CHAPTER 2
Signal Conditioning

2.1 Sensors

Sensors are elements for monitoring the performance of machines and processes. The common classification of sensors is: distance, movement, proximity, stress/ strain/force, and temperature. There are many commercially available sensors but we have picked the ones that are frequently used in mechatronic applications. Often, the conditioned signal output from a sensor is transformed into a digital form for display on a computer or other display units. The apparatus for manipulating the sensor output into a digital form for display is referred to as a

measuring instrument (see Figure 2.1 for a typical computer-based measuring system).
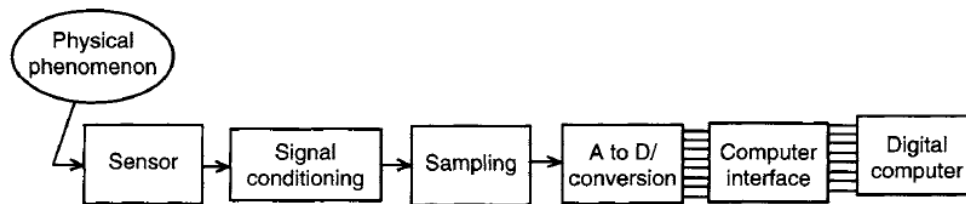

Figure 2.1 Measurement system.

2.1.2 Actuators

While a sensor is a device that can convert mechanical energy to electrical energy, an electrical actuator, on the other hand, is a device that can convert electrical energy to mechanical energy. All actuators are transducers (as they convert one form of energy into another form). Some sensors are transducers (e.g. mechanical actuators), but not all. Actuators are used to produce motion or action, such as linear motion or angular motions. Some of the important electrical actuators used in mechatronic systems include solenoids, relays, electric motors (stepper, permanent magnet, etc.). These actuators are instrumental in moving physical objects in mechatronic systems.

Mechanical actuators are transducers that convert mechanical energy into electrical energy. Some of the important mechanical actuators used in mechatronic systems include hydraulic cylinders and pneumatic cylinders.

2.2 Analog electronics
2.2.1 Introduction

Usually electrical signals in mechatronic and measuring systems come from transducers, which convert physical quantities (displacement, temperature, strain, flow, etc.) into voltages; the output of which is usually in analog signal form since it is continuous and varies with time. Often the signals from transducers need to be cleaned up because they could be distorted (noisy, too small, corrupted, or d.c. offset).

The primary purpose for the analog signal conditioning circuitry is to modify the transducer or sensor output into a form that can be optimally converted to a discrete time digital data stream by the data acquisition system. Some important input requirements of most data acquisition systems are:

• The input signal must be a voltage waveform. The process of converting the sensor output to a voltage can also be used to reduce unwanted signals, that is noise.

•The dynamic range of the input signal should be at or near the dynamic range of the data acquisition system (usually equal to the voltage reference level, Kref, or 2Fref). This is important in maximizing the resolution of the analog to digital converter (ADC).

• The source impedance, Zs , of the input signal should be low enough so that changes in the input impedance, Zin, of the data acquisition system do not affect the input signal.

• The bandwidth of the input signal must be limited to less than half of the sampling rate of the analog to digital conversion. Analog signal processing comprises of the following issues: signal isolation, signal preprocessing, and removal of undesirable signals. These are now briefly discussed.

2.2.2 Signal isolation

In many data acquisition applications it is necessary to isolate the sensor from the power supply of the computer. This is done in one of two ways: magnetic isolation or optical isolation. Magnetic isolation by means of a transformer is primarily used for coupling power from the computer or the wall outlet to the sensor. Optical isolation is used for coupling the sensor signal to the data acquisition input. This is usually done through the use of a light emitting diode and a photodetector often integrated into a single IC package.

2.2.3 Signal preprocessing

Many times it is desirable to perform preprocessing on the sensor signal before data acquisition. Depending on the application, this can help lower the required computer processing time, lower the necessary system sampling rate, or even perform functions that will enable the use of a much simpler data acquisition system entirely. For example, while an accelerometer system can output a voltage proportional to acceleration, it may be desired to only tell the computer when the acceleration is greater than a certain amount. This can be accomplished in the analog signal conditioning circuitry. Thus, the data acquisition system is reduced to only having a single binary input (and thus no need for an ADC).

2.2.4 Removal of undesired signals

Many sensors output signals that have many different components to them or other signals may corrupt the signal. It may be desirable or even necessary to remove such components before the signal is digitized. This *noise* can also be removed using analog circuitry. For example, 60 Hz interference can distort the output of low output sensors. The signal conditioning circuitry can remove this before it is amplified and digitized.

The simplest and the most common form of signal processing is amplification where the magnitude of the voltage signal is increased. Other forms of signal processing include inversion, addition, subtraction, comparison, differentiation, and integration. The operational amplifier is an integrated circuit that can perform these operations. Accordingly, in the remaining part of this section, simple circuit models of the operational amplifier (op amp) will be introduced. The simplicity of the models will permit the use of the op amp as a circuit element, or building block, without the need to describe its internal workings in detail. For the purpose of many instrumentation applications, the op amp can be treated as an ideal device.

2.2.5 Amplifiers

The op amp was designed to perform mathematical operations and is a common feature of modern analog electronics. The op amp is constructed from several transistor stages, which usually include a differential input stage, an intermediate gain stage and a push-pull output stage. The differential amplifier consists of a matched pair of bipolar transistors or FETs. The push-pull amplifier transmits a large current to the load and hence has a small output impedance. An op amp can perform a great number of operations such as addition, filtering, or integration, which are based on the properties of ideal amplifiers and of ideal circuit elements.

The op amp is a linear amplifier with *Vout* oc *V-m.* The d.c. open-loop voltage gain of a typical op amp is 102 to 106. The gain is so large that most often feedback is used to obtain a specific transfer function and control the stability. Inexpensive IC versions of op amps are readily available, making their use popular in any analog circuit. These operate from d.c. to about 20 kHz, while the highperformance models operate up to 50 MHz. A popular device is the 741 op-amp which drops off 6dB/octave above 5 Hz. Op amps are usually available as an IC in an 8-pin dual, in-line package (DIP). Some op amp ICs have more than one op amp on the same chip. Figure 2.1 shows the internal design of a 741 op amp IC.

Many sensors output a voltage waveform so no signal conditioning circuitry is needed to perform the conversion to a voltage. However, dynamic range modification, impedance transformation, and bandwidth reduction may all be necessary in the signal conditioning system depending on the amplitude and bandwidth of the signal and the impedance of the sensor. It is especially important to review the analysis of ideal op amp circuits.
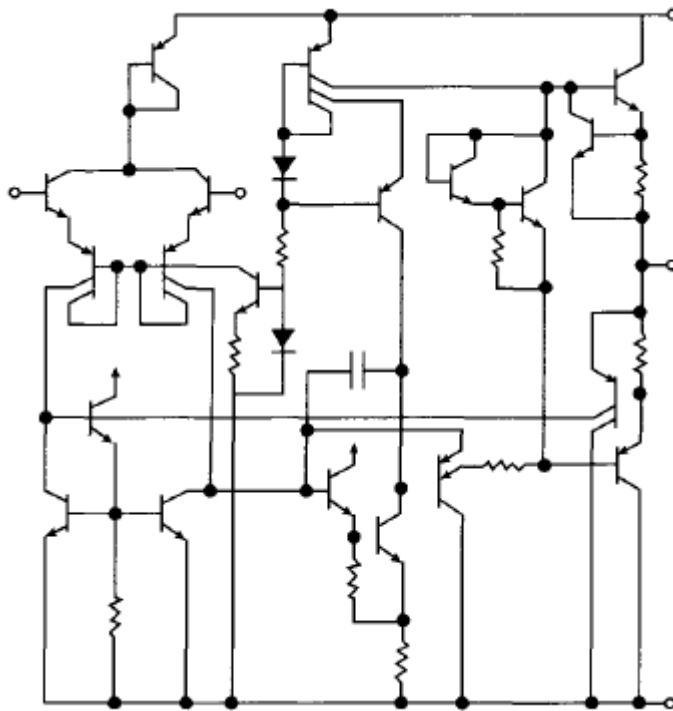

Figure 2.1 A 741 op amp internal design. (Courtesy of National Semiconductor Inc.)

2.3 The ideal operational amplifier model

The ideal op amp model is shown in Figure 2.2. Before continuing we define some terminology:

• **linear amplifier:** the output is directly proportional to the amplitude of input signal.
• **open-loop gain, *A*:** the voltage gain without feedback (^ 106) as shown in Figure 2.2(a).
• **closed-loop gain, *G*:** the voltage gain with negative feedback (approximation to *H(}co))* as shown in Figure 2.2(b).
• **negative feedback:** the output is connected to the inverting input forming a feedback loop (usually through a *feedback resistor RF)* as shown in Figure 2.2(b). Figure 2.2(c) shows an ideal model for analyzing circuits containing op amps. The ideal op am model is based on the following assumptions: it has infinite impedance at both inputs, consequently there is no current drawn from the input circuits;
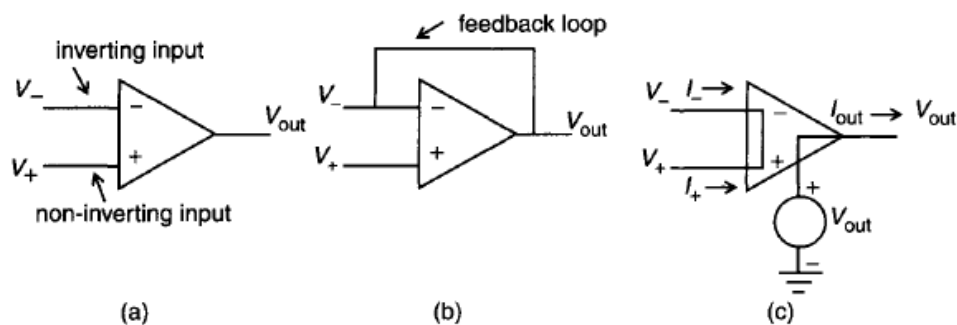

Figure 2.2 The op amp model: (a) open loop; (b) closed loop; (c) ideal.

• it has infinite gain, hence the difference between the input and output voltages is zero. This is denoted by short circuiting the two inputs;
• it has zero output impedance, so that the output voltage is independent of the output current.

All the signals are referenced to ground and feedback exists between the output and the inverting input in order to achieve stable linear behavior.

