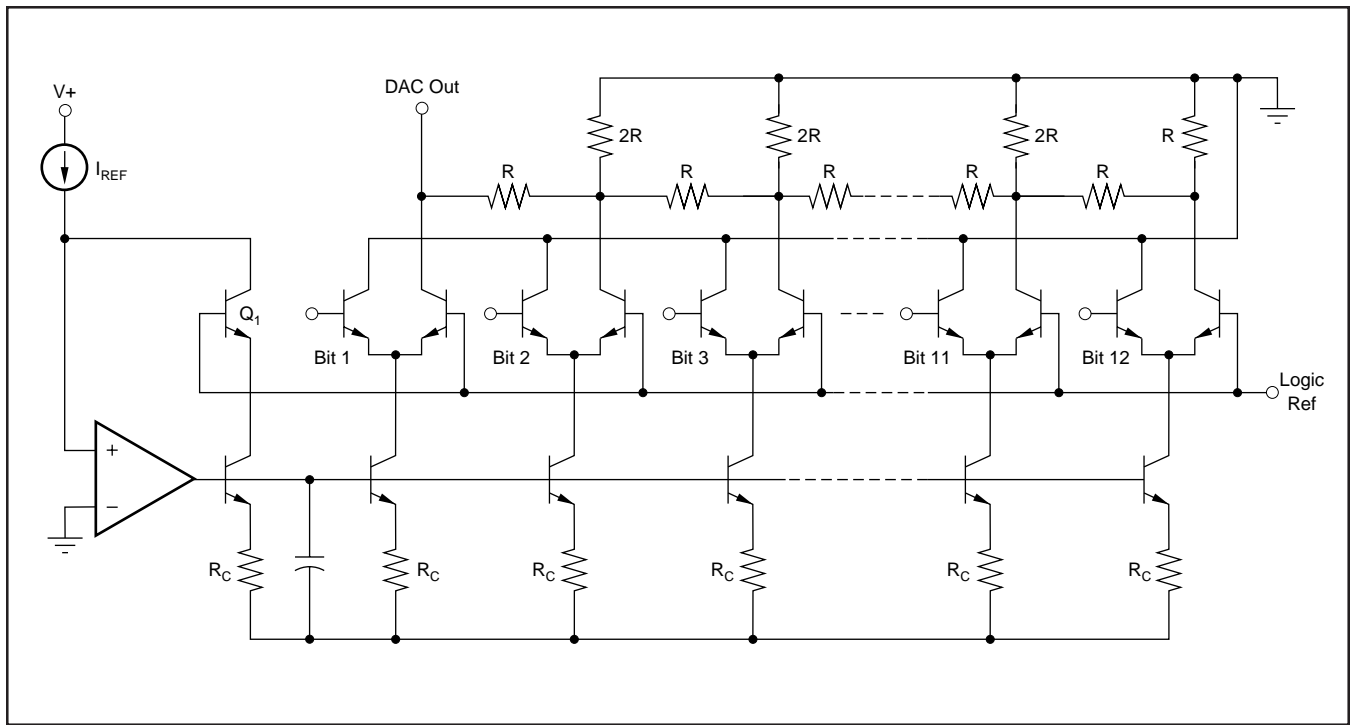


FIGURE 56. High Speed DAC with Equally Weighted Currents.

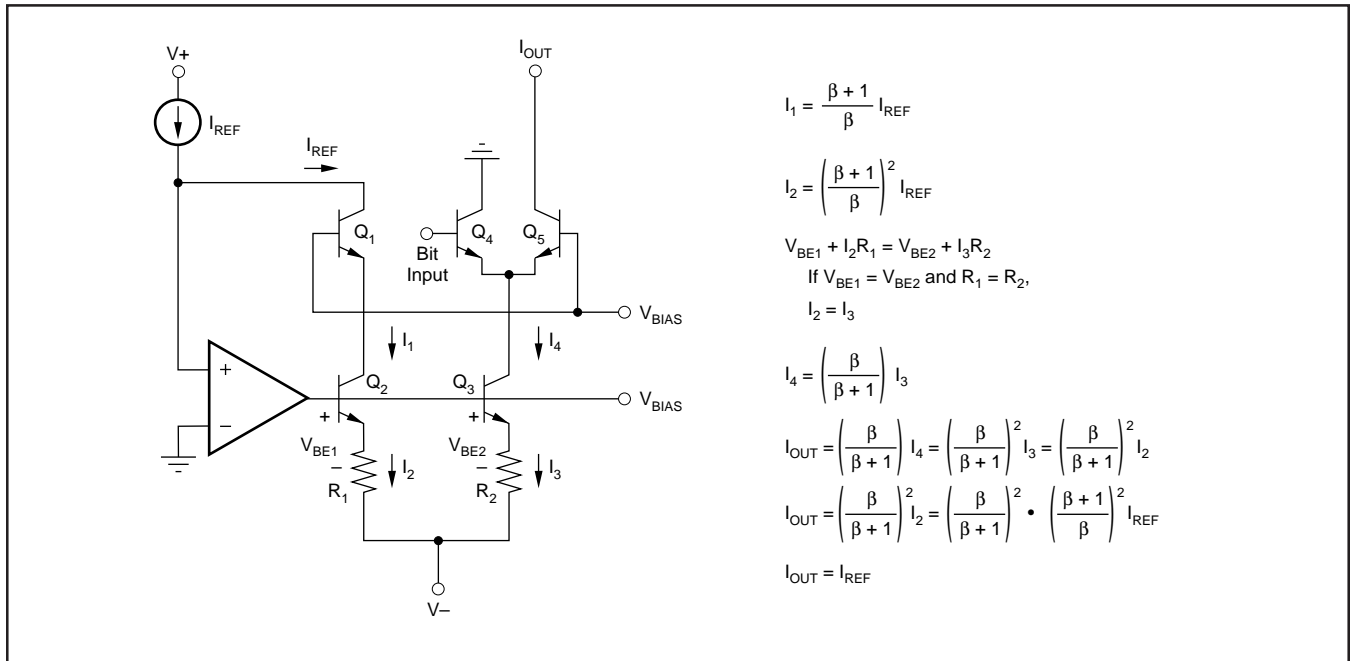


FIGURE 57. DAC Servo Loop.

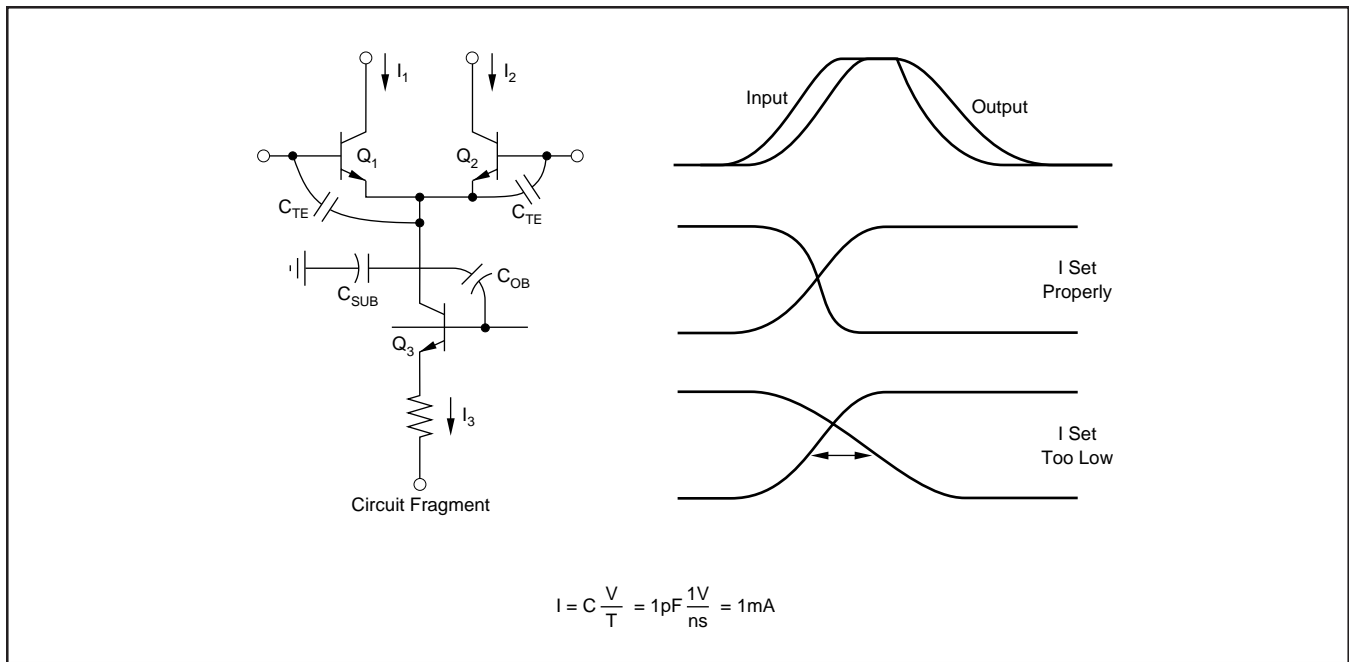


FIGURE 58. DAC Switching.

binarily weighted until the current becomes low enough to effect switching speed. Even though the MSB currents are not the same value as the LSB currents, matching is maintained as the current density is made to be the same. The current density is maintained by making the transistors that are to conduct larger currents physically larger, thereby causing the voltage drop associated with the transistor to be the same. This is similar to placing transistors in parallel.

Figure 58 shows a diagram depicting the switching of one DAC current switch. This type of an emitter coupled pair is capable of switching very rapidly in response to a positive going input logic change as Q_1 acts as an emitter follower which is capable of driving the capacitance attached to the common mode where the emitters of Q_1 and Q_2 are joined. If the base of Q_1 is driven from an ECL input the speed at the emitter of Q_1 is determined by the rate of change of the ECL input of about 1V/ns. To obtain low glitch performance it is necessary to have the DAC propagation delay to be equal for negative as well as for positive logic changes. Therefore, when the logic input makes negative going transition the current supplied by current source Q_3 will have to provide enough current to drive the node capacitance to allow the voltage change at the emitters to track the negative going input signal. Some digital to analog converter designs will drive the current switch differentially, which means that either side of the switch is capable of actively driving the node capacitance instead of depending upon the current that is being switched. While this approach solves the problem of providing high current for the lower order bits, it does so at the expense of providing a differential driver. If done so externally, the chip would have to have an additional twelve inputs as well as require that the user supply differential inputs. Alternatively, a differential driver could be placed on

the chip at the expense of loss in speed and extra circuit complexity. This particular approach is taken to emphasize simplicity. Returning to the design at hand: the amount of current necessary to follow a negative going logic change is given by the formula (see Figure 58):

$$I = C(V/T)$$

where C is the total node capacitance and (V/T) is the rate of change of the logic input.

Substituting:

$$I = 1\text{pF}(1\text{V/ns}) = 1\text{mA}$$

An extra amount of current is provided to assure equal propagation delays in both directions so that the minimum current that is set for the lower order bits is 1.25mA. The MSB current switch is scaled to be four times this value of 5mA and the next bit, Bit 2, is scaled at twice the minimum, or 2.5mA. Bits 3 through 12 are then set at 1.25mA. Bits 1 through 3 are connected together while Bits 4 through 12 are passed through the R-2R ladder to establish the proper binary weighting. In order to maintain high accuracy, Bits 1 and 2 are also physically scaled. Physical scaling can be thought of as placing unit current switches in parallel, thereby allowing proper matching and compensation by the servo amp. Figure 59 shows how this is done.

