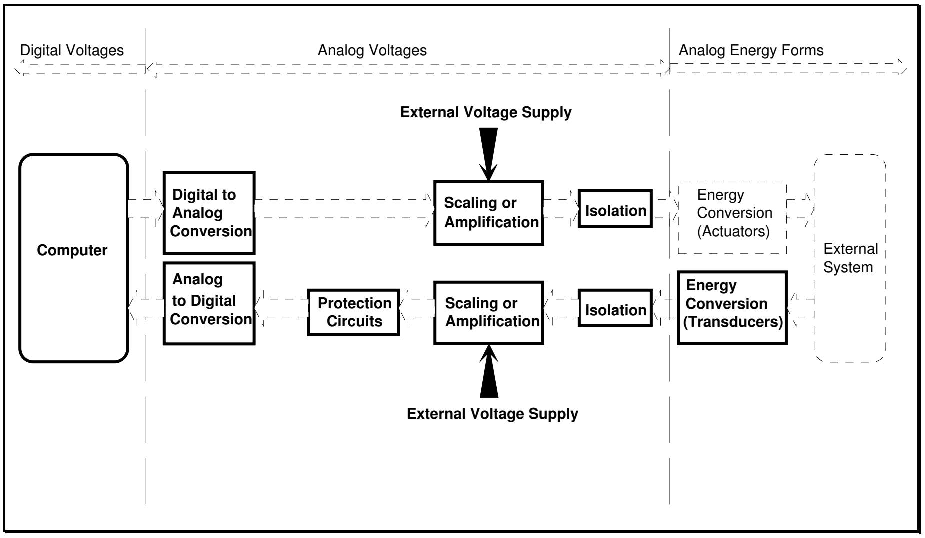
Chapter 7

Interfacing Computers to Mechatronic Systems

A Summary...

A chapter covering the basic technical issues related to designing and selecting hardware to interface computers to mechatronic devices. This chapter delves into the basic aspects of the closed control loop in terms of protection, isolation and signal conversion between external systems and the computer. The chapter also looks in some detail at the problems of digital control including quantisation error in A/D conversion and the problems associated with transducers.



7.1 Introduction

Most modern devices that engineers encounter in the industrial environment are mechatronic in nature. That is, they are a combination of some or all of the following elements:

- Mechanical components (both moving and stationary)
- Electro-mechanical components (motors, solenoids, relays, etc.)
- Traditional power circuits (single and three-phase power supplies)
- Power electronic circuits (servo drives, switched-mode power supplies, etc.)
- Digital circuits (discrete combinational logic and microprocessor logic)
- Analog circuits (transistor amplifiers, etc.)
- Energy conversion devices (transducers)
- Software.

In simple terms, mechatronic systems involve the application of digital and analog electronics and computer systems to the control of both traditional and modern mechanical devices. Since nearly every one of the elements in a mechatronic system is very nearly the basis of a professional discipline in its own right, the task of designing such devices and interfacing them to other computer systems can be quite daunting. No single text can therefore cover in absolute detail the spectrum of issues related to designing and interfacing to mechatronic systems and this text certainly does not endeavour to do so.

This text, and this chapter in particular, should be viewed as an index that gives the reader an overview of the problems involved in interfacing digital electronics (particularly computers) to industrial systems. The title of this chapter and this book are somewhat ambiguous and deliberately so. The reason for the ambiguity is that sometimes we need to interface computers to mechatronic systems that already have computer control and other times we need to interface computers to mechatronic systems that may have analog and digital power electronics, but no computer control. In both cases, the introduction of a computer ultimately creates a larger mechatronic system and hence the ambiguity.

Interfacing industrial signals to computers (or computer-based devices) can be a complex and time consuming task. There are many problems involved in isolating low-voltage, low-current computer circuits from the high voltages and currents that are prevalent in industrial environments. However, these problems are just one part of the overall interfacing dilemma. Converting signals from one energy form to another or reducing or amplifying signal levels is the other part of the problem.

In the engineering market-place there are an enormous number of commercially available (off-the-shelf) interfacing boards and transducers that have been designed to allow computers to interact with the outside world. These can satisfy the majority of our interfacing needs. Off-the-shelf interfacing boards are normally ideal for one-off applications because they save on development and testing costs. Unfortunately, commercial boards are normally "general-purpose" and can therefore be relatively expensive because they are designed to handle a range of interfacing needs. If we need to use large volumes of interfaces for specific applications, then we generally find that these commercial solutions are no longer viable. As with most commercial decisions, there is a natural "break-even" point and beyond this point, we discover that is necessary to develop our own, special-purpose interfaces from basic concepts.