2.4 The inverting amplifier

The most common circuit used for signal conditioning is the inverting amplifier circuit as shown in Figure 2.3. This amplifier was first used when op amps only had one input, the inverting (—) input. The analysis is done by noting that at the inverting input node, Kirchoff 's current law requires that

$$i_I + i_{out} = 0$$

$$\frac{V_i}{R_I} + \frac{V_{out}}{R_F} = 0.$$

Leading to the voltage gain

$$\frac{V_{out}}{V_i} = -\frac{R_F}{R_I}.$$

Thus the level of sensor outputs can be matched to the level necessary for the data acquisition system. The input impedance is approximately Ri and
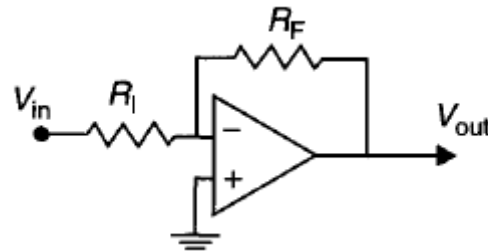

Figure 2.3 The inverting amplifier.

the output impedance is nearly zero, so this circuit provides impedance transformation between the sensor and the data acquisition system.

2.5 The non-inverting amplifier

To avoid the negative gain (i.e. phase inversion) introduced by the inverting amplifier, a non-inverting amplifier configuration is commonly used as shown in Figure 2.4. The analysis of the non-inverting amplifier is done in the same way as the inverting amplifier by noting that at the inverting input node, Kirchoff s
current law requires that

$$i_f + i_n - i_i = 0; \qquad (i_n \approx 0)$$

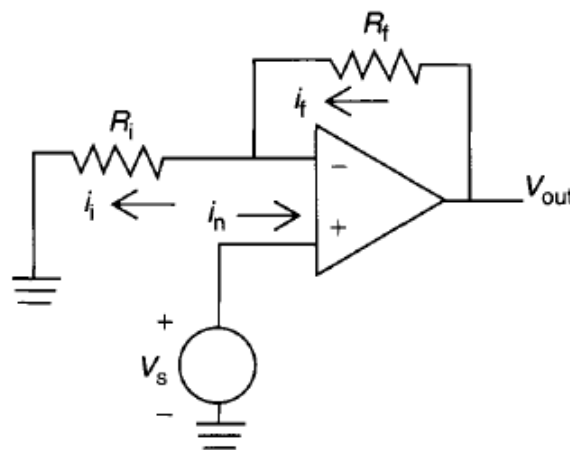$$\frac{V_{out} - V_s}{R_f} - \frac{V_s - 0}{R_i} = 0$$
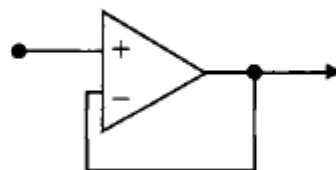

Figure 2.4 The non-inverting amplifier.

Figure 2.5 The unity-gain buffer

Leading to the voltage gain

$$\frac{V_{out}}{V_s} = 1 + \frac{R_f}{R_i}$$

The input impedance is nearly infinite (limited only by the op amp's input impedance) and the output impedance is nearly zero. The circuit is ideal for sensors that have a high source impedance and thus would be affected by the current draw of the data acquisition system.

The unity-gain buffer

If $R_f=0$ and $R_i=0$ and is open (removed), then the gain of the non-inverting amplifier is unity. This circuit, as shown in Figure 2.5 is commonly referred to as a unity-gain buffer or simply a buffer

The summing amplifier

The op amp can be used to add two or more signals together as shown in Figure 2.6. The analysis of the summing amplifier is done by noting that at the inverting input node, Kirchoff s current law requires that

$$i_1 + i_2 + ... + i_i + i_f - i_n = 0; \qquad (i_n \approx 0)$$

$$\frac{V_1}{R_1} + \frac{V_2}{R_2} + ... + \frac{V_i}{R_i} + \frac{V_{out} - 0}{R_f} = 0.$$

Leading to

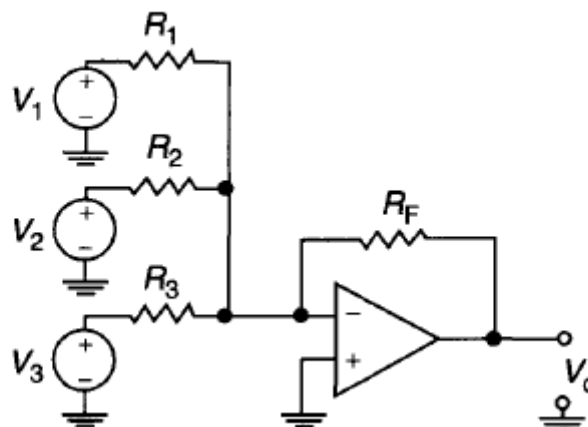$$V_{out} = -R_f \sum_{i=1}^{N} \frac{V_i}{R_i}$$

Figure 2.6 The summing amplifier.

$$V_{\text{out}} = -\sum_{i=1}^{N} V_i \quad \text{if} \quad R_f = R_i.$$

This circuit can be used to combine the outputs of many sensors such as a microphone array.

2.8 The difference amplifier

The op amp can also be used to subtract two signals as shown in Figure 2.7. The priniciple of superposition is used to analyze the difference amplifier. In the first step, we short $V2$ so that we have an inverter, and that the output to Vi is
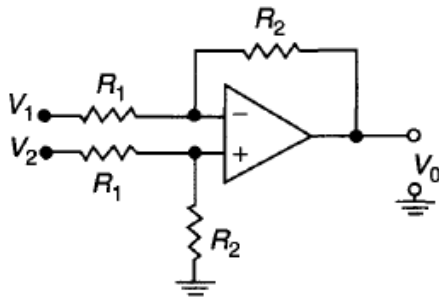
$$V_{o1} = -\frac{R_2}{R_1} V_1$$

Figure 2.7 The difference amplifier.

In the second step, we short we short $V\backslash$ so that we have a non-inverter. We then use the voltage divider principle to obtain the non-inverting voltage as

$$V_3 = \frac{R_2}{R_1 + R_2} V_2$$

$$V_{o2} = \left(1 + \frac{R_2}{R_1}\right) V_3 = \left(1 + \frac{R_2}{R_1}\right)\left(\frac{R_2}{R_1 + R_2}\right) V_2$$

The priniciple of superposition means that the output voltage is the sum of the input voltages:

$$V_\mathrm{o} = V_\mathrm{o1} + V_\mathrm{o2} = -\left(\frac{R_2}{R_1}\right)V_1 + \left(1 + \frac{R_2}{R_1}\right)\left(\frac{R_2}{R_1 + R_2}\right)V_2.$$

or

$$V_\mathrm{o} = V_\mathrm{o1} + V_\mathrm{o2} = \left(\frac{R_2}{R_1}\right)(V_2 - V_1).$$

Thus the difference amplifier magnifies the difference between the two input signals by the closed-loop gain *(R2/R\)*. This circuit is commonly used to remove unwanted d.c. offset. It can also be used to remove differences in the ground potential of the sensor and the ground potential of the data acquisition circuitry (so-called ground loops). In this case *V2* can be the output of the sensor and *V* can be the signal that is to be removed.

2.9 The instrumentation amplifier

When the input signals are very low level and also have noise, the difference amplifier is not able to extract a satisfactory difference signal. Possibly the most important circuit configuration for amplifying sensor output when the input signals are very low level is the instrumentation amplifier (IA). The requirements for an instrumentation amplifier are as follows:
• Finite, accurate and stable gain, usually between 1 and 1000.
• Extremely high input impedance.
• Extremely low output impedance.
• Extremely high common mode rejection ratio (CMRR).

The CMRR is defined as:

$$\mathrm{CMRR} = \left(\frac{A_\mathrm{vd}}{A_\mathrm{vc}}\right)$$

where

$$A_\mathrm{vd} = \frac{V_\mathrm{out}}{V^+ - V^-} = \text{differential-mode gain}$$

$$A_\mathrm{vc} = \frac{V_\mathrm{out}}{\frac{V^+ + V^-}{2}} = \text{common-mode gain}$$

That is, CMRR is the ratio of the gain of the amplifier for differential-mode signals (signals that are different between the two inputs) to the gain of the amplifier for common-mode signals (signals that are the same at both inputs). The difference amplifier described above, clearly does not satisfy the second requirement of high input impedance. To solve this problem, a non-inverting amplifier is placed at each one of the inputs to the difference amplifier as shown in Figure 2.8. Remember that a non-inverting amplifier has a nearly infinite input impedance. Notice that instead of grounding the resistors, the two resistors are connected together to create one common resistor, *RG.* The analysis of the instrumentation amplifier is simplified by taking advantage of symmetry: we recognize two symmetric halves at the input (draw a vertical line between *R* and *R3,* and a horizontal line across *RG).* This results in an

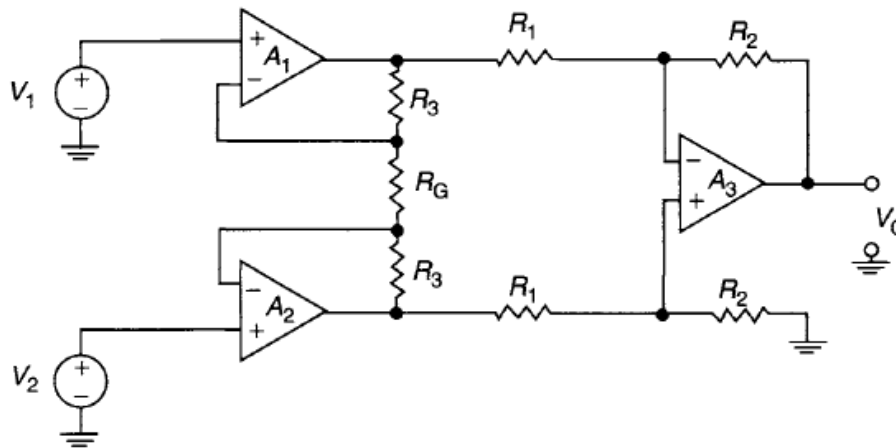equivalent circuit shown in Figure 2.9.