Typically the output resistance of the ladder is 250Ω (see Figure 60), so the DAC output voltage swing will be 2.5V. For greatest flexibility there is a resistor connected to the positive reference that allows that DAC output to be able to swing $\pm 1.25\text{V}$ around ground. In order to accommodate the negative level of -1.25V , care must be taken so the output transistor is not saturated. Figure 61 shows a circuit diagram that includes the parasitic collector resistances which must

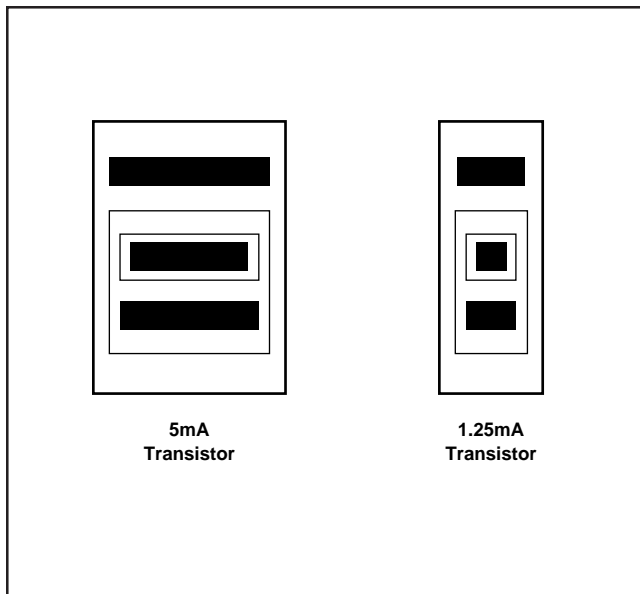


FIGURE 59. Scaled Transistors.

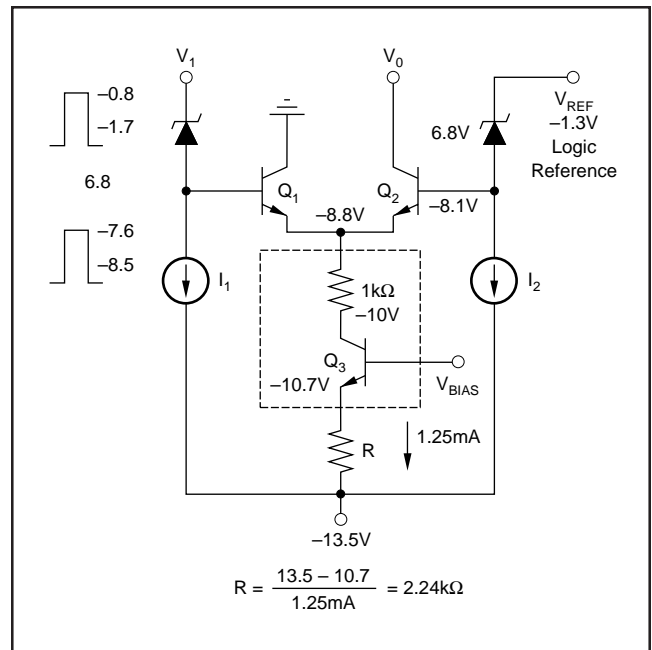


FIGURE 61. DAC Bias Voltages.

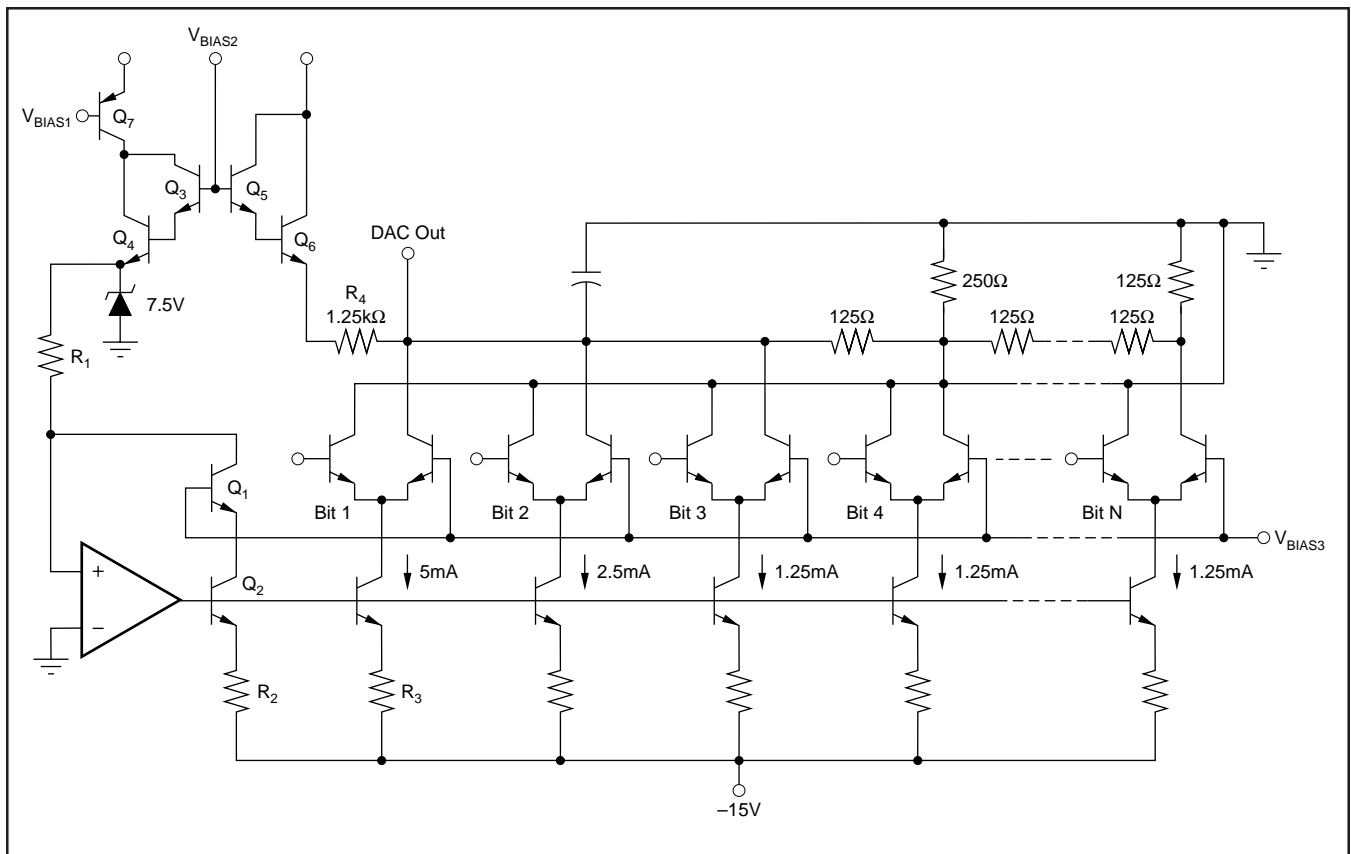


FIGURE 60. Practical DAC.

be accounted for. This diagram also shows the translation zener so that the DAC switch can properly interface to the ECL level inputs. A 6.8V zener is a useful voltage translation device as the impedance level is low, about 50Ω, which is necessary for maintaining high speed. When the ECL level is at a low of -1.7V, the voltage at the base of Q₁ will be -8.5V. Since the voltage at the base of Q₂ is -8.1V, Q₁ will be off and Q₂ will be on. Even though Q₁ is slightly forward biased, the amount of conduction is tolerable for 12-bit applications. The voltage at the emitter is then -8.8V when Q₂ is on. The saturation resistance for the transistors used in this DAC design is 1kΩ and since the current level is 1.25mA, the voltage at the actual collector is -10V. Under worst case conditions, the base voltage should not be allowed to become greater than -10V or the onset of saturation will begin. Therefore, under these conditions the emitter of Q₃ will be at -10.7V. It is always desirable for maximum accuracy (as will be shortly seen) to create as large an emitter degeneration voltage as possible. The largest voltage tolerable will be when the emitter voltage is -10.7V and the power supply voltage, which is nominally -15V, is at its lowest of -13.5V. Under these conditions the emitter degeneration voltage will be: 13.5 - 10.7 = 2.8V. The emitter degeneration resistance will then be (2.8V)/(1.25mA) = 2.24kΩ.

Refer to Figure 62, which shows an analysis of the principal error-producing elements of a typical DAC switch and current source. There are three error sources that can be eliminated by adjustments after the DACs are assembled and two sources of error that must be eliminated by design. The three sources of error that can be adjusted to zero or "trimmed out" are the beta and V_{BE} matching of the transistors and the matching of the thin film resistors. While these error producing effects can be corrected at room temperature, they will change over temperature. As an example: The beta of a transistor will be assumed to be 150 and to have a temperature coefficient of +7000ppm/°C. An uncompensated transistor collector current will experience a beta error of (7000/C)/150 = 47ppm/°C. This means that at room temperature the ratio of the collector to emitter current will be 150/151 = 0.99348 and at 125°C the ratio will be 0.99609. Due to the compensation action of the servo-loop, experience has shown that a further reduction by a factor of 200 can be attained so that the net drift over temperature due to this effect is 0.24ppm/°C. If transistors Q₂ and Q₃ are carefully matched, their V_{BE}s will track each other to 1μV/°C and the effect upon the accuracy of the switch will be (1μV/°C)/2.8V = 0.36ppm/°C. Lastly, resistor matching of 0.5ppm/°C is achievable if the resistors are laid out properly. Adding these three effects yields a net current source drift over temperature of:

$$\begin{aligned} \text{Change over temperature} &= (2)\text{Beta} + V_{BE} + \text{Resistor} \\ &= 0.47 + 0.36 + 0.5 = 1.33\text{ppm/}^\circ\text{C} \end{aligned}$$

Assuming that these errors can be laser-trimmed to arbitrary accuracy at room temperature, any bit switch over a 100°C temperature change will experience a 133ppm change which

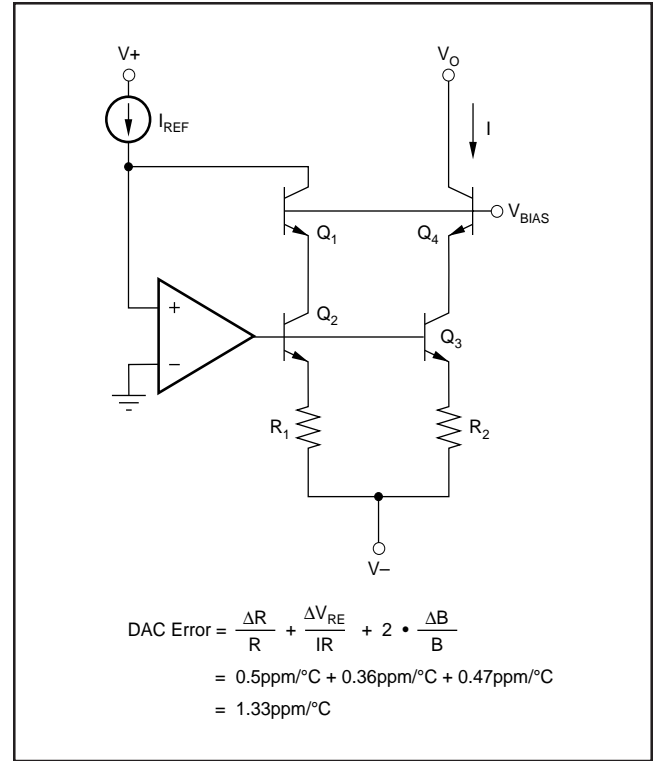


FIGURE 62. Trimmable DAC Errors.

implies that if these assumed tracking values were attained it would be difficult to produce with high yield a DAC that had ±1/2LSB linearity. The art of building high accuracy digital to analog converters is the ability of design, layout, processing, and manufacturing engineers to control the previously described elements to sufficient accuracy. In fact it is possible to manufacture 12-bit DACs from -55°C to +125°C within ±1/2LSB accuracy and the above assumed parameters are achievable.