There are a number of fundamental concepts that need to be understood, in terms of computer interfacing, and this text is designed to introduce you to some of the more important ones in the sense that they will pertain to mechatronic systems. In Chapter 3, we examined a range of different electrical and electronic devices that are used as mechanisms for interfacing computers to external systems. However, in order to understand how these mechanisms fit in to the interfacing process, we need to understand the process itself. To this end, our task, for the remainder of this chapter, will be to examine the global objectives of interfacing and to see how the sorts of devices, introduced in Chapter 3, can be used to interface computers to external systems. If you have not already done so, then you should go through the concepts espoused in Chapter 3 before proceeding with this chapter.

By the time you have completed this chapter, you should clearly understand that there is actually nothing unique about the problem of interfacing a computer to a mechatronic system. The issues involved in interfacing to chemical systems, large-scale power systems, etc., are generic and all require an application of the basic principles that will be introduced in this chapter.

7.2 The Interfacing Process

The purpose of computer interfacing is to make the software on a computer system communicate with an external system that provides or accepts information in various physical forms and sizes. More precisely, we can say that signals must be transferred from the external system to the computer's CPU and memory devices via the data bus in order to create a feedback path. Signals must be transferred from the CPU and memory devices, via the data bus, to the external system to create a driving force. This input and output is normally used to create a closed-loop control system, as shown schematically in Figure 7.1 or a data-acquisition (monitoring) system if the computer is not required to provide any driving force as an output.

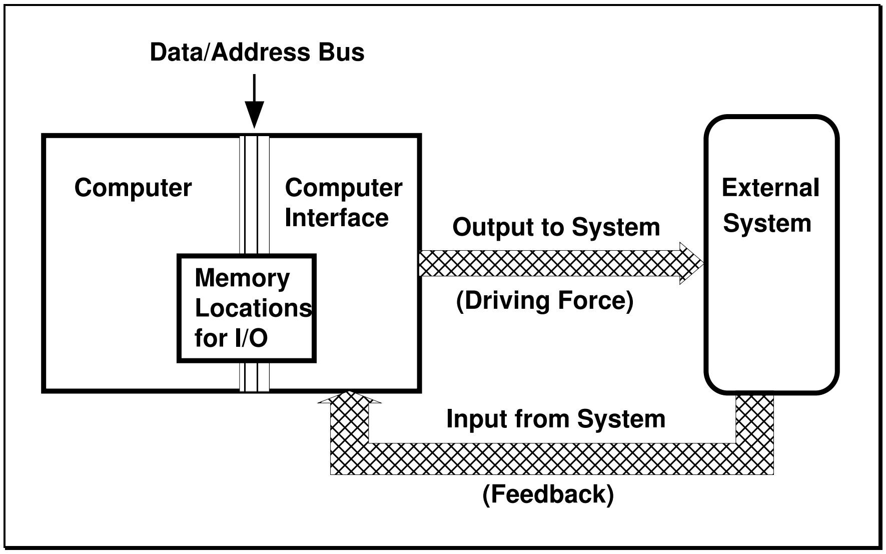


Figure 7.1 - The Computer Interfacing Process - Closed Loop Control

The computer in Figure 7.1 could be a general-purpose personal computer, a specially designed (one-off) microprocessor or for that matter, a mini-computer or main-frame system. The problems involved in interfacing any of these devices to the external system of Figure 7.1 are numerous. The key point to remember in understanding these problems is that the circuits in computer systems are generally designed to:

- Respond to small, digital voltages (typically less than ten volts) and not currents or pressures or temperatures
- Provide small digital voltage outputs (typically less than ten volts) and very small current outputs (typically less than one milli-Amp).

However, our knowledge of most physical systems requiring control tells us that the external systems:

- Do not generally provide voltage feedback and are not necessarily voltage driven (other energy and signal forms are also prevalent including current, temperature, pressure, capacitance, inductance, etc.)
- Often do not provide or require digital signals (many feedback and drive signals are analog in nature)
- Generally provide and require signals with energy levels having magnitudes not compatible with the computer system (signals may be too large or too small and need to be converted)

As a result of this incompatibility, there are a number of steps involved in the interfacing process. These are shown in Figure 7.2 for the closed-loop system.