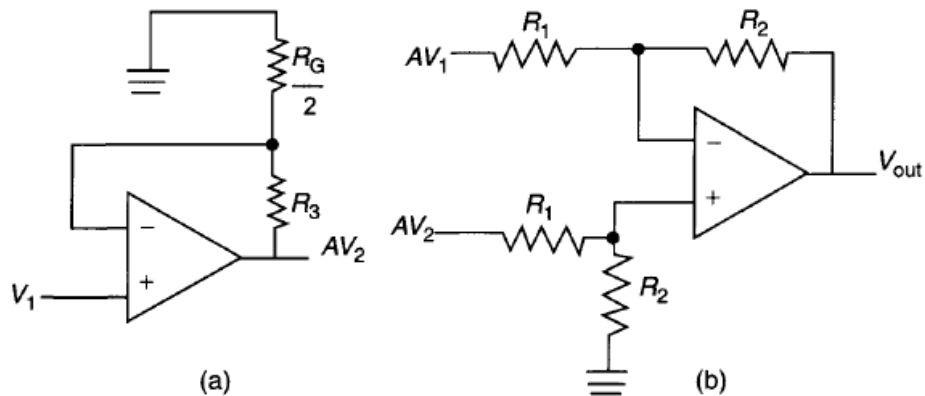**Figure 2.8** The instrumentation amplifier.

Figure 2.9 Equivalent instrumentation amplifier for analysis: (a) step 1; (b) step 2.

The analysis is done in two steps:

■ Step 1: Using Figure 5.9(a), which is a non-inverting amplifier,

$$A = 1 + \frac{R_3}{R_G/2} = 1 + \frac{2R_3}{R_G}.$$

■ Step 2: Using Figure 5.9(b), which is a difference amplifier,

$$V_{out} = \left(\frac{R_2}{R_1}\right)(AV_2 - AV_1) = \left(\frac{R_2}{R_1}\right)\left(1 + \frac{2R_3}{R_G}\right)(V_2 - V_1).$$

■ The overall differential gain of the circuit is

$$A_{vd} = \left(\frac{R_2}{R_1}\right)\left(1 + \frac{2R_3}{R_G}\right).$$

2. 10 The integrator amplifier

The op amp in Figure 2.10 provides an output that is proportional to the integral of vin(r), an arbitrary function of time (e.g. a pulse train, a triangular wave, or a square wave).
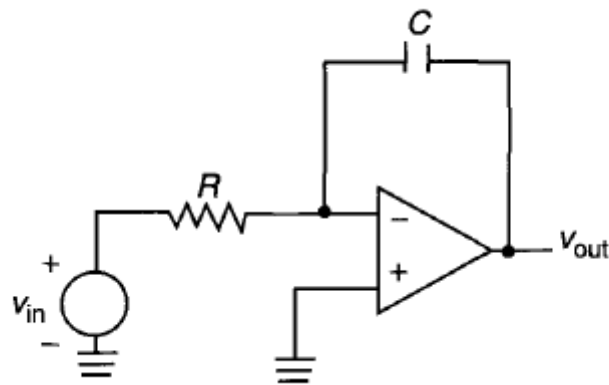
Figure 2.10 The integrating amplifier.

The analysis of the differential circuit is based on the observation that

$$i_{in}(t) = -i_f(t)$$

$$i_{in}(t) = \frac{v_{in}(t)}{R}$$

and

$$i_f(t) = C\frac{dv_{out}(t)}{dt}$$

$$\frac{1}{RC}v_{in}(t) = -\frac{dv_{out}(t)}{dt}$$

$$v_{out}(t) = -\frac{1}{RC}\int_{-\infty}^{t} v_s(t)\,dt.$$

2.11 The differentiator amplifier

The op amp in Figure 2.11 provides an output that is proportional to the integral of vin(f), an arbitrary function of time (e.g. a pulse train, a triangular wave, or a square wave).
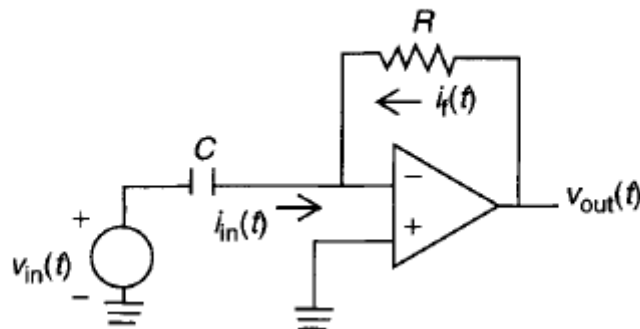

Figure 2.11 differential amplifier

The analysis of the differential circuit is based on the observation that

$$i_{in}(t) = -i_f(t)$$

$$i_{in}(t) = C\frac{dv_{in}(t)}{dt}$$

and

$$i_f(t) = \frac{v_{out}(t)}{R}$$

$$v_{out}(t) = -RC\frac{dv_{in}(t)}{dt}.$$

2.12 The comparator

Sometimes there is no need to send the entire range of voltages from a sensor to an analog-to-digital converter (ADC). Instead, often a sensor is used simply as a switch. Figure 2.12 functions as a *comparator* which takes an analog sensor
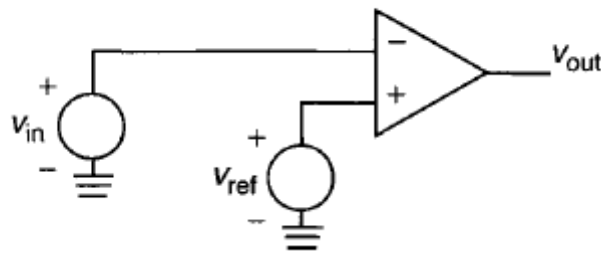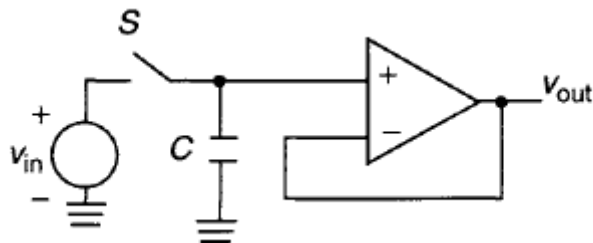

Figure 2.12 The comparator.

Figure 2.13 The sample and hold amplifier.

2.13 The sample and hold amplifier

The purpose of sample and hold circuitry is to take a snapshot of the sensor signal and hold the value. An ADC must have a stable signal in order to accurately perform a conversion. An equivalent circuit for the sample and hold is shown in Figure 2.13. The switch connects the capacitor to the signal conditioning circuit once every sample period. The capacitor then holds the measured voltage until a new sample is acquired. Often, the sample and hold circuit is incorporated in the same integrated circuit package as the amplifier.

2.13.1 Problems with sample and hold amplifiers

• **Finite aperture time:** The sample and hold takes a period of time to capture a sample of the sensor signal. This is called the aperture time. Since the signal will vary during this time, so the sampled signal can be slightly inaccurate. • **Signal feedthrough:** When the sample and hold is not connected to the signal, the value

being held should remain constant. Unfortunately, some signal does bleed through the switch to the capacitor, causing the voltage being held to change slightly. • **Signal droop:** The voltage being held on the capacitor starts to slowly decrease over time if the signal is not sampled often enough.
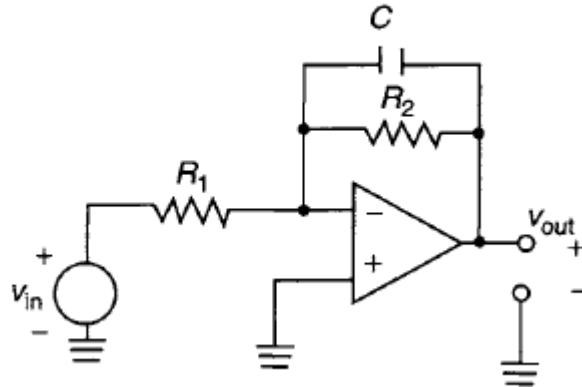

Figure 2.14 Low-pass active filter.

4 Active filters

There are many practical applications that involve filters of one kind or another; for example, filters to eliminate impurities from drinking water, sunglasses to reduce the light intensity reaching the eye, etc. Similarly, in electric circuits, it is possible to attenuate or reduce/eliminate the amplitude of unwanted frequencies caused by electric noise or other forms of interference. This section treats the analysis of electric filters.

1 The low-pass active filter

The inverting amplifier configuration can be modified to limit the bandwidth of the incoming signal. For example, the feedback resistor can be replaced with a resistor/capacitor combination as shown in Figure 2.14. We now analyze the filter as follows:

$$Z_F = Z_R \parallel Z_C = R_2 \parallel \frac{1}{j\omega C}$$

$$Z_F = \frac{R_2 \frac{1}{j\omega C}}{R_2 + \frac{1}{j\omega C}} = \frac{R_2}{1 + j\omega C R_2}$$

$$Z_i = Z_R = R_1.$$

The gain of this filter is given by:

$$A_{LP}(j\omega) = -\frac{Z_F}{Z_i} = -\frac{R_2/R_1}{1+j\omega CR_2} = \left(\frac{-R_2}{R_1}\right)\frac{1}{1+j\frac{\omega}{\omega_0}}$$

$$A_{LP}(j\omega) = H_0\frac{1}{1+j\frac{\omega}{\omega_0}},$$

where

$$H_0 = \left(\frac{-R_2}{R_1}\right); \omega_0 = \frac{1}{CR_2}; f_0 = \frac{1}{2\pi R_2 C}$$

As $\omega \to 0$, $A_{LP}(j\omega) \to H_O$

As $\omega \to \infty$, $A_{LP}(j\omega) \to 0$.

When $\omega = \omega_0$, $A_{LP}(j\omega) = H_0\frac{1}{1+j}$.

Under the last condition, the filter *rolls off* and

$$A_{LP}(j\omega) = H_0\frac{1}{\sqrt{2}}; \quad |A_{LP}(j\omega)|_{dB} = 20\ \log_{10}H_0 - 20\ \log_{10}\sqrt{2}$$

$$\therefore |A_{LP}(j\omega)|_{dB} = 20\ \log_{10}H_o - 3\ dB.$$

This means that the filter rolls off at 20 dB per 10-times increase in frequency (20 dB/decade) times the order of the filter. That is

the rate of attenuation = (order of filter) × (20 dB/decade).

Thus a first order filter rolls off at 20 dB/decade as shown in Figure 5.15.
This is sometimes expressed as a roll off of 6 dB/octave (where an octave is a doubling of frequency).