There are two other sources of error that can only be eliminated by proper design: the output impedance and superposition error. Due to the cascoded nature of the DAC switch, the output resistance is given by (see Figure 63):

$$\begin{aligned} R_{OUT} &= (\text{Beta})(V_A/I) \quad (V_A = \text{Early voltage}) \\ &= (150)(200/10) \\ &= 3\text{Meg} \end{aligned}$$

Since the ladder impedance is 250Ω, the output resistance represents an error of 83.3ppm which is below the error budget of 122ppm needed for a 12-bit design. The output resistance causes a non-linear error since there is a difference in this value when the switch is on compared to the off value.

The last error source that needs to be considered is superposition error. Superposition error occurs when the individual bits do not add up to the proper sum defined by their values when they are individually turned on. Superposition error can have many causes, but one of the most prevalent causes for the type of DAC being discussed is the offset created by

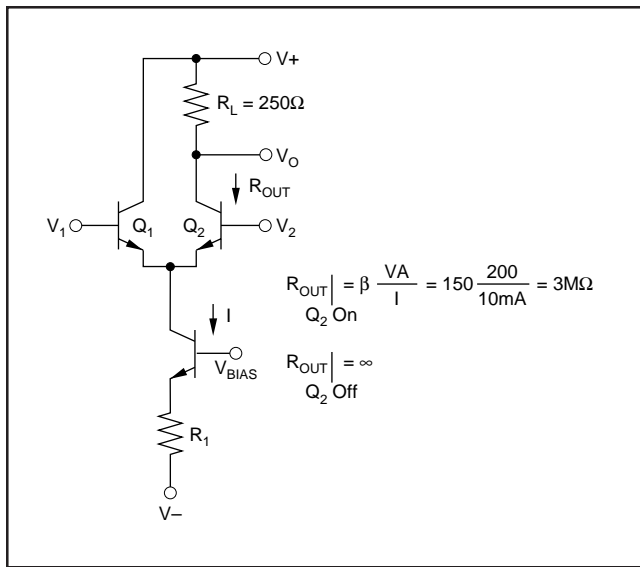


FIGURE 63. Error Due to R_{OUT} .

the resistance in the return line of the R-2R ladder. Figure 64 illustrates how the DAC offset can be markedly different as a function of how many bit switches are turned on. If any individual bit is turned on, the offset will be equal to the bit current multiplied by the value of the resistance in the ladder return. As long as only one bit switch is turned on the offset voltage will be constant. However, when multiple bit switches are turned on this error will not be constant. Take the case when the DAC makes a 1LSB transition around the MSB. When the MSB is on, the offset voltage is the small value defined by only one switch being on. However, 1LSB below

the MSB occurs when the lower eleven bits are turned on with the MSB off, and this offset voltage will now be eleven times greater compared to when only the MSB is turned on. This effect can be minimized by making the ladder return impedance as small as possible and by returning the opposite side of the bit switch back to the same point as the on side is to be returned to. This has the effect of keeping this offset voltage constant for any digital code combination. It is important to sense the voltage at the true reference point on the ladder to achieve maximum when the DAC is trimmed at the factory level.

This digital to analog converter has the capability of settling to $\pm 0.01\%$ accuracy for a full scale change in about 26ns. The settling time is primarily determined by the ladder impedance and the total capacitance that is accumulated on the output node. The combined capacitance of the R-2R ladder, the offset resistor, the output transistors, and the load capacitance is about 10pF. The propagation delay from the digital input to the actual current switch is 3ns. The remaining part of the settling time is due to the voltage settling of the output time constant formed by the ladder impedance of 250Ω and the node capacitance of 10pF which forms a 2.5ns time constant. Settling to $\pm 0.01\%$ accuracy requires $(2.5) \ln(1/0.01\%) = 23\text{ns}$, and when the digital propagation delay is added to the voltage settling, the total becomes 26ns. Achieving fast and accurate settling times requires paying attention to several other aspects of the design that will be described. Improperly designed reference and servo-amp circuitry can lead to a DAC that will not achieve the previously calculated 26ns. Figure 65 shows a path of how the digital input coupled onto the fence line which deter-

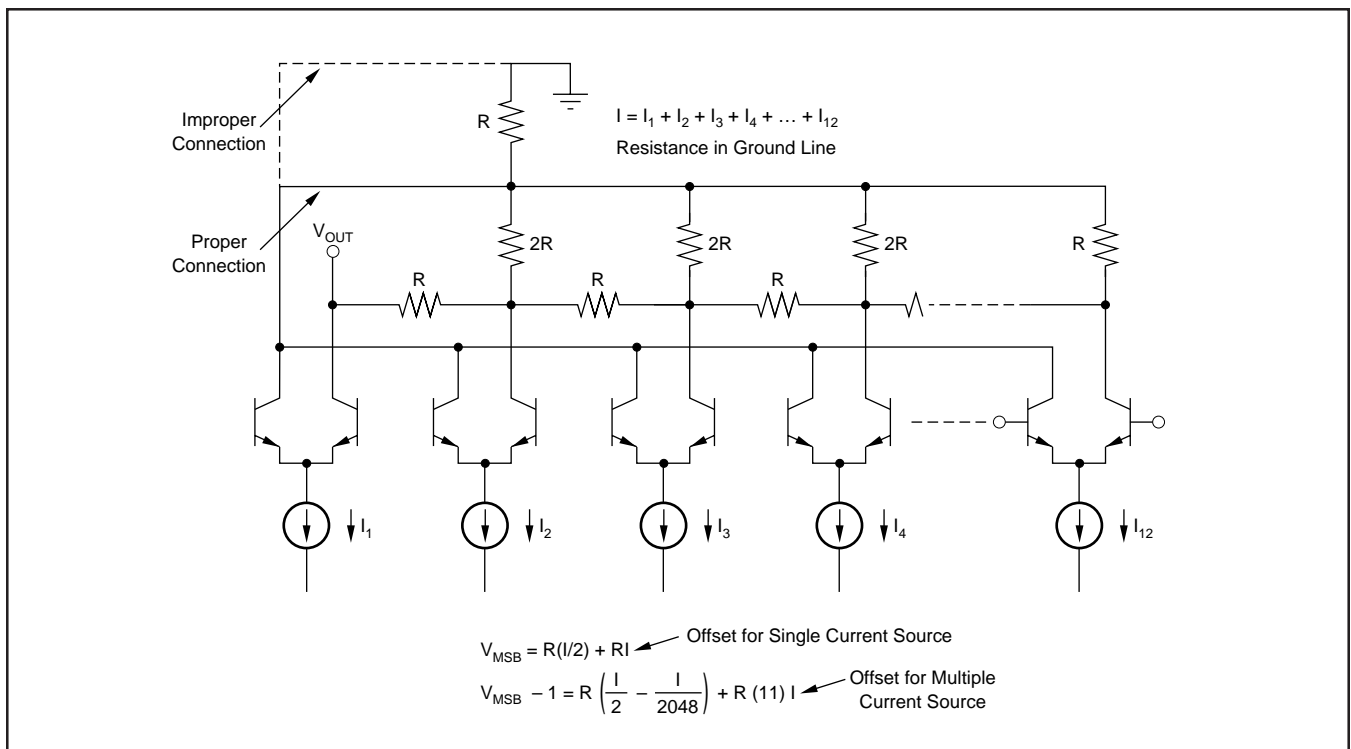


FIGURE 64. Superposition Error.

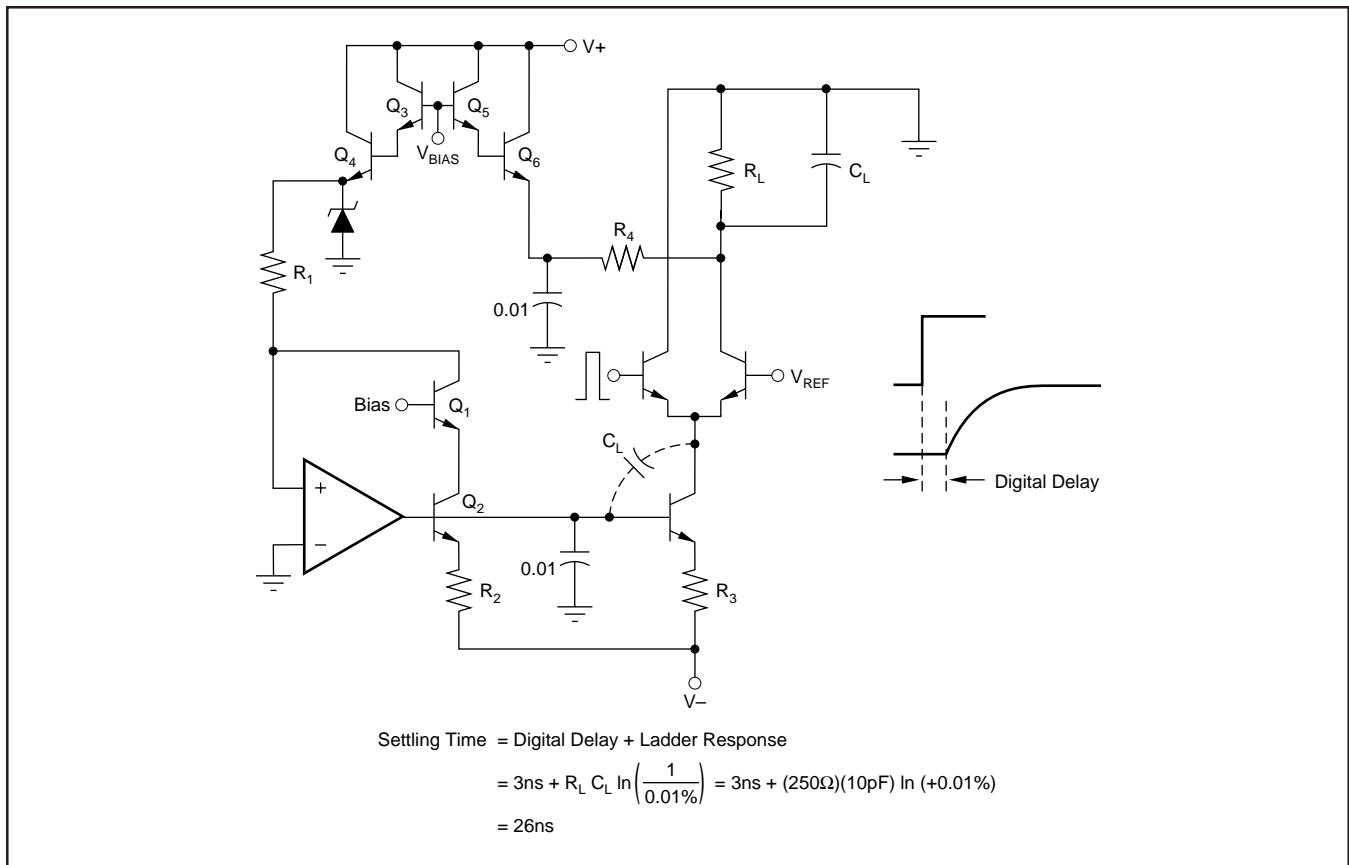


FIGURE 65. Settling Time.

mines the value of the bit currents. The servo operational amplifier would have to have a bandwidth of 25MHz to 50MHz to be able to respond and to settle to the capacitively injected transient onto its output. This is not a practical requirement since only the fastest op amp can settle in 24ns even if separately designed and not part of a DAC chip. A more practical solution would be to place a 0.01pF capacitor on this reference line to absorb the transient and then design a low frequency op amp that was stable.