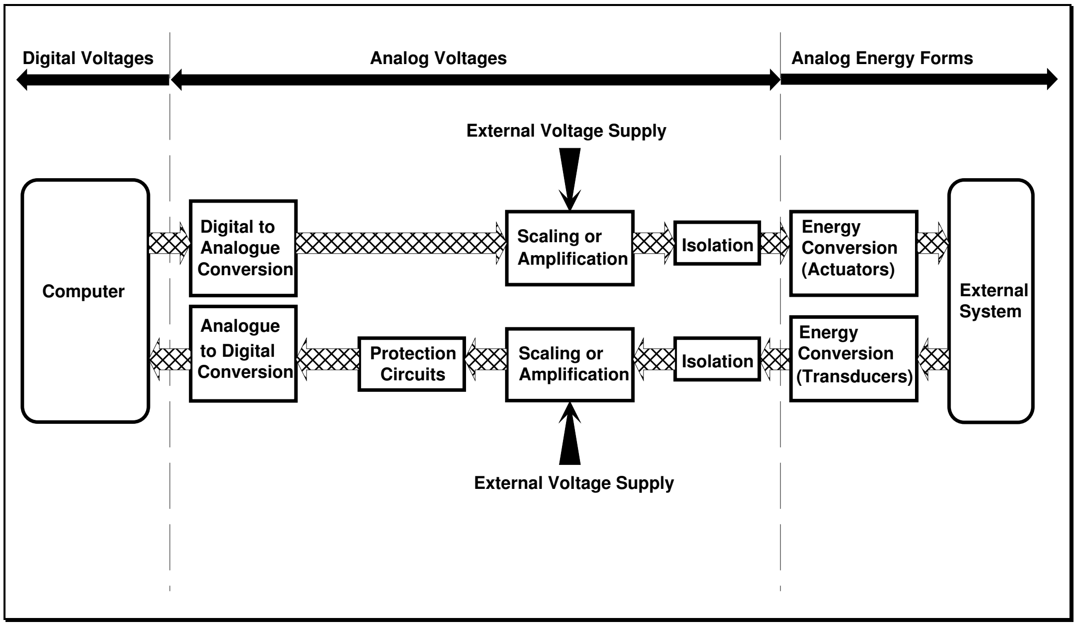


Figure 7.2 - Basic Steps in the Computer Interfacing Process

For the signals fed from the external system to the computer, the following steps are generally required:

(i) Conversion from initial energy form to voltage (via transducers)
(ii) Conversion of raw voltage to levels appropriate to computer circuits
(iii) Protection and isolation of the computer circuits from raw signals where spuriously high signals can occur in the external system
(iv) Conversion of analog signal voltages to digital form.

For the signals fed from the computer to the external system, the following steps are generally required:

(v) Conversion from digital form to analog form
(vi) Conversion of raw analog output to appropriate voltage and current levels (amplification and/or transformation)
(vii) Conversion of output currents and voltages to appropriate drive energies (magnetic field, mechanical force, etc.).

However, in examining Figure 7.1 and Figure 7.2, it must also be remembered that any signals fed back from the external system into the computer are effectively "asynchronous" or random. In other words, their timing has no relationship to the carefully timed activities within the computer system and its data and address bus structure. The first objective of computer interfacing is to therefore provide a mechanism that will accept this random, incoming information and retain it until it can be released by normal, internal addressing and timing techniques to the data bus and CPU.

As far as the computer software is concerned, all incoming and outgoing data must ultimately appear in memory locations which are mapped within the normal computer address/data bus structure. In order to get an interface to provide this functionality, we have to convince the CPU that the entire interface is nothing more than a collection of memory addresses.

There are a range of special chips whose internal registers can be mapped into a computer system (just like memory chips) and are equipped with input/output ports so that a range of asynchronous, incoming digital signals can be received or transmitted. These devices are often referred to as "Programmable Parallel Interfaces" (abbreviated to PPI), "Peripheral Interface Adaptors" or "Parallel Interface Adaptors" (normally abbreviated to PIA). The actual name usually depends upon the manufacturer The integration of these chips into a computer system is shown schematically in Figure 7.3.

The specific functionality of each PPI chip depends upon the manufacturer and a range of different chips are available to meet different requirements in terms of inputs and outputs, number of internal registers and so on. Generally, a PPI provides a number of "ports" to which external system connections are made. Each port consists of a number of digital lines that can be used for input or output. Each port has a corresponding register that is mapped onto the address/data bus structure of the computer system, as schematically shown in Figure 7.3.