2.14.2 The high-pass active filter

The input resistor of the inverting amplifier is replaced by a resistor/capacitor pair to create a high-pass filter as shown in Figure 2.16.
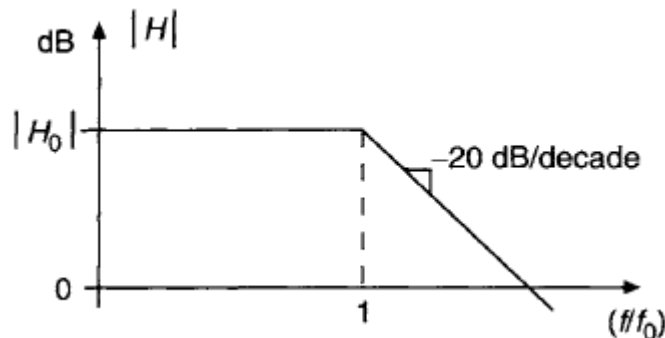We now analyze the filter as follows:

$$Z_F = Z_R = R_2$$

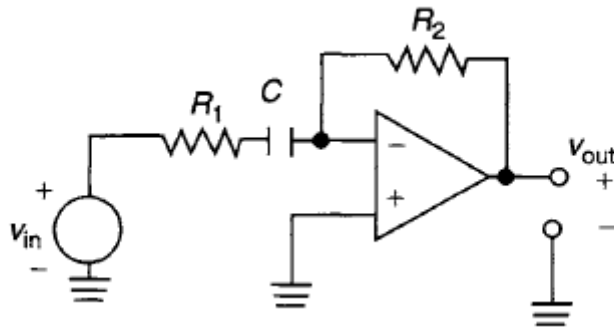**Figure 2.15** Frequency response of a single pole low-pass filter.

Figure 2.16 High-pass active filter.

$$Z_i = Z_R + Z_C = R_1 + \frac{1}{j\omega C}$$

$$A_{HP}(j\omega) = -\frac{Z_F}{Z_i} = \frac{-R_2}{R_1 + \frac{1}{j\omega C}} = \frac{-j\omega CR_2}{1 + j\omega CR_1}$$

$$A_{HP}(j\omega) = \frac{-j\omega CR_2}{1 + j\frac{\omega}{\omega_0}}.$$

The gain of this filter is given by:

$$\therefore A_{HP}(j\omega) = H_0 \frac{j\omega CR_2}{1 + j\frac{\omega}{\omega_0}},$$

where

$$H_0 = \left(\frac{-R_2}{R_1}\right); \quad \omega_0 = \frac{1}{CR_1}; \quad f_0 = \frac{1}{2\pi R_2 C}$$

As $\omega \rightarrow 0$, $A_{HP}(j\omega) \rightarrow 0$

As $\omega \rightarrow \infty$, $A_{HP}(j\omega) \rightarrow H_0$.
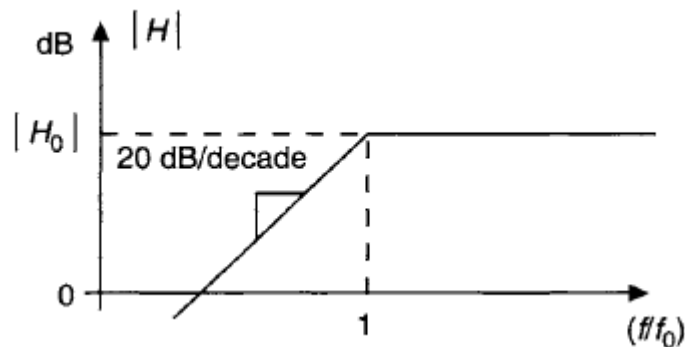

Figure 2.17

Frequency response of a single pole high-pass filter. Under the last condition, the filter acts as a linear amplifier. The frequency response of this filter is shown in Figure 2.17. The frequency ω_0 is called the 'cutoff frequency'; this is the point where the filter begins to filter out the higher-frequency signal.

2.14.3 The band-pass active filter

The band-pass active filter is a combination of the low-pass active filter and the high-pass active filter as shown in Figure 2.18. We analyze the circuit as follows:

$$Z_F = Z_R \| Z_C = R_2 \left\| \frac{1}{j\omega C_2} \right.$$

$$Z_F = \frac{R_2 \frac{1}{j\omega C_2}}{R_2 + \frac{1}{j\omega C_2}} = \frac{R_2}{1 + j\omega C_2 R_2}$$
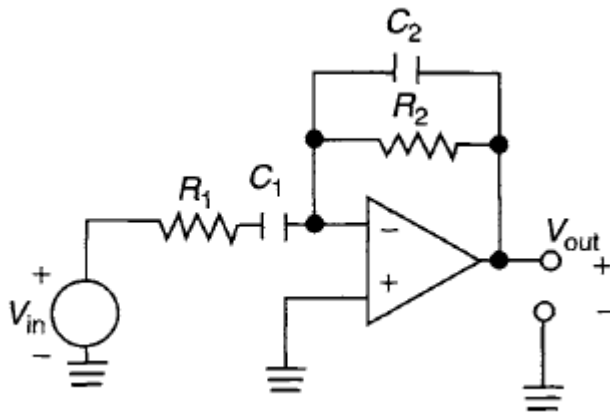

Figure 2.18 Band-pass active filter.
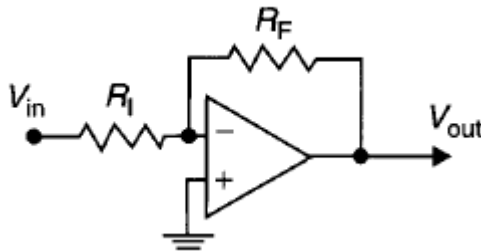
Figure 2.19 Circuit for Example 2.1.

$$Z_i = Z_R + Z_C = R_1 + \frac{1}{j\omega C_1} = \frac{1 + j\omega C_1 R_1}{j\omega C_1}.$$

The gain of this filter is given by:

$$A_{LP}(j\omega) = -\frac{Z_F}{Z_i} = -\frac{j\omega C_1 R_2}{(1 + j\omega C_2 R_2)(1 + j\omega C_1 R_1)},$$

where

$$\omega_1 = \frac{1}{C_1 R_2}; \quad \omega_{LP} = \frac{1}{C_2 R_2}; \quad \omega_{HP} = \frac{1}{C_1 R_1}; \quad \omega_{HP} > \omega_{LP}$$

EXAMPLE 2.1 In the circuit shown in Figure 2.19,

$$R_F = 8\,k\Omega; \quad R_1 = 4\,k\Omega; \quad V_i = 5\,mV.$$

Determine the output voltage.
Solution
This is an inverting amplifier, so from Equation 2.3:

$$\frac{V_{out}}{V_i} = -\frac{R_F}{R_I}$$

$$V_{out} = -\frac{R_F}{R_I} V_i = -\frac{8}{4} \times 5 \, mV = -10 \, mV$$

EXAMPLE 2.2
In the circuit shown in Figure 2.20,

$$R_F = 8 \, k\Omega; \quad R_I = 4 \, k\Omega; \quad V_s = 10 \, mV.$$

Determine (a) the voltage gain of this amplifier; (b) the output voltage.
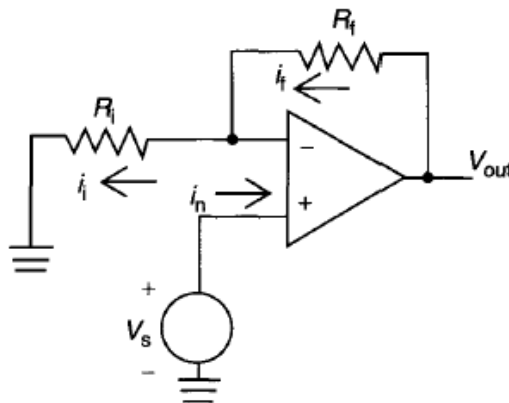

Figure 2.20 Circuit for Example 2.2.

Solution
This is a non-inverting amplifier, so from Equation 2.6:
(a) The voltage gain of this amplifier is

$$1 + \frac{R_f}{R_i} = 1 + \frac{8}{4} = 3.$$

(b) The output voltage is

$$V_{out} = \left(1 + \frac{R_f}{R_i}\right) V_s = 3 \times 10 \, mV = 30 \, mV.$$

EXAMPLE 2.3

In the circuit shown in Figure 2.21,

$$R_F = 10\,\text{k}\Omega; \quad R_1 = 6\,\text{k}\Omega; \quad R_2 = 6\,\text{k}\Omega; \quad R_3 = 6\,\text{k}\Omega;$$

$$V_1 = V_2 = V_3 = 5\,\text{V}.$$

Determine the output voltage.

Solution

This is a summing amplifier, so from Equation

$$V_{\text{out}} = -R_f \sum_{i=1}^{N} \frac{V_i}{R_i} = -10\left(\frac{1}{6} + \frac{1}{6} + \frac{1}{6}\right) \times 5\,\text{V} = -25\,\text{V}.$$
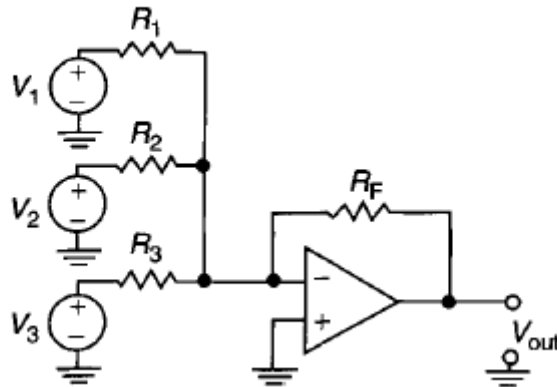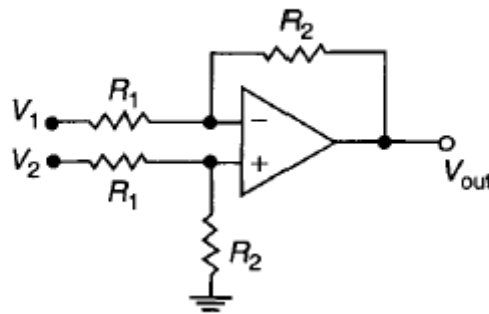

Figure 2.21 Circuit for Example 2.3.

Figure 2.22 Circuit for Example 2.4.

EXAMPLE 2.4

In the circuit shown in Figure 2.22,

$$R_1 = 5\,\text{k}\Omega; \quad R_2 = 15\,\text{k}\Omega;$$
$$V_1 = 10\,\text{V}; \quad V_2 = 15\,\text{V}.$$

Determine the output voltage for these values and for when $R^\wedge = R2$.

Solution

This is a difference amplifier, so from Equation 5.15:

$$V_{out} = \left(\frac{R_2}{R_1}\right)(V_2 - V_1) = \left(\frac{15}{5}\right)(15 - 10) = 15\,V.$$

When $R_1 = R_2$, the output voltage is 5 V.