The switches that form the DAC are unipolar and in its natural form the DAC has an output that swings from ground to some negative voltage. For maximum flexibility it is desirable to have a bipolar DAC which requires a means of translating the output voltage in a positive direction. This is accomplished by connecting a 1.2kΩ resistor back to the 7.5V reference voltage. Examination of Figure 65 reveals a buffer compound emitter follower that is used to isolate the low current offset resistor. This buffer isolates the reference offset current change from entering the low bandwidth reference and servo-amp circuitry. A 0.01µF capacitor is added to the isolation circuit to prevent transients from entering the low frequency servo-loop.

Ordinarily a designer would not consider the use of an ECL DAC but there are several reasons that the DAC previously described will have superior performance compared to a TTL DAC. Briefly, ECL has a lower logic delay than TTL,

is less noisy, and ECL data registers have lower data skew. Data skew occurs when all the digital inputs do not change at exactly the same time and is defined as the difference between $T_{PD}(+)$ and $T_{PD}(-)$. $T_{PD}(+)$ is the positive going propagation delay while $T_{PD}(-)$ is its negative counterpart. As an example of this phenomenon, consider the major carry change for a 12-bit DAC. For a 1LSB change around the MSB, the code would change from 0111 1111 1111 to 1000 0000 0000 under ideal conditions. With the presence of data skew all bits might not change at the same time and an intermediate code could exist. Consider what happens if the MSB changed more rapidly compared to the rest of the bits, so that the code transition pattern would be:

0111 1111 1111 → 1111 1111 1111 → 1000 0000 0000
code before intermediate code code after

See Figure 66 which shows a timing diagram depicting data skew. Therefore, for a period of time equal to the data skew, the DAC output would start to head in the direction of an output that was considerably different than a 1LSB change from the previous code. This large transient-like waveform that is created by data skew is often referred to as a DAC output “glitch.” A convenient way to specify the glitch is by measuring the area of the glitch in units of LSB-ns. This is a more effective method for specifying the glitch than if it were defined as a voltage amplitude, as one could not

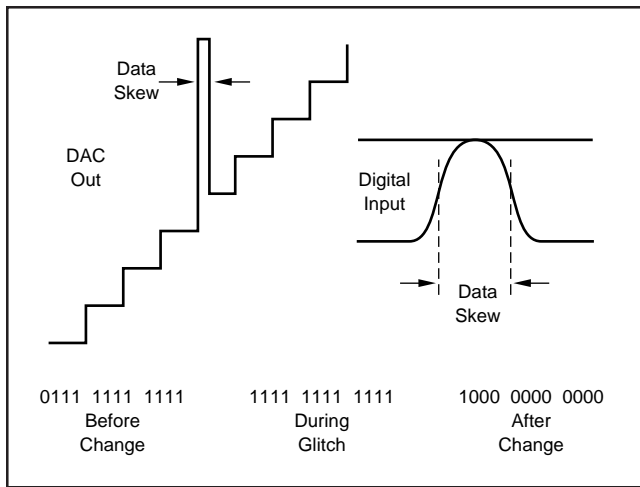


FIGURE 66. Skew.

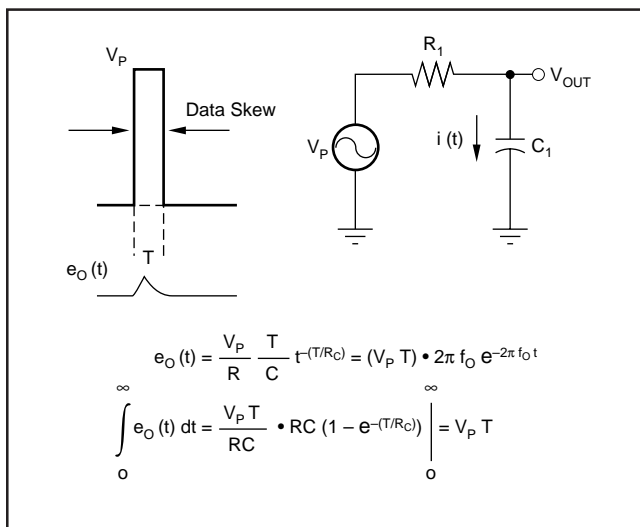


FIGURE 67. Glitch Response.

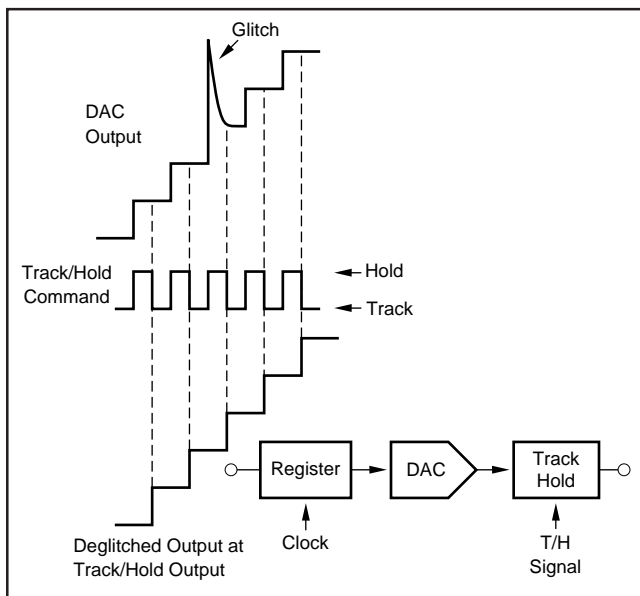


FIGURE 68. Deglitched DAC.

compare DACs with different full scale output levels. Additionally, if the DAC's output were processed by a lower bandwidth amplifier, the peak amplitude of the glitch would change but the area under the curve would not. Assume that the glitch response of the DAC is a pulse with a width equal to the data skew and with an amplitude of 1/2 full scale. Figure 67 shows that the area under the time response of the glitch is constant. Examination of the time response of the glitch indicates that the peak glitch amplitude is a function of the bandwidth of the amplifier. For that reason, a more reliable way to specify the glitch performance of the DAC is by the ET product, or in LSB-ns as ET is independent of the bandwidth. Further note that the average value of $e_O(t)$ is equal to ET and is independent of the bandwidth of the amplifier. What this means is that, as the bandwidth is reduced, the peak amplitude will diminish but the effect of the glitch will last longer.

An ECL DAC will generate a lower glitch than a TTL DAC but there are systems where the glitch has to be further reduced. For these applications the DAC is followed by a track and hold as shown in Figure 68. The track and hold is placed in the hold mode prior to the register being clocked. After the register is clocked and the digital to analog converter is allowed to settle, the track and hold is then placed back into the track mode. In many systems the non-uniform nature of the glitch response creates distortion and harmonics and even though the track and hold may actually have a greater glitch than the DAC, the glitch is uniform for all code combinations and will manifest itself at the system level as an offset or gain error but not as a code dependent non-

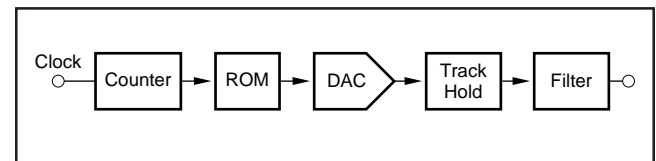


FIGURE 69. Arbitrary Waveform Generator.

linearity.

Figure 69 shows a system with a high speed digital to analog converter that can be used to generate a precise arbitrary waveform. While there are many ways to accomplish this with lower frequency circuitry, the use of a high speed DAC is an attractive alternative. A high frequency DAC is capable of being updated at a 50MHz rate, which will substantially ease the subsequent analog filtering requirements. Since the waveform is effectively sampled at a 50MHz rate it would be possible to create a waveform with frequency components up to the Nyquist rate of 25MHz. Generating an arbitrary waveform is the inverse of digitizing a waveform with an analog to digital converter and the same sampling considerations apply. Figure 70 shows an arbitrary analog waveform that is to be synthesized. If the waveform were sampled at periodic intervals, the synthesized waveform would be created. The synthesis procedure consists of math-

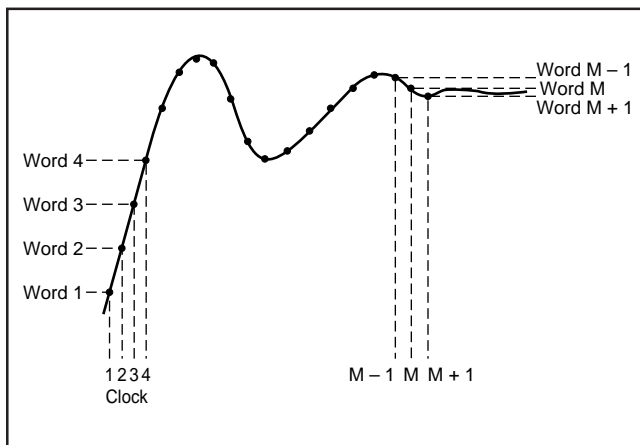


FIGURE 70. Arbitrary Waveform.

ematically computing the closest 12-bit approximation to each sample point which would be used to generate the encoding table for the ROM. Refer to Figure 69 which shows a simplified block diagram of a system that will generate a synthesized waveform. The sample points would correspond to the ROM address while the ROM output would be the associated code at each one of these addresses.

DIFFERENT HIGH SPEED ADC ARCHITECTURES

This section will compare the performance features, and trade-offs, of three commonly found architectures of high speed analog to digital converters to gain an understanding of how resolution, speed, and complexity interact in the design of an analog to digital converter. These three architectures form the basis of most high speed ADCs that are on the market, although there are many variations of these basic circuit arrangements due to the nature of particular technologies. It is useful, though, to gain an understanding of the

architectures in their most elementary form before an appreciation of the variations can be gained. Each architecture has distinct characteristics that need to be properly understood to maximize the benefits of the chosen analog to digital converter with the application.

The three types of designs that will be compared are flash, successive approximation, and sub-ranging. Each method of conversion has strengths and weaknesses which will be clearly contrasted. This section will compare the relative merits of each converter with respect to accuracy, dynamic characteristics, aperture effects, simplicity, and cost. A description of each analog to digital converter will first be given which will then be followed by the performance features of each architecture.