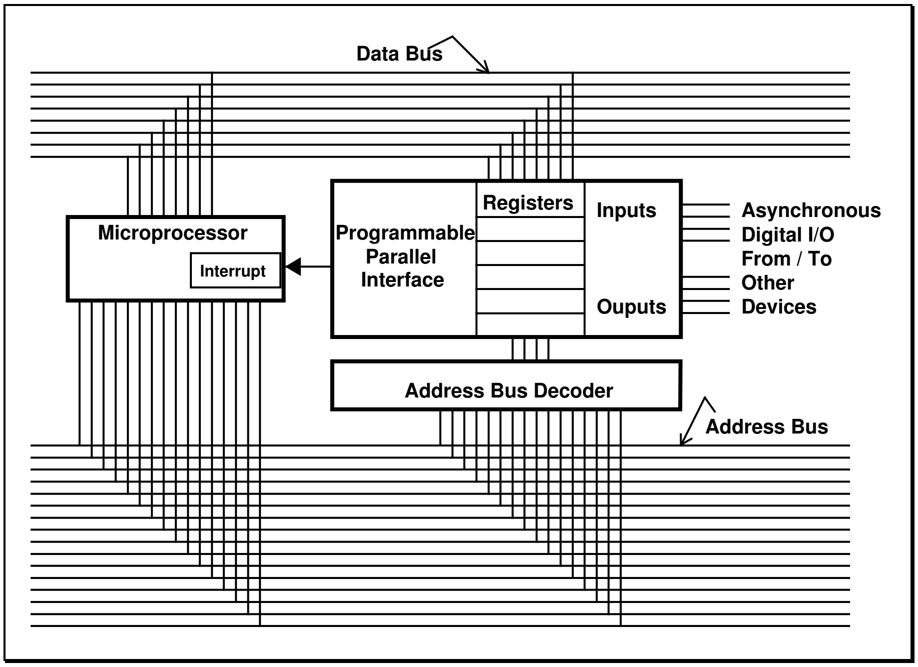


Figure 7.3 - Interfacing External Asynchronous Signals to the Synchronous Internal Environment via a Programmable Parallel Interface

The CPU reads from the registers (which the CPU views as nothing more than memory locations) in order to get the current status of input ports and writes to registers in order to change the status of output ports. The PPI is therefore a device which performs the conversion between the asynchronous and synchronous data forms found outside and inside the computer, respectively. Most PPI devices have additional registers that provide the CPU with information about the status of incoming data and the PPI itself. Additional registers are also normally provided so that the operation of the PPI can be changed by the CPU. This attribute is common to many peripheral devices (including UARTs, etc.) whose characteristics can be programmed by changing the contents of internal registers.

We have already noted that a substantial proportion of external feedback signals will be analog and similarly, that the driving forces required by external systems will also, in general, be analog. The PPI essentially only provides digital inputs and outputs. However, there are numerous single-chip devices available to perform the conversion between analog and digital (and vice-versa). The self-evident titles of these devices are "Analog to Digital (A/D) Converters" and "Digital to Analog (D/A) Converters". The functionality of these devices is shown schematically (in conjunction with the PPI chip) in Figure 7.4.