EXAMPLE 2.5

For the instrumentation amplifier circuit shown in Figure 2.8,

$$R_1 = 5\,k\Omega; \quad R_2 = 20\,k\Omega; \quad R_3 = 15\,k\Omega; \quad R_G = 3\,k\Omega;$$
$$V_1 = 10\,V; \quad V_2 = 15\,V.$$

Determine (a) the gain; (b) the output voltage.

Solution

$$\text{(a)} \quad G = \left(\frac{20}{5}\right) \times \left(1 + \frac{2 \times 15}{3}\right) = 44.$$

$$\text{(b)} \quad V_{out} = \left(\frac{20}{5}\right)\left(1 + \frac{2 \times 15}{3}\right)(15 - 10) = 220\,V.$$

EXAMPLE 2.6 For the low-pass active filter circuit shown in Figure 2.14,

$$R_1 = 5\,k\Omega; \quad R_2 = 20\,k\Omega;$$

$$C = 2\,\mu F.$$

Determine the closed-loop gain and the decibel value at which the filter rolls off.

Solution

$$H_0 = \left|\frac{-R_2}{R_1}\right| = \frac{20}{5} = 4$$

$$\omega_0 = \frac{1}{2 \times 10^{-6} \times 20 \times 10^3} = 25$$

$$f_0 = \frac{25}{2\pi} = 3.98\,Hz$$

The filter roll off is given as

$$A_{LP}(j\omega) = H_0 \frac{1}{\sqrt{2}}; \quad |A_{LP}(j\omega)|_{dB} = 20\,\log_{10} H_0 - 20\,\log_{10}\sqrt{2}$$

$$\therefore |A_{LP}(j\omega)|_{dB} = 20\,\log_{10} H_0 - 3\,dB = 20\,\log_{10}(4) - 3\,dB = 9.04\,dB$$

CHAPTER 3
Data acquisition

3.1 Introduction

The purpose of most electronic systems is to measure or control some physical quantity, hence a system will need to acquire data from the environment, process this data and record it. Acting as a control system it will also have to interact with the environment.

Data acquisition can be divided into the steps shown in Figure 3.1. The flow of information in a typical data acquisition system can be described as follows:

1. The input transducers measure some property of the environment.
2. The output from the transducers is conditioned (amplified, filtered, etc.).
3. The conditioned analog signal is digitized using an analog-to-digital converter (ADC).
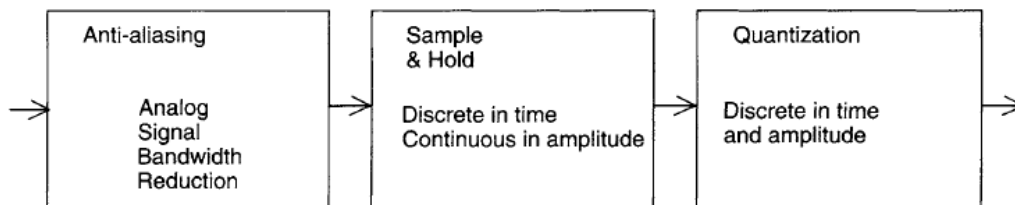4. The digital information is acquired, processed and recorded by the computer.

Figure 3.1 Data acquisition.

5. The computer may then modify the environment by outputting control signals. The digital control signals are converted to analog signals using a digital-to-analog converter (DAC).
6. The analog signals are conditioned (e.g. amplified and filtered) appropriately for an output transducer.
7. The output transducer interacts with the environment.

3.2 Sampling and aliasing

In many common situations in engineering, a function $f(i)$ is sampled. When a function is evaluated by numerical procedures, it is always necessary to sample the function in some manner, because digital computers cannot deal with analog, continuous functions except by sampling them. Often the function is not even explicitly defined, but only known as a series of values recorded on tape, by a data logger or a computer. Examples of where sampled time domain data is used are areas where computers and microcontrollers are used to automate processes and react to input data from the processes. It is also used in engineering applications, which include simple and complex vibration analysis of machinery, as well as measurements of other variables such as boiler pressures, temperatures, flow rates and turbine speeds and many other machine parameters.

3.2.1 Sampling

If A is the time interval between consecutive samples, then the sampled time data can be represented as:

$$h_n = h(\Delta n) \quad n = 0,\ \pm 1,\ \pm 2, \dots$$

Consider an analogue signal $x(t)$ that can be viewed as a continuous function of time, as shown in Figure 3.2(a). We can represent this signal as a discrete time
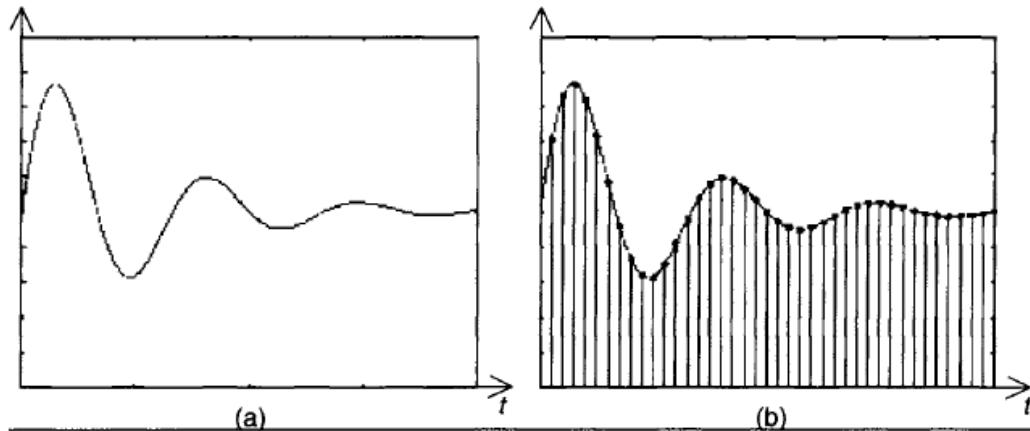
Figure 3.2 Digitizing a waveform: (a) analog signal; and (b) sampled equivalent.

signal by using values of $x(t)$ at intervals of nTs to form $x[n]$ as shown. In this case, we are mapping points from the function $x(t)$ at regular intervals of time, Ts, called the sampling period.

It is usual to specify a sampling rate or frequency f_s rather than the sampling period. The frequency is given by $f_s = 1/T_s$, where f_s is in Hertz. If the sampling rate is high enough, then the signal $x(t)$ can be constructed from $x[n]$ by simply joining the points by small linear portions, thus approximating to the analog signal. The discrete samples are digitized for processing by a computer or similar device (Figure 7.2(b)).

3.2.1.1 *The sampling theorem*
The sampling theorem, or more correctly Shannon's sampling theorem, states that we need to sample a signal at a rate at least twice the maximum frequency
component in order to retain all frequency components in the signal. This is expressed as

$$f_s > 2f_{max},$$

where fs is the *sampling rate* (frequency), $fmax$ is the highest frequency in the input signal, and the minimum required rate $(2fmax)$ is called the *Nyquist frequency*. The time interval between the digital samples is

$$\Delta t = \frac{1}{f_s}.$$

3.2.2 Aliasing
One would expect that if the signal has significant variation then Ts must be small enough to provide an accurate approximation of the signal $x(t)$. Significant signal variation usually implies that high frequency components are present in the signal. It could therefore be inferred that the higher the frequency of the components present in the signal, the higher the sampling rate should be. If the sampling rate is not high enough to sample the signal correctly then a phenomenon called aliasing occurs as

shown in Figure 3.3. In other words, we do not correctly obtain the frequency in the original signal if a signal is sampled at less than twice its maximum frequency component.

The term aliasing refers to the distortion that occurs when a continuous time
signal has frequencies larger than half the sampling rate. The process of aliasing describes the phenomenon in which components of the signal at high frequencies are mistaken for components at lower frequencies. The frequency domain view of sampling is that, when a continuous time signal is sampled, its spectrum will show the aliasing effect if aliasing occurs because regions of the frequency domain will be shifted by an amount equal to the sampling frequency.
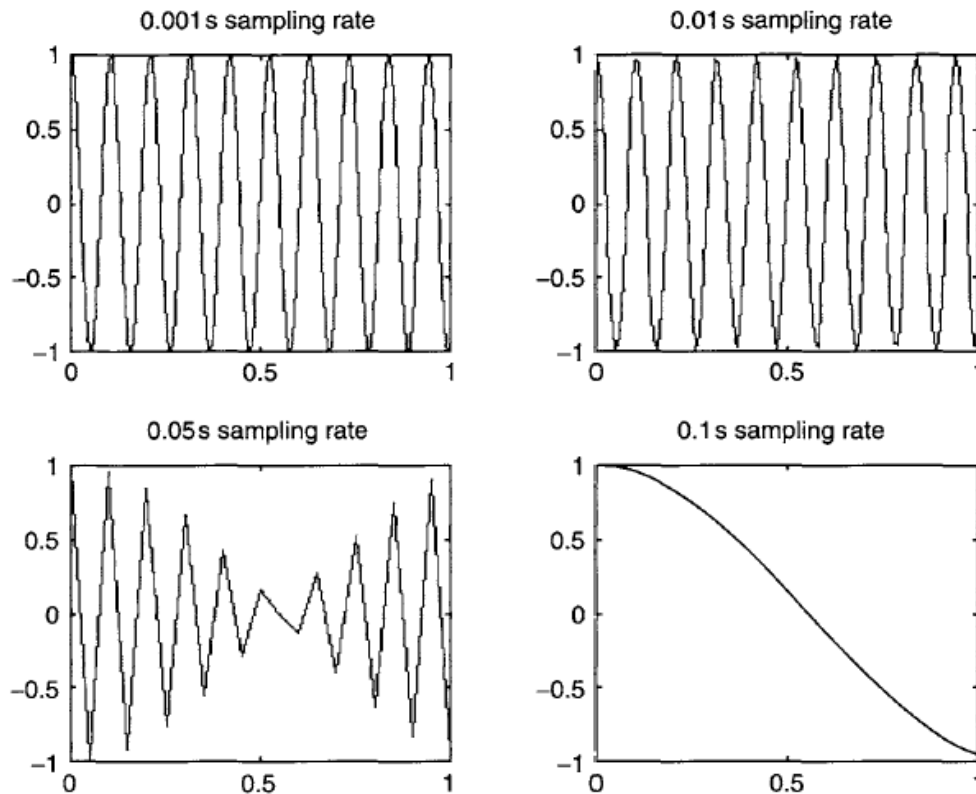

Figure 3.3 The effect of different sampling rates when sampling the waveform cos(600-

3.2.2.1 *Anti-aliasing and the Nyquist sampling theorem*

The Nyquist sampling theorem states that to avoid aliasing occurring in the sampling of a signal, the sampling rate should be greater than or equal to twice the highest frequency present in the signal. This is referred to as the Nyquist sampling rate. Many scientific observation instruments critically sample data to permit high throughput scanning and minimal data transfer. Minimal sampling refers to sampling at a rate that corresponds to twice the highest frequency content in the source signal. The frequency that corresponds to the sampling rate, 1/2, is known as the Nyquist frequency (1/7). Instruments that minimally sample a signal may pose one of the most difficult problems associated with the activity of signal enhancement. Resolution enhancement increases the high frequency content from observed data, since we are narrowing the narrowest features in the observation. Potential problems exist when we increase the resolution of observed data. The required sampling interval shortens to maintain the critical sampling criteria. If the data is sampled at the critical sampling

rate then that sampling rate is less than that necessary to properly sample the enhanced data, then we must be concerned with our effects on the observed data. It is important to understand what effect aliasing has on our observations. In some instruments, data is sampled at a rate less than that necessary to properly characterize data at the Nyquist frequency, that is, the data is under sampled. It has been the general consensus that a mild under sampling of the observable does not produce significant fictitious information to appear in the under-sampled observation.