FLASH ADC

The fastest of all types of high speed analog to digital converters, and perhaps the easiest to understand, is the flash or parallel type of converter. The flash converter is considered to be the fastest because the conversion takes place in a single cycle, hence the name "flash." The resolution of flash converters is typically 8 bits, although expensive or experimental designs have been reported with up to 10 bits of resolution. Flash converters are very appealing to monolithic designers due to the highly repetitive nature of the design. Refer to Figure 71 which shows a block diagram of a flash converter. Speeds of up to 500MHz have been achieved and conversion times of up to 200MHz are readily available on the commercial market. Bipolar technology is used for the fastest designs with CMOS achieving conversion rates of up to 30MHz. The resolution of a flash converter tends to be limited to 8 bits due to the fact that the amount of circuitry doubles every time the resolution is increased by 1 bit. The input comparators are arranged in a "thermometer" code fashion with each comparator's refer-

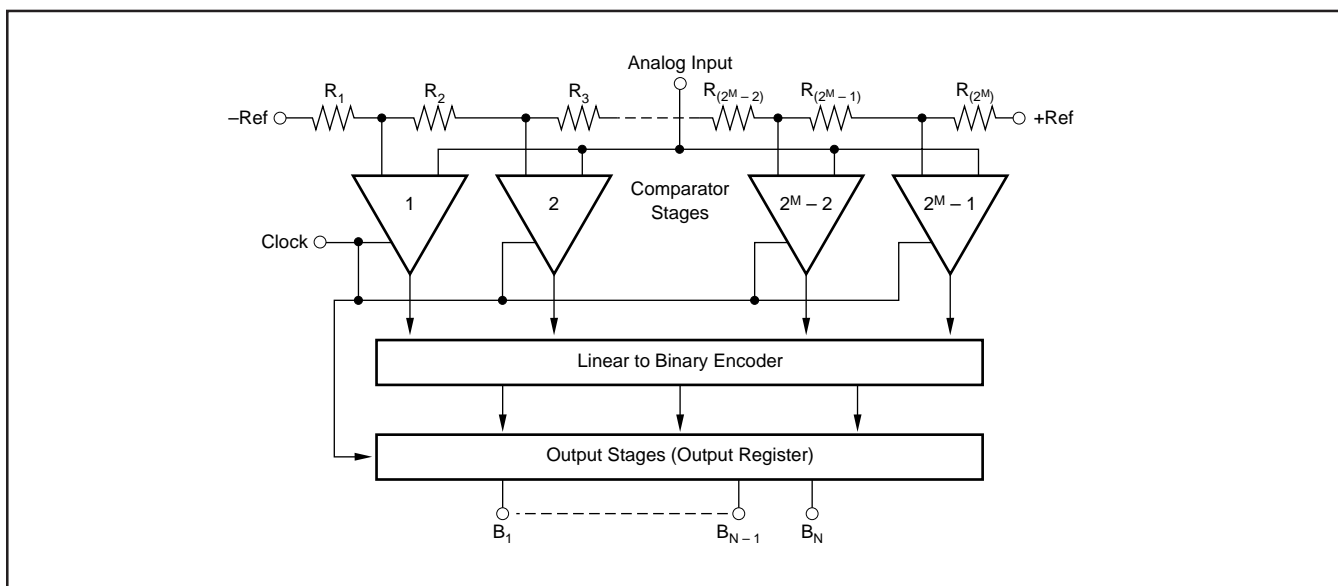


FIGURE 71. Block Diagram of a Flash Encoder.

ence biased 1LSB higher than that of the adjacent comparator. The reference for each one of the comparators is derived from a series connection of a string of resistors that is placed between the negative and positive reference. This resistor string is monotonic by design but it is possible for the entire flash converter not to be monotonic due to the comparator offset. This condition could possibly occur if the reference voltage is set too low thereby enabling the offset of the comparator to dominate the effective reference level which is the sum of each. Figure 72 illustrates this point. Suppliers of flash converters are able to produce monotonic results and still maintain LSB weight of 5mV, although LSB weights of 10mV are required for $\pm 1/2$ LSB linearity. The output of the comparators must be converted to a more economical digital code to be convenient to use. The thermometer code is typically converted to a conventional binary output. To achieve high sampling rates, digital pipelining is often employed in the design of the flash converter. This has the benefit of enabling a new sample to take place before the previous binary code has been formed.

The design of a comparator that is often employed in a flash converter is somewhat different compared to the design of a stand-alone comparator. The comparator input stage is configured to have a low gain state while in the tracking mode, and a high gain state while it is making the transition to the held state, the held state being the result of the comparison between the two inputs at the moment of sampling. Sampling takes place when a strobe pulse initiates positive feedback thereby causing regenerative action to take place which then sets the output of the comparator based upon the condition of the input. This method of design is necessary to achieve the simplicity required for a high resolution flash converter. Figure 22 shows a circuit diagram of a typical comparator stage of a flash ADC.

As previously mentioned, CMOS technology is employed in the design of low power flash converters with conversion

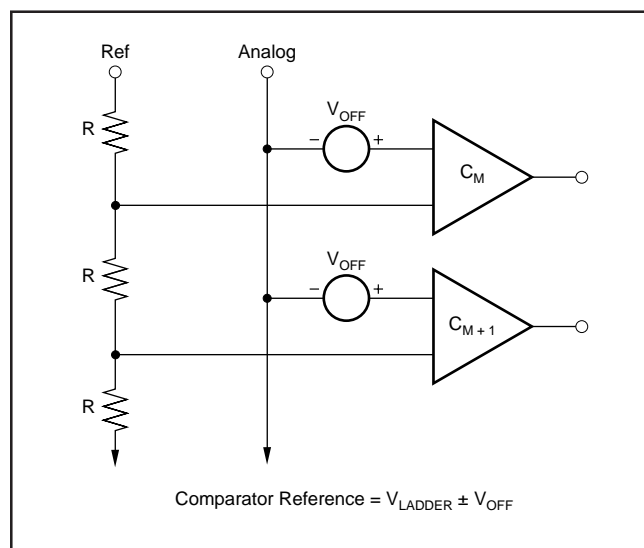


FIGURE 72. Comparator Offset.

rates of up to 20MHz. One of the drawbacks of CMOS comparators is that their offsets are much higher compared to bipolar comparators. Low offset comparators can be achieved in CMOS by use of the “auto-zero” technique. Figure 73 shows a diagram of an auto-zero comparator. Designing an auto-zero comparator is practical in CMOS because of the high impedance nature of CMOS. The auto-zero comparator operates by going into the auto-zero mode for part of the conversion time and staying in the measure mode during the remaining time. While in the auto-zero mode the input coupling capacitor charges up to the comparator offset so that when the comparator is placed back in the measure mode, the voltage on the capacitor is in such a direction to cancel the offset of the uncompensated comparator.

Dynamic performance is the one area that separates the performance of one flash converter from that of another. Dynamic performance is a measure of how a flash converter is able to accurately digitize a high frequency signal. This requires that the user understand how aperture jitter, aperture delay distortion and input bandwidth affect overall system performance. Input bandwidth is easily understood as this specification is similar to that of any band-limited device. The input bandwidth of a flash converter consists of both a small and large signal component that must be separately specified. Sometimes the large signal bandwidth is not directly specified but can be determined from the input slew rate. The input capacitance of flash converters can be high; therefore it is necessary to drive the encoder from a low impedance source to achieve high bandwidth.

Another phenomenon that limits the high frequency performance of a flash converter is aperture time. Aperture time is defined as the effective point where the comparator makes its decision. It should be noted that the aperture time is actually the difference between the delay in the path that is processing the compared signal and the delay in the path that

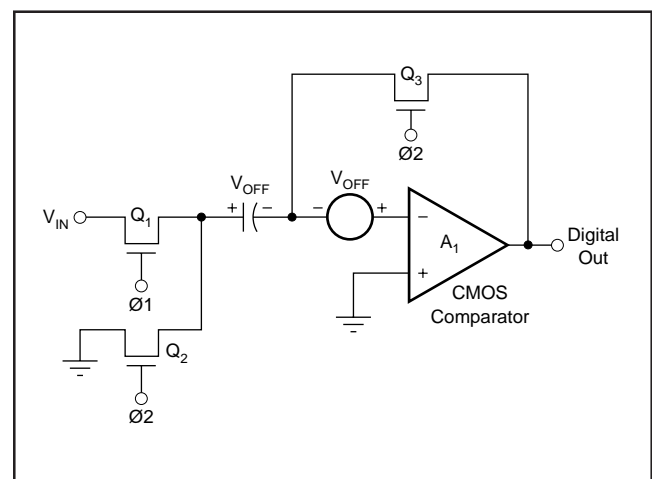


FIGURE 73. Auto Zero Comparator.

processes the strobe. This can become a serious source of distortion if the aperture delay of each comparator within the flash converter is different. As an example of this effect, consider how closely the aperture delay of an 8-bit, 200MHz flash converter needs to be matched to digitize a signal at the Nyquist rate while making only a 1LSB error. Aperture error is given by:

$$T_A = E_n / (D_{Fs} / D_T)$$

Where:

T_A = aperture time

E_n = allowable noise = 1LSB

F_s = signal frequency

D_{Fs} / D_T = max signal rate of change = $(2^N)(LSB)(7)(F_s)$

Substituting:

$$T_A = LSB / (2^N)(LSB)(\pi)(F_s) = 1 / (256)(\pi)(100E6) = 12.5ps$$

If the effective analog bandwidth of each comparator were 1GHz, the propagation delay of each comparator stage would be in the 100ps to 200ps range. It would then be necessary to match the delay of each comparator to 12.5ps to preserve the accuracy. Since flash converters can easily be 250mils in length and the signal could take as long as 400ps to propagate the length of the chip. The physical layout of the chip is extremely important to achieve acceptable high speed performance.

The high frequency performance of most analog to digital converters can be improved by conditioning the input signal by a sample and hold. This happens as the aperture distortion occurs due to the time delay of the individual comparators within the flash encoder not being matched. Since the sample and hold utilizes a single switch, the aperture performance of the combined system will be improved. One of the methods that can be used to determine the existence of aperture induced distortion is to measure the spectral response of the ADC by performing an FFT. This should first be performed at a low frequency to eliminate static accuracy as the source of the distortion. Aperture induced distortion will then be noted as the component of the distortion that increases with frequency. Even when a sample and hold is not required, interfacing an analog signal to a flash encoder deserves serious consideration. Both the input capacitance and resistance vary with signal level so it is important to drive these types of high speed converters with a low impedance source that can be supplied either from an op amp or from a buffer. This solution is not without its difficulties as high speed op amps are prone to oscillating when required to drive large capacitive loads characteristic of flash encoders. Usually, high speed op amps and buffers are capable of driving low resistive loads so it is possible to decouple the capacitive load from the driving source by placing a small resistor between the two. The resistor has the effect of making the impedance seen by the buffer look resistive thereby preventing an oscillatory condition. Setting the value of the resistor between 10Ω to 50Ω has minimal effect on the system bandwidth. Reference to the beginning of this study will show numerous buffers and amplifiers that could be suitable for interfacing to a flash

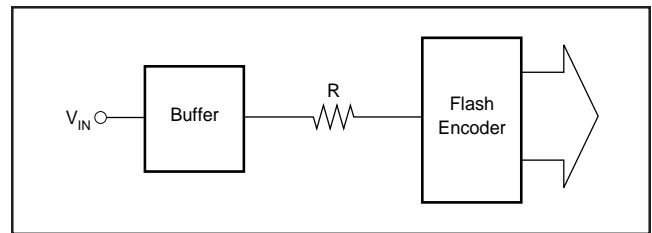


FIGURE 74. Buffer Driving Flash Encoder.

encoder. Refer to Figure 74 which shows how the coupling resistor helps stabilize the driving source.

It is straightforward to increase the resolution of a flash encoder by stacking two together as shown in Figure 75. It should be noted that two encoders stacked together in this manner will have poor aperture performance as matching the aperture delays of two separate encoders is difficult. This

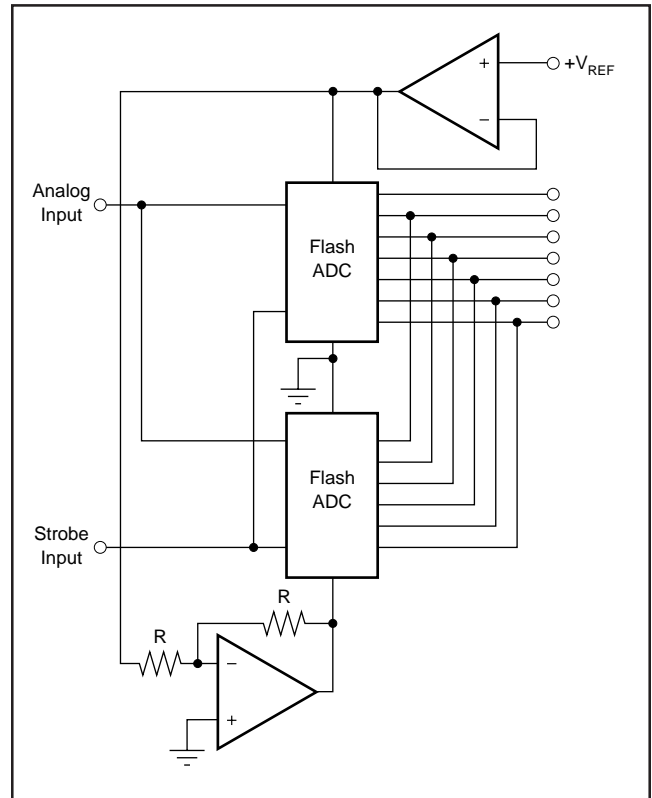


FIGURE 75. Stacked Flash Encoder.

can be connected by driving the stacked flash encoders by a sample and hold.