The Nyquist criterion dictates that all signals must be bandlimited to less than half the sampling rate of the sampling system. Many signals already have a limited spectrum, so this is not a problem. However, for broad spectrum signals, an analog low-pass filter must be placed before the data acquisition system. The minimum attenuation of this filter at the aliasing frequency should be at least:

$$A_{\min} = 20 \log\left(\sqrt{3} + 2^B\right),$$

3.2.2.2 *Problems with the anti-aliasing filter*

• **Time response:** In designing an anti-aliasing filter, there is a temptation to have it's attenuation roll-off extremely quickly. The way to achieve this is to increase the order of the filter. A so-called brick-wall filter (one with infinitely high order), however, causes a sine function time response that decays proportionally to *l/t*. What this means is that an extremely high order filter that eliminates all signals above the cut-off frequency will cause signals that change rapidly to ring on for a long time. This has a very undesirable effect. • **Phase distortion/time delay:** Most analog filters have a non-linear phase response. This a problem since non-linear phase causes an unequal time (group) delay as a function of frequency. The higher frequency signals will arrive later than low frequency signals. This can especially be a problem when multiple sensor outputs are compared such as when using a microphone array.

• **Amplitude distortion:** By definition, the filter will modify the frequency structure of the sensor signal which is usually not desired. The solutions to these problem include:

• Increase the sampling rate of the ADC. This allows the anti-aliasing filter to have a higher cut-off frequency and still eliminate aliasing. This enables the following:

1. the filter roll off can be more shallow, allowing a better time response;
2. the frequency response of the filter does not attenuate the lower sensor frequencies of interest;
3. phase distortion is strongest around the cut-off frequency of the filter so if this is pushed higher, it will not affect the sensor frequencies.

• Use linear phase filters. This, of course, will reduce the phase distortion problems.

3.3 Quantization theory

Analog to digital conversion is a two-step process, which changes a sampled analog voltage into digital form. These processes are quantization and coding.

Quantization is the transformation of a continuous analog input into a set of data represented by discrete output states. *Coding* is the assignment of a digital code word or number to each output state as shown in Figure 3.4. The number of possible states N is equal to the number of bit combinations:

$$N = 2^n,$$
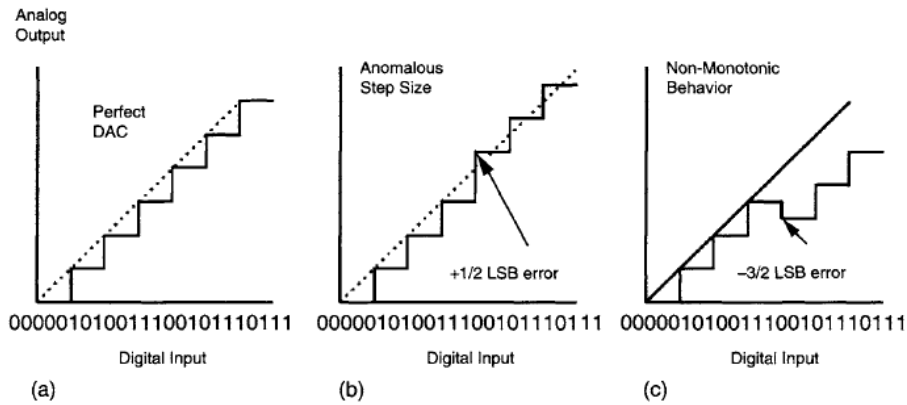
where n is the number of bits.

Figure 3.4 Quantization.

The *analog quantization size* (or *resolution) Q* is defined as the full scale range of the ADC divided by the number of output states: Knowing the resolution, *Q,* and range, *R,* of an ADC, we can determine the number of bits required as follows:

$$Q = \frac{(V_{\max} - V_{\min})}{2^n - 1} = \frac{R}{2^n - 1}$$

$$2^n = 1 + \frac{R}{Q}$$

$$n = \frac{\log\left(1 + \frac{R}{Q}\right)}{\log 2}$$

EXAMPLE 3.1
Determine the smallest step size (resolution) of a 4-bit ADC, which has a maximum output voltage of 12 V.

Solution

$$Q = \frac{(12-0)}{2^4-1} = \frac{12}{15} = 0.8\,\text{V}$$

EXAMPLE 3.2

Determine the smallest step size (resolution) of an 8-bit ADC in the range _5Vto+5V.

Solution

$$Q = \frac{(5-(-5))}{2^8-1} = \frac{10}{255} = 0.04\,\text{V}$$

EXAMPLE 3.3

An ADC has a range of 5 V and a resolution of 5 mV. Determine the number of bits required.

Solution

$$n = \frac{\log\left(1 + \frac{5}{0.005}\right)}{\log 2} = 9.97 = 10\,\text{bits}$$

3.4 Digital-to-analog conversion hardware

The process of converting a number held in a digital register to an analog voltage or current is accomplished with a digital-to-analog converter (DAC). The DAC is a useful interface between a computer and an output transducer.

3.4.1 Binary-weighted ladder DAC

Current summing and IC devices are used to build DACs. DACs are nothing more than operational amplifiers whose gains can be programmed digitally. An inverting op amp has a selection of fixed resistances which scales a varying input voltage.

DACs fix the input voltage but switch a series of input resistors to vary the gain and voltage output level.

DACs are normally switched current devices designed to drive the currentsumming junction of an operational amplifier. In Figure 3.5 all the input voltages are equal to the reference voltage *VR*. Each input resistor is twice as large as the one preceding it. Thus the construction is called a binary-weighted ladder. Each current is proportional to the value of a bit position in a binary number. The output voltage from each input, is one-half the value produced by the preceding input:

$$\text{1-bit only} \quad V_{out} = -\frac{V_R}{2}$$

$$\text{2-bit only} \quad V_{out} = -\frac{V_R}{4}$$

$$\text{3-bit only} \quad V_{out} = -\frac{V_R}{8}$$

$$\vdots \qquad \vdots$$
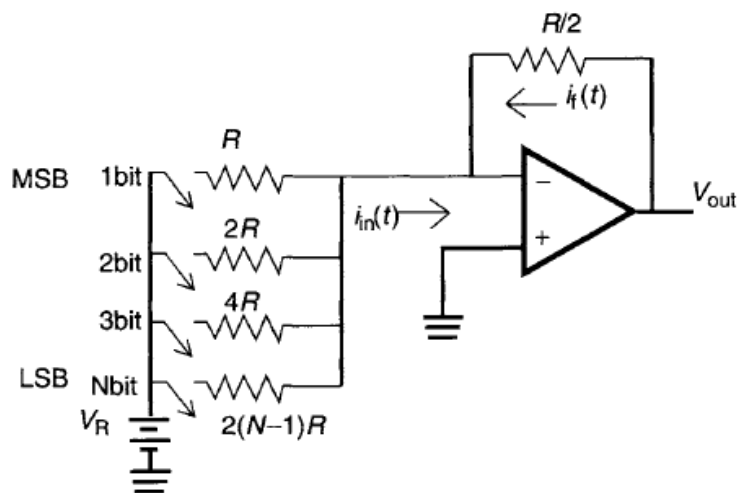
$$N\text{-bit only} \quad V_{out} = -\frac{V_R}{2^N}$$

Figure 3.5 The binary-weighted ladder DAC.

The LSB, or bit N, produces the minimum step size, called the resolution,

$$Q = \frac{V_R}{2^N}$$

The full-scale output occurs when all bits are closed:

$$V_{out} = V_R\left(\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \cdots + \frac{1}{2^N}\right) = -V_R\left(1 - \frac{1}{2^N}\right).$$

As the number of bits increases, the full scale output approaches the reference voltage but is never equal to it. Generalizing for an arbitrary binary input, we have where $b\backslash =$ MSB and $bN=$ LSB. The design requires a number of different precisely defined resistor values. We can improve the circuit by replacing it with a circuit that requires fewer distinct resistor values.

EXAMPLE3.4
Determine the output of the DAC for an input of 10011001.

Solution

$$V_{\text{out}} = V_R\left(\frac{1}{2} + \frac{0}{4} + \frac{0}{8} + \frac{1}{16} + \frac{1}{32} + \frac{0}{64} + \frac{0}{128} + \frac{1}{256}\right) = -\frac{153}{256}V_R$$

3.4.1.1 *Limitations of binary-weighted ladder DAC*
The limitations of binary-weighted ladder DAC are as follows: • resistance values require precision trimming (difficult to achieve); • increasing the resolution leads to the RN becoming very large; • their practical realization has stray capacitance (pF), which combined with very large resistances can cause undesirable RC time constants or slow conversion times.

3.4.2 Resistor ladder DAC
The simplest type of DAC converter and perhaps the most popular single package DAC is the resistor ladder network connected to an inverting summer op amp circuit. It requires two precision resistance values *(R and 2R)* and resolves the problems of the binary-weighted ladder. Each digital input bit in the circuit controls a switch between ground and the inverting input of the op amp as shown in Figure 3.6. If the binary number is 0001, only the *b0* switch is connected to the op amp, and the other bit switches are grounded. The resistance between node *V0* and ground is *R* since it is the parallel combination of two *2Rs*. Therefore, *V0* is the result of voltage division of *Vx* across two series resistors of equal value *R: V0= V\/2*. Similarly, we can obtain *V\ = V2/2*
and F2=F3/2.

Consequently, $V_0 = V_3/8 = V_s/8$.
Since the gain of the inverting amplifier is

$$-\frac{R}{2R} = -\frac{1}{2},$$

the analog output voltage corresponding to the binary input 0001 is $V_{out0} = -V_s/16$.

Similarly, for the inputs 0010, 0100, and 1000, $V_{out1} = -V_s/8$, $V_{out2} = -V_s/4$, and $V_{out3} = -V_s/2$ respectively.

The output for any combination of bits is $V_{out} = b_3 V_{out3} + b_2 V_{out2} + b_1 V_{out1} + b_0 V_{out0}$.
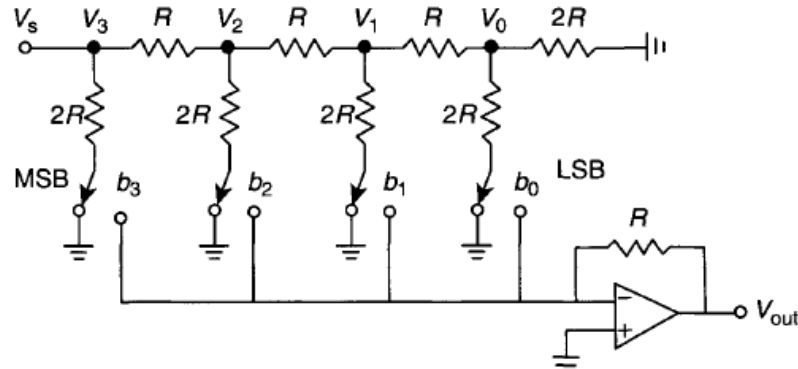

Figure 3.6 The resistor ladder DAC.

3.4.3 DAC limitations

The output of a DAC can only assume discrete values. The relationship between the input binary number and the analog output of a perfect DAC is shown in Figure 3.4. The figure shows output signals from DACs showing (a) the ideal result, and (b) a differential non-linearity or (c) non-monotonic behavior, both caused by imperfectly matched resistors. Common DAC limitations are an anomalous step size between adjacent binary numbers, non-monotonic behavior, or a zero output.

3.5 Analog-to-digital conversion hardware

The purpose of the analog to digital converter is to digitize the input signal from the sample and hold circuit to 2^B discrete levels, where B is the number of bits of the ADC. The input voltage can range from 0 V to V_{ref} (or $-V_{ref}$ to $+V_{ref}$ for a bipolar ADC). What this means is that the voltage reference of the ADC is used to set the range of conversion of the ADC. For a monopolar ADC, a 0 V input will cause the converter to output all zeros. If the input to the ADC is equal to or larger than V_{ref} then the converter will output all ones. For inputs between these two voltage levels, the ADC will output binary numbers corresponding to the signal level.

There are three widely used ADC technologies today: (1) flash or parallel converter, (2) successive-approximation converters and (3) dual slope converters. These are discussed in the following sections.

To properly convert an analog voltage signal for computer processing the following elements must be properly selected and used in this sequence (Figure 3.7):

1. Buffer amplifier (chosen to provide a signal in a range close to but not exceeding the full input voltage range of the ADC).
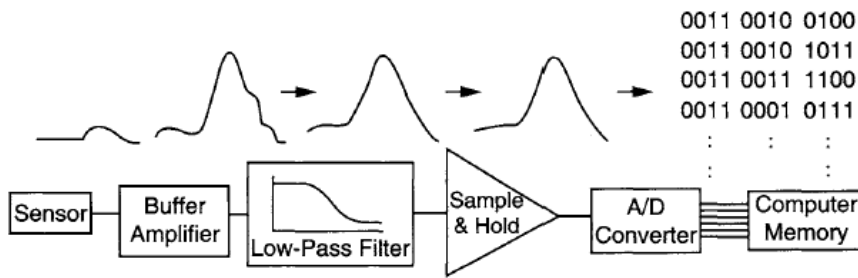
Figure 3.7 ADC block diagram.

2. Low-pass filter (to remove undesirable high-frequency components in the signal that could produce aliasing).

3. Sample and hold amplifier (to maintain a fixed input value during the short conversion time).

4. Analog to digital converter (ADC)

5. Computer

Other chapters discuss the buffer amplifier and the low-pass filter. Here we look at the sample and hold amplifier.

3.5.1 The sample and hold circuit

The sample and hold circuit takes a 'snapshot' of a signal level at a particular time in readiness for digitization. A signal value (sample) is stored in a capacitor which is prevented from discharging using FETs. An illustrative sample and hold circuit made from discrete components is shown in Figure 3.8. A dual op amp is ideal for this purpose, because their input bias currents are practically zero. The FET is used to isolate the capacitor from the source used to charge it. A 'hold' is initiated by connecting HOLD to —12 V, and leaving it disconnected for a 'sample'. Note the back-to-back signal diodes at the output of the first op amp (any diodes can be used). The purpose of these diodes is to 'catch' the output when the feedback loop is broken when the FET turns OFF. The resistor R isolates this action from the output feedback loop. If these diodes are not present, the op amp saturates at the supply rail and the FET will not turn off. It is important to test the circuit, and to make sure it *holds* properly, and observe the droop of the output. The choice of hold capacitor is important. The leakage of an electrolytic and the transient behavior of ceramics rule them out completely in this application.
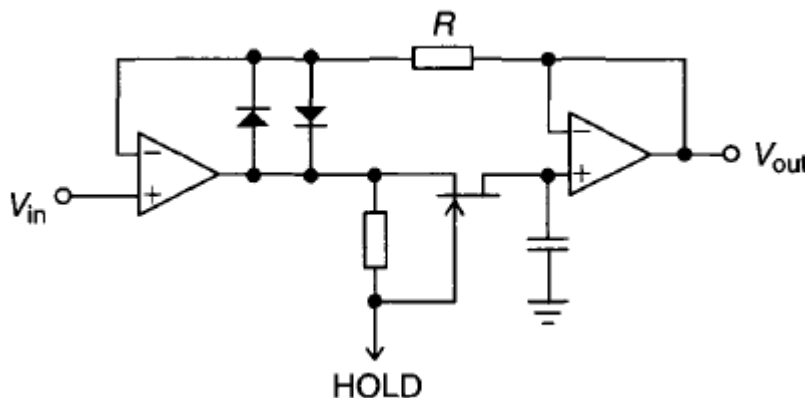

Figure 3.8 Sample and hold circuit.

The best choice is probably polypropylene, and after that polystyrene or Mylar. Polycarbonate is the least preferred. The greatest problem (after leakage, which

should be practically zero) is dielectric hysteresis in which the voltage changes on charge and discharge are not the same. There is also dielectric absorption, where there is a 'memory' of past states. A capacitor freshly discharged may acquire a small voltage as time goes by. All of these phenomena are the result of the complexity of dielectric structure and behavior. Some properties of sample-and-hold circuits are important in critical, dynamic applications. The hold step is the change in output voltage when the circuit is switched OFF, the result of various capacitive effects. The settling time is the time required after the HOLD command for the output to stabilize. The aperture time is the time after the HOLD command at which changes in the input have no effect. The acquisition time is the time at which the output settles after a change at the input. Finally, the dynamic sampling error is the difference between the voltage held and the instantaneous input voltage at the instant of the *hold* command. Everything necessary for a sample-and-hold except the hold capacitor can be put on chip, so monolithic sample and hold circuits are available and very easy to use. The sample/hold command is given through a digital logic level, so these circuits interface directly with logic.

3.5.1.1 *Limitations of a sample and hold*
• Finite aperture time: The sample and hold takes a period of time to capture a sample of the sensor signal. This is called the aperture time. Since the signal will vary during this time, the sampled signal can be inaccurate. • Signal feedthrough: When the sample and hold is not connected to the signal, the value being held should remain constant. Unfortunately, some signal does bleed through the switch to the capacitor, causing the voltage being held to change slightly. • Signal droop: The voltage being held on the capacitor starts to slowly decrease over time if the signal is not sampled often enough.
• The main solution to these problems is to have a small aperture time relative to the sampling period. This means that if the designer uses a high sampling rate, the aperture time of the sample and hold must be quite small.

3.5.2 Noise problem
Because the ADC outputs only 2^B levels there is inherently noise in the quantized output signal. The ratio of the signal to this quantization noise is called SQNR. The SQNR (in dB) is approximately equal to six times the number of bits of the ADC: 20 log(SQNR) = 6 x bits. Therefore, for a 16-bit ADC this means that the SQNR is approximately equal to 96 dB. There are, of course, other sources of noise that corrupt the output of the ADC. These include noise from the sensor, from the signal conditioning circuitry, and from the surrounding digital circuitry. The key to reducing the effects of the noise is to maximize the input signal level. What this means is that the designer should increase the gain of the signal conditioning circuitry until the maximum sensor output is equal to the Fref of the ADC. It is also possible to reduce Fref down to the maximum level of the sensor. The problem with this is that the noise will corrupt the small signals. A good rule of thumb is to keep Fref at least as large as the maximum digital signal, usually 5 V.