As previously mentioned, the resolution of flash converters is generally not greater than 8 bits. If the resolution of the converter were to increase by 1 bit, the amount of additional circuitry would have to double. Therefore, a 10-bit converter operating at the same speed as an 8-bit one would be four times as large and dissipate four times as much power. A conflict now develops when the designer attempts to use smaller geometry devices to reduce the size of the chip. With the use of smaller devices comes less accuracy which then compromises the possibility of achieving a 10-bit flash converter design. Similarly, the speed tends to be reduced to

avoid excessive power dissipation on the chip.

SUCCESSIVE APPROXIMATION ADC

One of the most popular architectures that is employed for the design of analog to digital converters is successive approximation. Successive approximation has achieved this high degree of popularity because this type of design has the highest degree of performance for the cost. Figure 76 shows a block diagram of a successive approximation ADC. As can be seen from the block diagram, the circuit design is straightforward, employing only a single comparator along with a digital to analog converter and the successive approximation logic. The previously mentioned comparators and digital to analog converter would serve as suitable subassemblies for a successive approximation ADC. Performance varies widely for designs employing this type of architecture, ranging

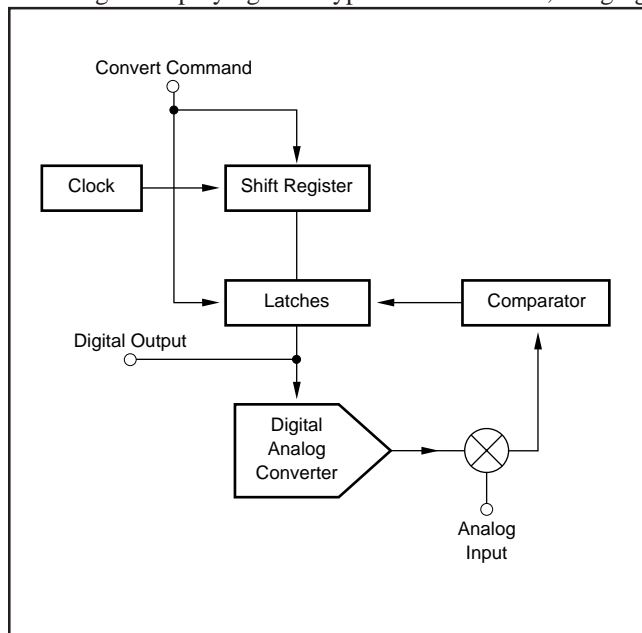


FIGURE 76. Block Diagram of Successive Approximation Analog-to-Digital Converter.

from 8 to 16 bits of resolution with conversion rates from 400ns to 25 μ s.

Hybrid, discrete, and monolithic technologies are all used in the manufacture of this type of converter, with hybrid designs dominating the high performance sectors. Recently monolithic designs have been introduced that were formerly the exclusive province of converters designed with hybrid technology. Performance levels with 12 bits of resolution at conversion times of 3 μ s can now be achieved in monolithic form with speeds down to 1 μ s on the “drawing board.” These single chip 12-bit ADCs employ bipolar, CMOS, and BiMOS for their design. Through BiMOS it is possible to use the most desirable features of CMOS and bipolar technologies. BiMOS processing offers both technologies on the same process. CMOS is optimum for achieving high speed logic with very little power dissipation, and bipolar technol-

ogy is better suited for low noise and high speed which is required by the analog section of the ADC.

Some of the most recent advances with these types of converters have been the introduction of CMOS technology for either monolithic converters or for supplying the logic function in hybrid designs. This is a very important development as CMOS offers lower power dissipation and current drain than bipolar logic. Both of these features are important to a system user as lower power dissipation leads to a lower temperature rise, greater reliability, and fewer problems with warm-up and temperature drift. Reduced current drain will enhance system accuracy as noise due to common analog and digital current paths will be reduced. At the present time successive approximation designs are dominated by the conventional R-2R ladder approach used by the digital to analog converter that lies within the ADC. Just over the horizon, several manufacturers have designs based upon charge distribution techniques employing CMOS. These newer CMOS designs also hold the potential of error correction and self-calibration that will enable converters to achieve greater stability with time than can be achieved with bipolar converters. This arises from the fact that higher circuit density can be achieved with CMOS which is required to implement the error correction function.

The successive approximation process begins with a start conversion pulse, setting the most significant bit to the “on” state with the remaining least significant bits in the “off” state. The output of the digital to analog converter is sent to one of the inputs of the comparator. The other input to the comparator is the analog signal that is to be digitized. After allowing an adequate amount of time for the digital to analog converter to settle, the output of the comparator is read into a latch where the decision is made whether to keep the bit on or not. If the input signal exceeds the weight of the MSB, the decision is made to keep the bit on. During the next trial period, Bit 2 is turned on and added to the result of the initial MSB comparison. In the event that the signal was greater than the MSB but not as great as the sum of the MSB and Bit 2, the MSB would be left on with Bit 2 being left off. This process of adding one more bit and testing the state of the comparator continues until all the bits of the digital to analog converter have been exercised. Figure 77 illustrates this process. Figure 78 shows a timing diagram of the successive approximation conversion cycle. One manufacturer adds digital correction to the conventional successive approximation algorithm. The first eight bits are converted only to 8-bit accuracy when the converter goes into a tracking mode to correct the conversion to 12-bit accuracy. This correction capability allows the first eight bits to operate at a higher sampling rate compared to conventional successive approximation.

Successive approximation has achieved wide popularity due to the simplicity of the design. The linearity of the ADC only depends on linearity of the digital to analog converter which is typically not true of the other ADC architectures being compared. The offset of the comparator creates an overall

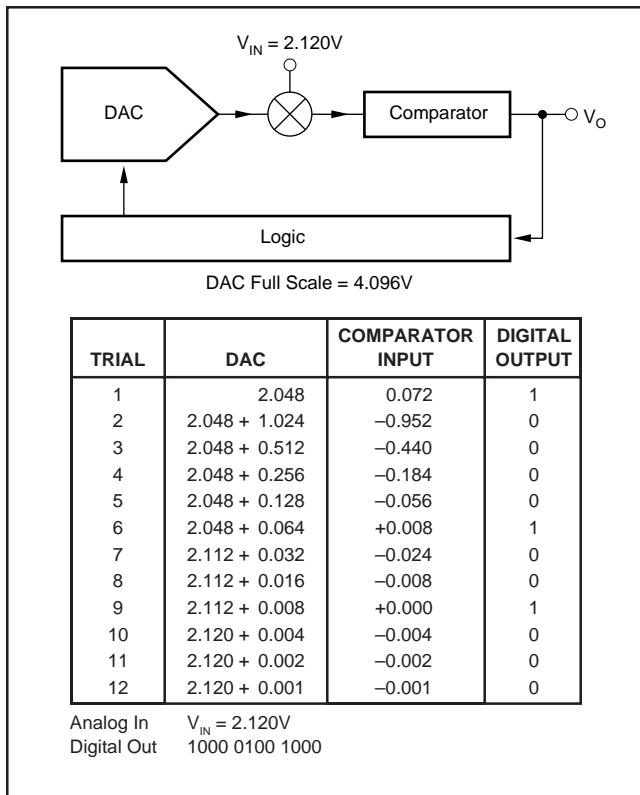


FIGURE 77. Successive Approximation Process.

offset but not a linearity error as would be created in a flash converter. The state of the art in producing accurate digital to analog converters is highly developed, which directly benefits successive approximation converters. ADCs employing these DACs will have correspondingly wide temperature ranges. Typically, successive approximation ADCs will operate over a wider temperature range, compared to other techniques, and designs are even available that operate at 200°C. Sub-ranging analog to digital converters have additional sources of error that distort the linearity, as will be explained later. Since there is only a single comparator, more power can be applied to the DAC and comparator to reduce the overall conversion time. Additionally, a successive approximation converter will naturally produce a serial form of the converted output. The serial output feature is very useful for economical digital transmission. Also, it lends itself for optical isolation techniques which helps reduce the interaction between the analog and digital sections of the processing system. With this design, only the DAC has to settle to the final accuracy. This is not true of the sub-ranging arrangement as will be explained later. Assuming that the conversion rate could be achieved, successive approximation would be the architecture of choice as a higher level of performance can be obtained at a given selling price for a particular conversion rate. Another attractive feature of successive approximation is that speed can be

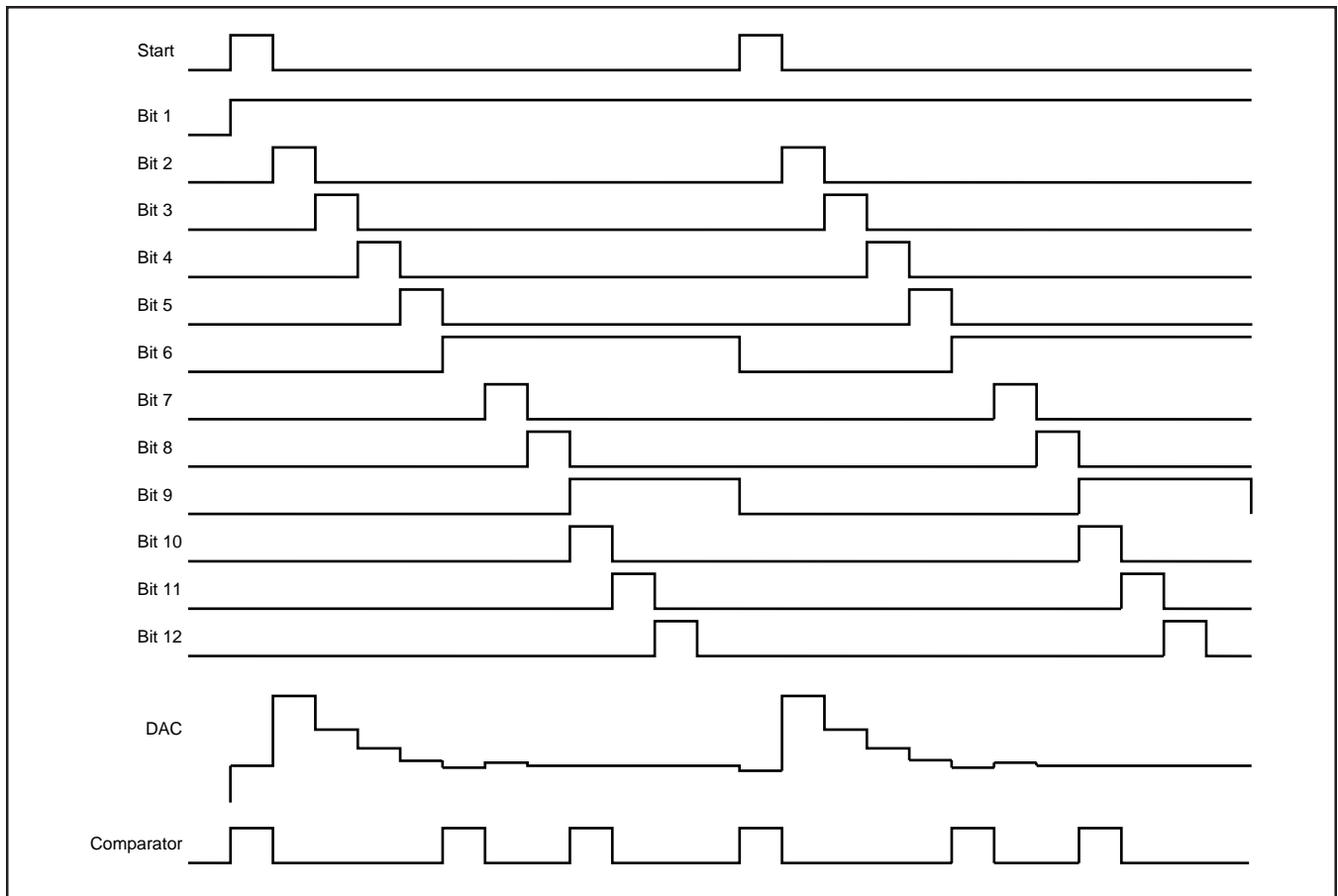


FIGURE 78. Timing Diagram of a Successive Approximation Converter.

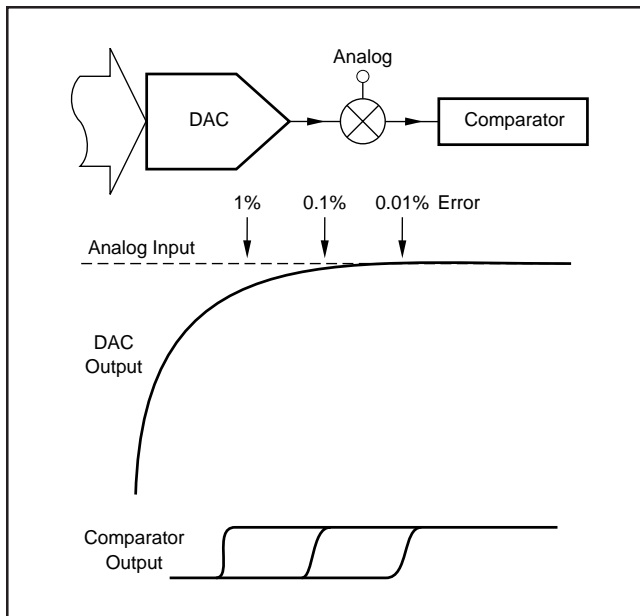


FIGURE 79. Speed vs Accuracy.

traded off against accuracy. Accuracy will degrade gradually due to the DAC not settling and the additional overdrive needed to switch the comparator more rapidly (see Figure 79). This tradeoff happens gradually, and very often a designer can increase the throughput rate of the system with only a moderate decrease in accuracy. This tradeoff is not possible with the other two architectures being compared, because once the stated conversion rate is exceeded, the accuracy degrades rapidly. Figure 80 shows this characteristic, which is common to successive approximation ADCs.

The aperture time of a successive approximation ADC is the conversion time. This occurs because it is necessary to hold the signal constant during the time when the conversion is taking place. If this does not occur, serious linearity errors

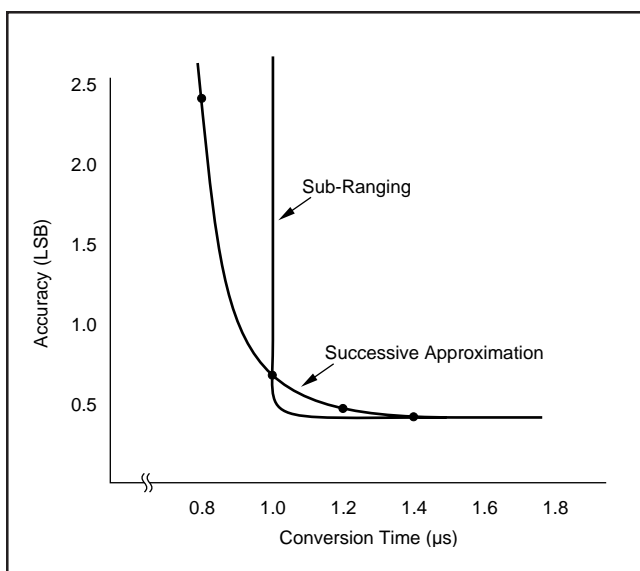


FIGURE 80. Speed vs Accuracy.

will develop. This results from the fact that the conversion is not essentially instantaneous as it is with the other converter types that are being compared. It then becomes necessary to condition the signal that is to be digitized by a sample and hold and not by the ADC. This is true whether the input signal is making rapid changes or not. The sample and hold must hold the signal constant during the ADC conversion time. Therefore, the only effect that is experienced by the ADC is varying signal levels on the comparator input which generally has a rapid recovery time.

A potential source of error that needs to be considered is when the comparator has been converting at one extreme of the input range and the signal changes to the opposite extreme. If the comparator is not properly designed, the analog to digital converter will experience a thermally induced offset that will cause multiple conversions not to be the same until thermal equilibrium is established. The system designer must also allow for sample to hold settling to take place as this could become a source of error—the input signal will not be the same during the conversion cycle time. Before the section on successive approximation converters is completed, an application problem needs to be mentioned. As the converter is going through the conversion cycle it is possible to inject a transient waveform into the source which is generated by the DAC. As the DAC value is being programmed by the logic the “summing junction” at the input to the comparator is not balanced and a signal will be injected into the sources. It is therefore important to drive high speed successive approximation ADCs with sources that have adequate settling performance or the rated linearity performance will not be achieved. The use of a common mode comparator arrangement can help to alleviate this problem by isolating the DAC from the input by the comparator’s common mode nature. See Figure 81.

SUB-RANGING ADC

The final architecture to be studied uses the sub-ranging or

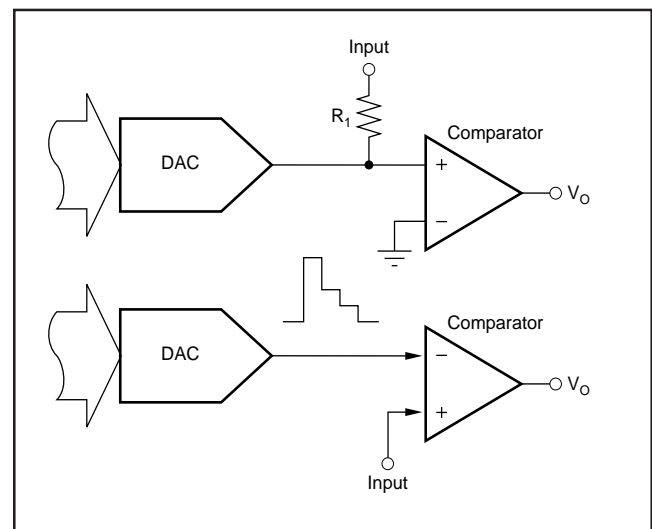


FIGURE 81. SAR ADC Input.

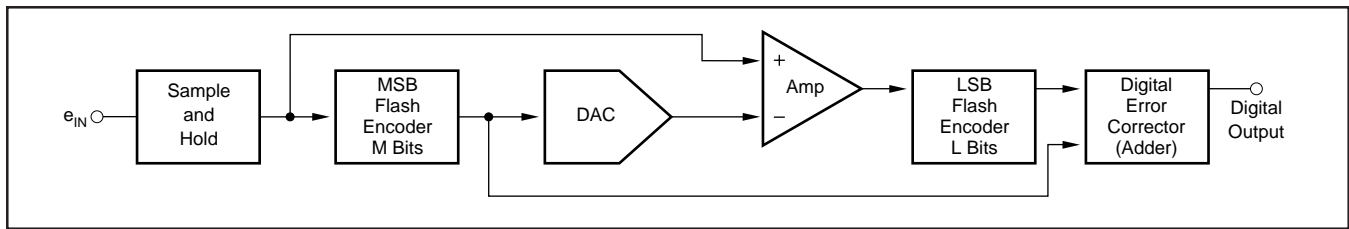


FIGURE 82. Block Diagram of Sub-Ranging ADC.

two-step technique. Both names are descriptive of these types of analog to digital converters. Sub-ranging converters are considered by designers when high resolution is required for conversion rates that are faster than can be achieved with successive approximation. As an example, a two-step design becomes the approach of choice when the system engineer requires 12 bits of resolution at conversion rates lower than $1\mu s$. This transition point between successive approximation and sub-ranging changes somewhat when lower resolution is required. Ten-bit performance can be achieved at conversion rates lower than $0.5\mu s$ using successive approximation.

Sub-ranging combines the elements of the two previously mentioned design techniques. All technologies are employed to produce sub-ranging designs varying from monolithic to modular. Flash converters only require one conversion cycle although 2^N comparators are necessary. Successive approximation uses only one comparator but N conversion cycles are needed. Sub-ranging is a mixture of the two, as an N -bit converter would use two cycles of an $N/2$ -bit flash converter. As an example, a 10-bit flash encoder would use 1023 comparators while a successive approximation type would use one comparator and a sub-ranging design would use 62 comparators. It should be noted that the sub-ranging ADC to be discussed uses only two ranges or conversion cycles. In general, more steps can be used and often find their way into higher resolution converters. The principle of operation is similar and for purposes of simplicity only the two-step version will be explained.

Refer to Figure 82 which shows a block diagram of a sub-

ranging converter. The analog signal is initially sent to a sample and hold to reduce aperture effects and to optimize AC performance. The output of the sample and hold then goes to an M -bit flash encoder and to a subtractor. After the sample and hold has acquired the signal and the sample to hold transient has decayed, the first encoder is strobed. The first encoder output determines the initial coarse approximation to the input signal. The digital output from the first encoder is sent to a digital to analog converter where it is converted back to analog form. This signal is then subtracted from the output of the sample and hold. The subtracted signal is then amplified before being applied to the second encoder which has L bits of resolution. The second encoder is also strobed with each encoder's output being sent to a digital adder where the final output word is created.

For a sub-ranging design to operate properly, it is necessary for $(M + L) > N$. These extra bits are used to encode internally developed errors which are capable of being corrected by a simple algorithm. Figure 83 helps to explain the operation of digital correction within a sub-ranging converter. The simplified analysis shown, in Figure 83, shows that the output of the ADC, with error correction, does not contain the error of the MSB encoder. This means that 12-bit accuracy can be achieved even though the MSB converter only has 8-bit accuracy. The output only contains the error of the LSB encoder which is reduced by the gain of the amplifier that precedes it. For simplicity the DAC error has been left off of the diagram shown in Figure 83 but has been shown in Figure 82. The DAC error has been

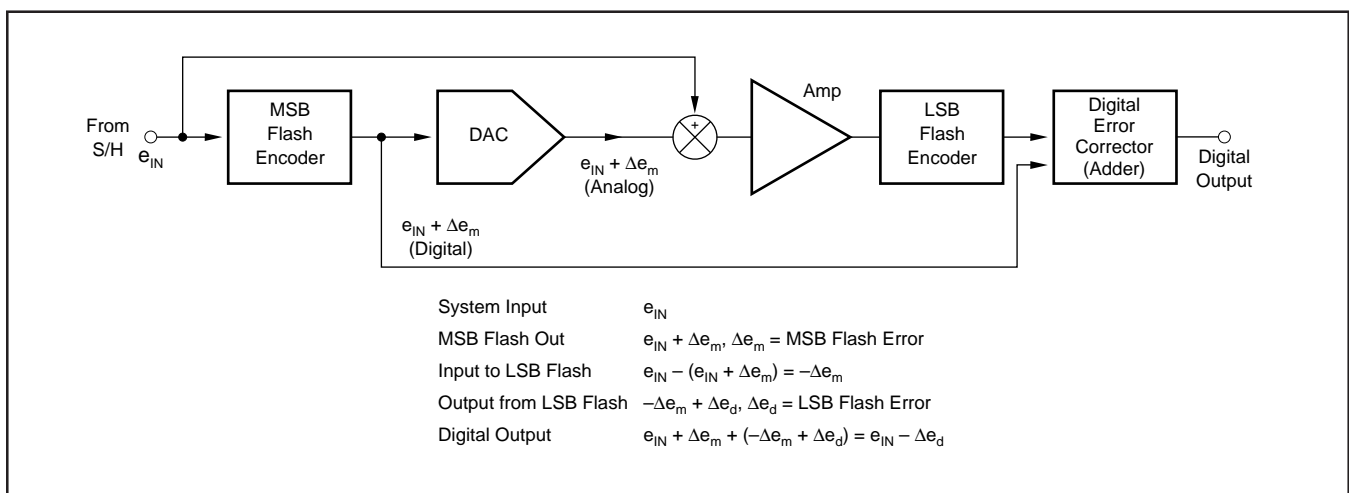


FIGURE 83. Error Correction.