3.5.3 The parallel-encoding (flash) ADC
The parallel-encoding or flash ADC design provides the fastest operation at the expense of high component count and high cost (Figure 3.9). For a 2-bit converter, the relationship between the code bit Gt and the binary bits Bt is

$$B_0 = G_0 \cdot \overline{G}_1 + G_2 \text{ and } B_1 = G_1$$

Table 3.1 shows the comparator output codes and corresponding binary outputs for each state, when an input voltage range of 0 V to 4 V is assumed.
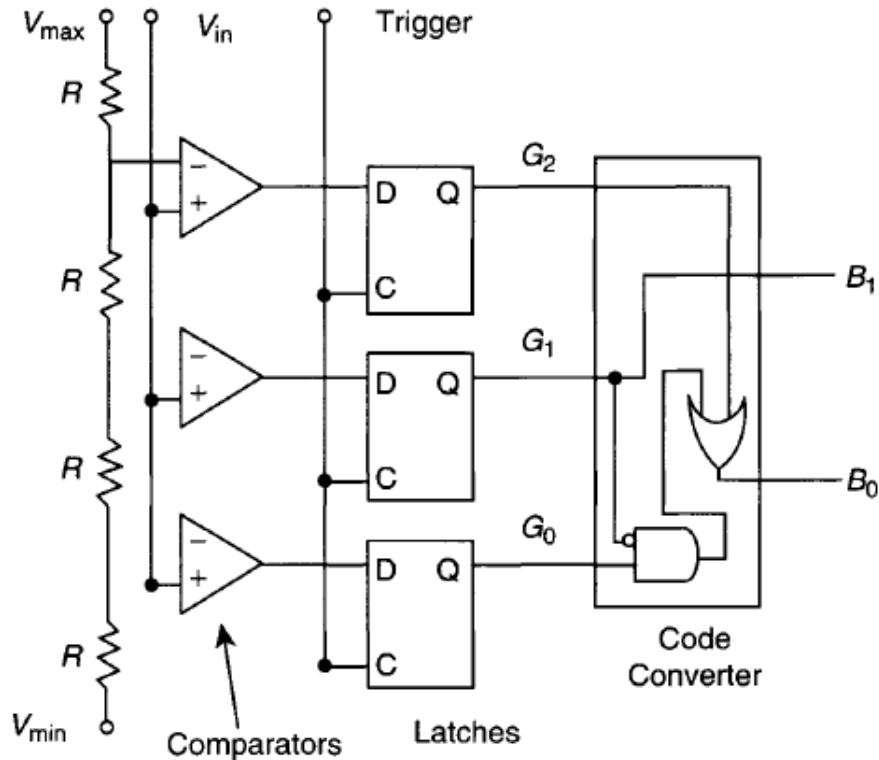

Figure 3.9 The flash ADC.

State	Code $(G_2 G_1 G_0)$	Binary $(B_1 B_0)$	Voltage range
0	000	00	0–1
1	001	01	1–2
2	011	10	2–3
3	111	11	3–4

3.5.4 The successive-approximation ADC

The successive-approximation ADC is the most commonly used design (Figure 3.10). This design requires only a single comparator and will be only as good as the DAC used in the circuit. The analog output of a high-speed DAC is compared against the analog input signal. The digital result of the comparison is used to control the contents of a digital buffer that both drives the DAC and provides the digital output word. The successive-approximation ADC uses fast control logic, which requires only *n* comparisons for an «-bit binary result.
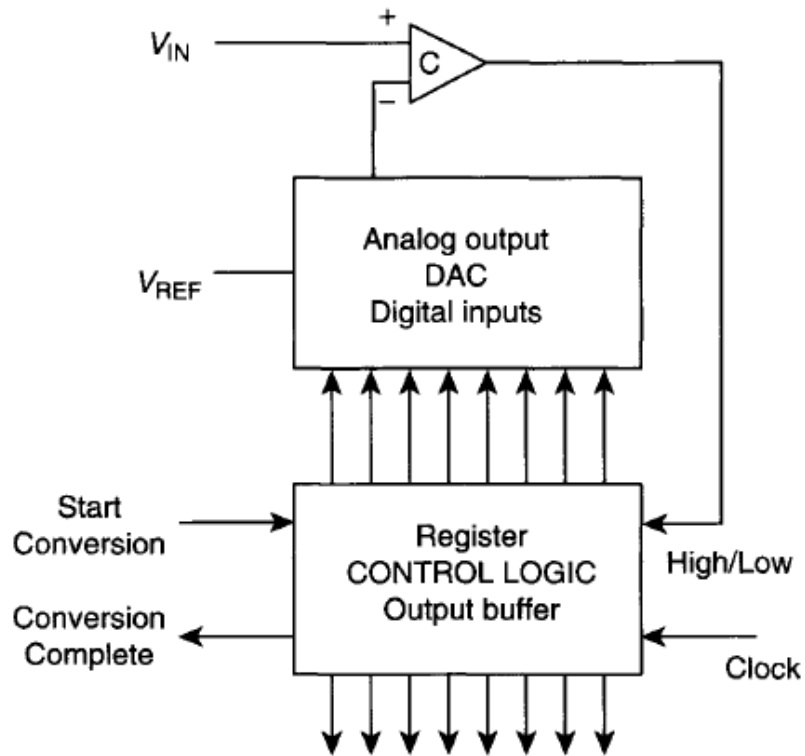
Figure 3.10 Block diagram of an 8-bit successive-approximation ADC.

3.5.4.1 *The successive-approximation procedure*

The procedure involved in a successive-approximation ADC (see Figure 3.11(a)) is as follows:

• When the start signal is applied, the sample and hold (S&H) latches the analog input.

• The control unit begins an iterative process where the digital value is approximated, converted to an analog value with the DAC, and compared to the analog input with the comparator.
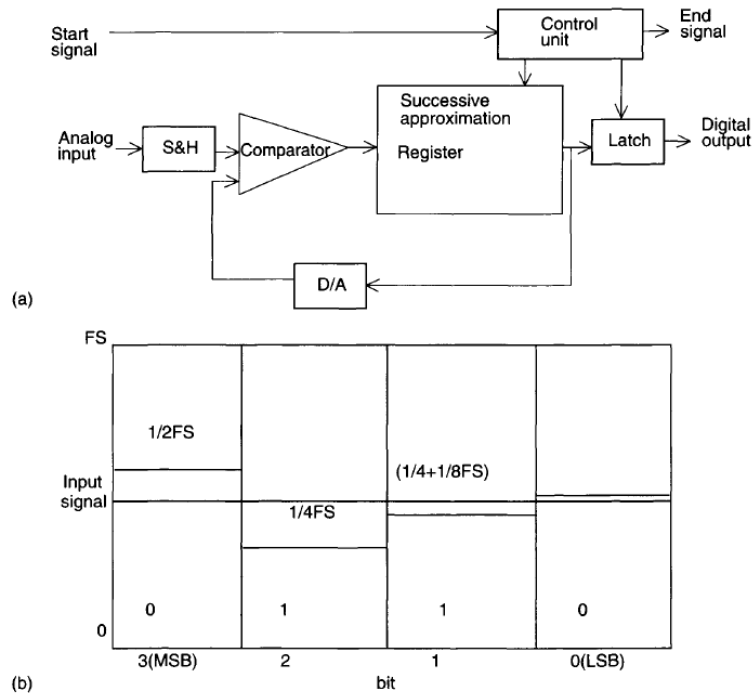
Figure 3.11 Example of a 4-bit successive-approximation ADC.

3.5.5 The dual-slope ADC

The limitations associated with the DAC in a successive-approximation ADC can be avoided by using the analog method of charging a capacitor with a constant current; the time required to charge the capacitor from zero to the voltage of the input signal becomes the digital output. When charged by a constant current the voltage on a capacitor is a linear function of time and this characteristic can be used to connect the

- When the DAC output equals the analog input, the control unit sets the end-signal and the correct digital output is available at the input.
- The input is compared to standard values in a decreasing binary sequence defined by $1/2, 1/4, 1/8, \ldots, 1/n$ of the full scale (FS) value of the ADC.

We illustrate the successive approximation method using an example shown in Figure 7.11(b), in which the MSB is $1/2$ FS, which in this case is greater than the signal; therefore, the bit is turned off. The second bit is $1/4$ FS and is less than the signal, so it is turned on. The third bit gives $1/4 + 1/8$ of FS, which, is still less than the input signal, so the third bit is turned on. The fourth provides $1/4 + 1/8 + 1/16$ of FS and is greater than the signal, so the fourth bit is turned off and the conversion complete.

We note that an n-bit ADC has a conversion time of $n\Delta T$, where ΔT is the cycle time for the DAC and control unit.

A *unipolar* output is either positive or negative, but not both. A *bipolar* output ranges over negative and positive values. We note that a single ADC can digitize several analog signals, using an analog *multiplexer*, which simply switches between several analog inputs. Important parameters in selecting an ADC are the input voltage range, output resolution, and conversion time.

analog input voltage to the time as determined by a digital counter. Although there are several versions of the integrating method: single slope, dual slope, and multiple, we discuss the dual-slope ADC, which is shown in Figure 3.12 since it is the most common.

An unknown voltage is applied to the input where an analog switch connects it to an integrator. The integrator drives a comparator. Its output goes *high* as soon as the integrator output is more than several millivolts. When the comparator output is *high,* an AND gate passes clock pulses to a binary counter. The binary counter counts pulses until the counter overflows. This time period Tx is fixed by the clock frequency fc and counter size M, and is independent of the unknown voltage. However, the integrator output voltage is proportional to the unknown voltage:

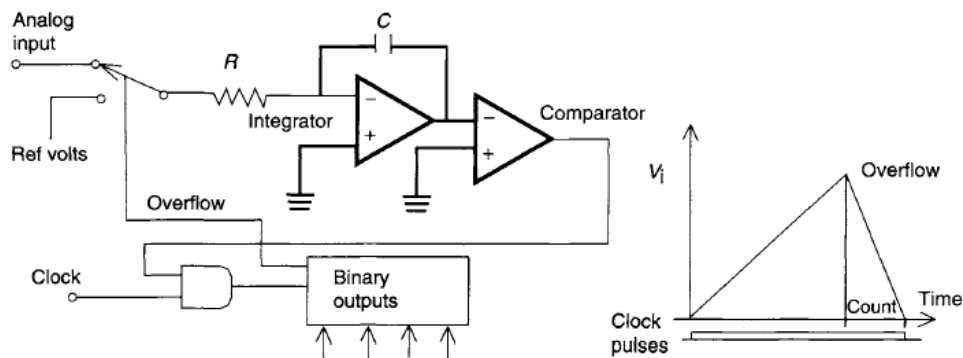$$V_1 = \left(\frac{T_1}{RC}\right)V_{in} = \frac{MV_{in}}{f_c RC}$$

Figure 3.12 Dual-slope ADC.

When the counter overflows, it resets to zero and sends a signal to the analog switch, which disconnects the unknown voltage and connects a reference voltage. The polarity of the reference voltage is opposite that of the unknown voltage. The integrator voltage decreases at a rate proportional to the reference voltage. In the meantime the counter has started counting again from zero at the instant the reference voltage is applied to the integrator. When the integrator output reaches zero, the comparator goes *low,* bringing the AND gate LOW. Clock pulses no longer pass through, and the counter stops counting at *m.*

$$V_1 = \left(\frac{T_2}{RC}\right)V_{ref} = \frac{mV_{ref}}{f_c RC}.$$

The base of the triangle, $V\backslash$ (Figure 3.12(b)) is common to both, hence, the count is

$$m = \left(\frac{M}{V_{ref}}\right)V_{in}.$$

The accuracy of the dual slope method depends only upon Fref. The integrator action of the converter averages out random negative and positive contributions over the sampling period $T\backslash$, and hence it is excellent in rejecting noise.