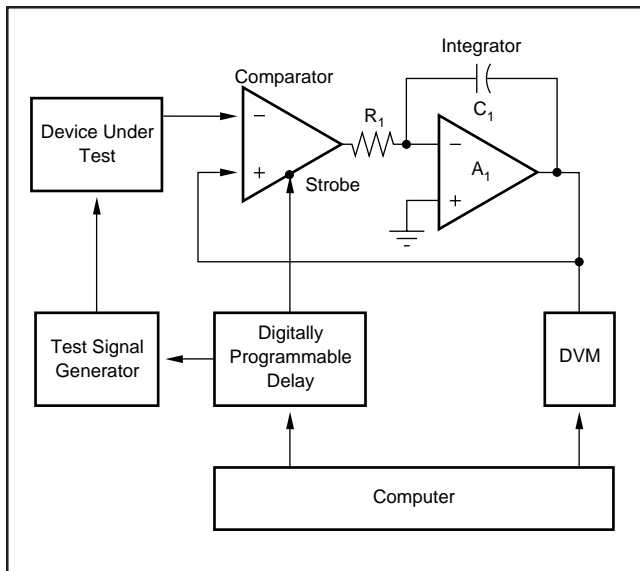


FIGURE 84. Waveform Digitizer.

omitted from the analysis because it has been assumed that the DAC is perfectly accurate. High speed DACs can achieve 14-bit accuracy, so this is a reasonable assumption. Another attractive feature of digital correction is that sample to hold settling errors can be corrected and will not lead to linearity errors as they do with successive approximation converters. The sample to hold settling error would be included as part of the MSB error. The M and L bit lines from each of the lower resolution encoders are then combined in the digital adder to form the final output word. To maintain high throughput rate the combining takes place during the next conversion cycle while the next data sample is being taken.

The adder, registers, and timing are grouped together and play the same role as does the successive approximation register of the single comparator design. Sub-ranging converters have achieved 12 bits of resolution with sampling rates to 20MHz.

TEST TECHNIQUES

There are numerous methods for evaluating the performance

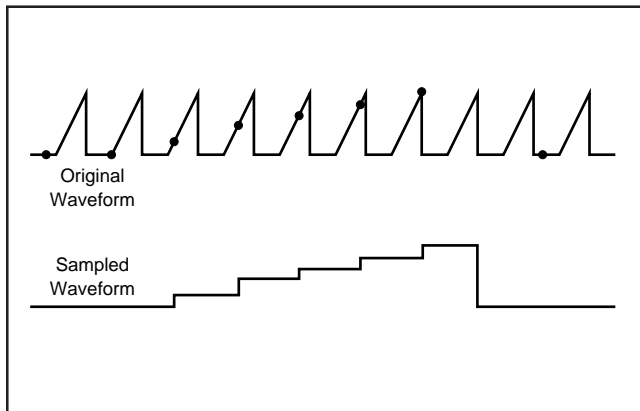


FIGURE 85. Waveform Sampling Process.

of the many components that comprise a data acquisition and conversion system. A few useful techniques will be discussed that may not be commonly mentioned but are useful to obtain experimental performance results. There may be other techniques that are not mentioned, although the technical literature is replete with these other techniques.

SETTLING TIME

The most powerful technique than can be employed to evaluate the settling time of a DAC or an amplifier is to digitize the waveform under test. Once the waveform is digitized, the waveform can be sent to a computer where software routines can be used to determine the performance of the device under test. Digitizing the waveform is superior to hardware-oriented instrumentation because of the versatility associated with a computer. Once the waveform has been digitized, any property of the waveform can be analyzed with the same hardware. Figure 84 is a block diagram of the digitizer. The waveform under test is fed to the inverting comparator input. The comparator's digital output is integrated by the op amp and fed back to the input. Figure 85 is an illustration of the sampling or digitization process. The sampled waveform shown in Figure 85 is a crude representation of the original signal and was done this way for purposes of the illustration. In actual practice, sampling is performed in fine increments to achieve high accuracy. Sampling of the waveform under test is accomplished by repeatedly strobing the comparator at a selected time point, until the integrator feedback forces the comparator reference input to equal the sampled value of the input signal. Once the loop settles, this value is read by the DVM and sent to the computer. The sample is then incremented by the computer through the programmable delay.

APERTURE JITTER

Determining the aperture jitter of an ADC or sample and hold can be accomplished by the block diagram shown in Figure 86. This system avoids introducing any additional error due to instrumentation induced jitter as the added delay

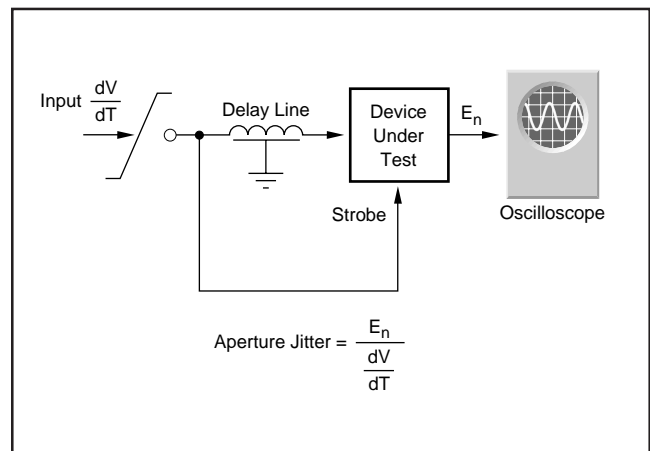


FIGURE 86. Aperture Jitter Measurement.

is a passive delay line. When the sampling signal is the signal being sampled there is no possibility of instrument induced noise affecting the measurement. Once the delay is adjusted so the maximum rate of change section of the signal is being sampled, the aperture jitter T_A is determined by the relationship:

$$T_A = E_n / (dV/dt)$$

Where:

E_n = measured noise

(dV/dt) = input rate of change

BEAT FREQUENCY TESTING

Beat frequency tests are qualitative tests that provide a quick, simple visual demonstration of dynamic ADC performance. Figure 87 shows a block that is used to perform a beat frequency evaluation. An input frequency is selected that provides the worst case change. This usually occurs at the Nyquist rate. The name “beat frequency” describes the nature of the test. The sample frequency is chosen to be a multiple of the input frequency plus a small incremental frequency (see Figure 88). By choosing a low beat frequency, the dynamic performance of the DAC does not affect the accuracy of the measurement. With the block diagram shown, the output of the ADC is resampled at 1/2 the data rate to enable evaluation at the Nyquist rate. The beat frequency is set so that many samples are taken at each code. The beat frequency test should not be used as a substitute for more accurate methods for determining high frequency performance such as FFT measurements or histogram testing, but it provides a very effective method for optimizing the dynamic performance during the development stage of a project. The design engineer will get instant visual feedback, via the oscilloscope, to help pinpoint a circuit defect. This type of cause and effect relationship is not as easy to establish using more complex computer-oriented tests. Burr-Brown, of Tucson, Arizona, offers Application Note AN-133, which describes many other dynamic tests for evaluating ADC performance.

SERVO LOOP TEST

Figure 89 shows the block diagram of a system that can be used to evaluate the DC integral and differential linearity along with the gain and offset of an analog to digital converter. The desired code that is to be measured is loaded into the digital comparator from the computer. Based upon the results of the comparison between the output of the ADC and the desired code, the comparator will command the integrator to slew until a balance is reached. Loop balance will be established when the output of the integrator produces a voltage that is equal to the code transition voltage.

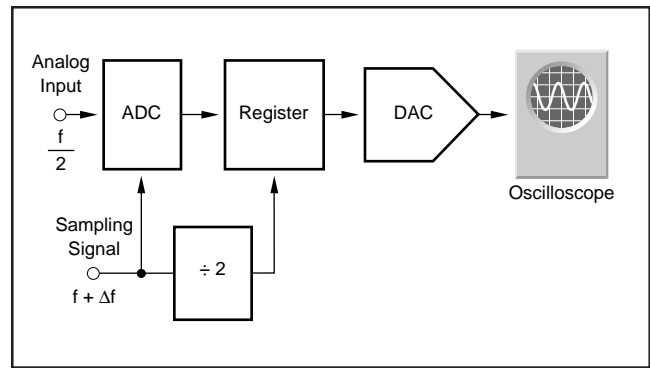


FIGURE 87. Block Diagram of Beat Frequency Testor.

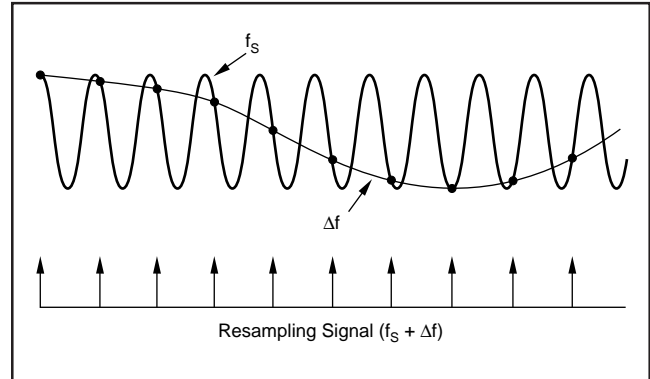


FIGURE 88. Beat Frequency Waveforms.

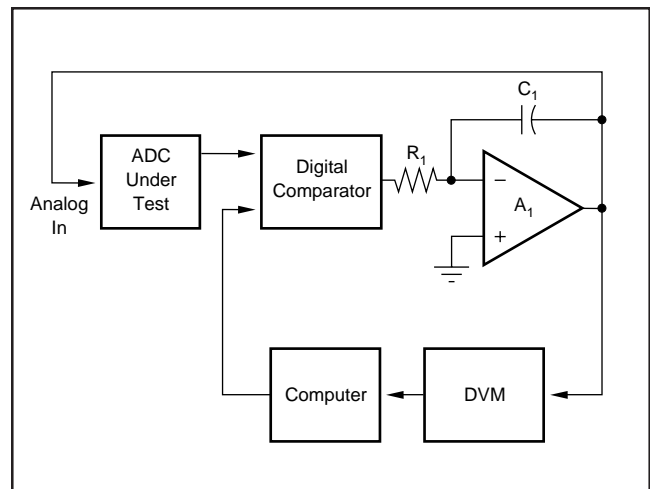


FIGURE 89. Block Diagram of Servo Loop Test.

The DVM reads this voltage and in this manner a measure of all the ADC's code transition points are established. Software can then be written to determine the ADC performance. The accuracy of this technique is dependent upon the DVM which can approach seven digits.