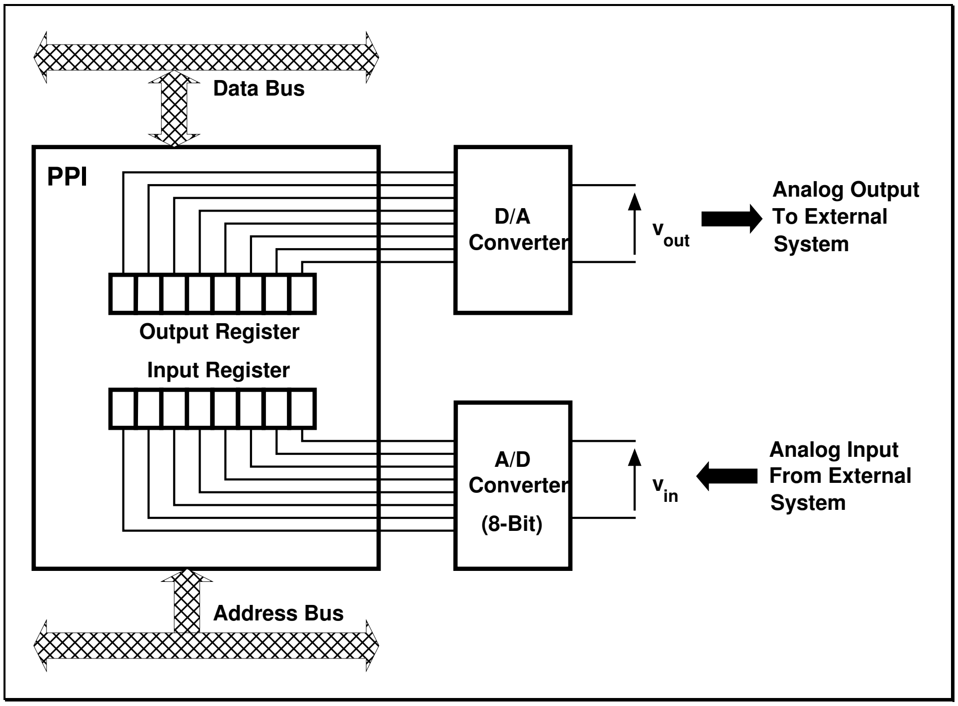


Figure 7.4 - Functionality of A/D and D/A Devices Interacting with a PPI Device

In Figure 7.4, we have shown an A/D converter which takes in an analog voltage (within a limited range) and converts it to an 8-bit form that can be fed into an input port of the PPI device. Similarly, we have shown an output port of the PPI feeding 8 bits of data (in parallel) into a D/A converter that then provides a voltage proportional to the number represented by the 8 bits. The actual number of bits to which analog to digital (& vice-versa) conversion occurs clearly defines the accuracy of the incoming data or outgoing driving force and is referred to as the "resolution" of the device. As a general rule, the higher the resolution of the device, the higher the cost. The conversion of data from and to an analog form takes time, and again, the faster the conversion, the higher the cost of the device.

As a rule, the cost of A/D conversion is normally much higher than that of D/A transformation. As a result, it is often necessary to share a single A/D chip amongst a number of incoming analog voltages and to selectively switch between them. The device that performs this selective switching is referred to as a "multiplexer". We shall examine the A/D and D/A processes in more detail in section 7.3 because they are most important to the stability of a control system and the accuracy of a data acquisition system.

Another interesting point to note is that it is possible to purchase the combined functionality of analog to digital conversion with the functionality of the programmable parallel interface, thereby having a single, memory-mapped device that is capable of accepting external analog signals. These inputs are accessed through registers in the same way as those in the discrete PPI device.

Once the PPI chip is in place within a computer system, the "control software" developer needs to generate a range of sub-programs (procedures) which allow the main program to read and write to those "memory locations" (which are in fact Input/Output registers of the adaptor). These programs normally have to be written in a low-level assembly language.

We have already noted that as far as the internal computer hardware is concerned, external feedback signals arrive at essentially random times. As far as the software is concerned, we generally have little or no control over the rate at which feedback data enters the registers in the PPI device. Normally the control software running on the CPU also has to perform many functions in addition to reading and using data from the interface adaptor. It therefore takes a finite amount of time between successive readings of data from the adaptor. A program which regularly checks inputs, in between performing other functions, is said to be "polling". The polling technique is simple and efficient, provided that inputs don't change too quickly. However, if the time taken to use the incoming data is longer than the refresh time then data in the I/O registers may be lost (ie: over-written).

The "interrupt programming technique" is more complex than the polling technique but is used to avoid data loss in real-time systems. In an interrupt-driven system, the PPI "interrupts" normal program execution whenever new data arrives. This can be done (as in Figure 7.3) by directly interrupting the CPU or (more commonly) by asserting an interrupt on a special "interrupt controller" chip. Control of the CPU is then transferred to an Interrupt Service Routine (ISR), whose task it is to read data from the PPI and place it into a conventional area of memory (a buffer which is later accessed by the main program). Once control of the CPU is returned to the main program, it can use the data at its own pace. The development of an appropriate ISR is therefore an important part of the total interfacing process. The polling and interrupt techniques have already been discussed in some detail in section 6.7. A number of other issues related to software development in control systems will be covered in more detail in Chapter 8.

As we progress through the remainder of this chapter, we will examine the major issues related to the hardware interfacing of a computer system to an external system as we have cited them in Figure 7.2. After we have examined these aspects in more detail, we shall return to the overall interfacing problem in section 7.8, where we can bring the basic elements together with a greater understanding of the broader problems.

7.3 A/D and D/A Conversion

The analog to digital (A/D) and digital to analog (D/A) conversion processes are two of the most important aspects of computer control. Although both of these devices are most commonly fabricated into commercially-available, single chip implementations, an understanding of their design and behaviour is important to selecting the appropriate devices for a particular task. The two devices are shown schematically in Figure 7.5.

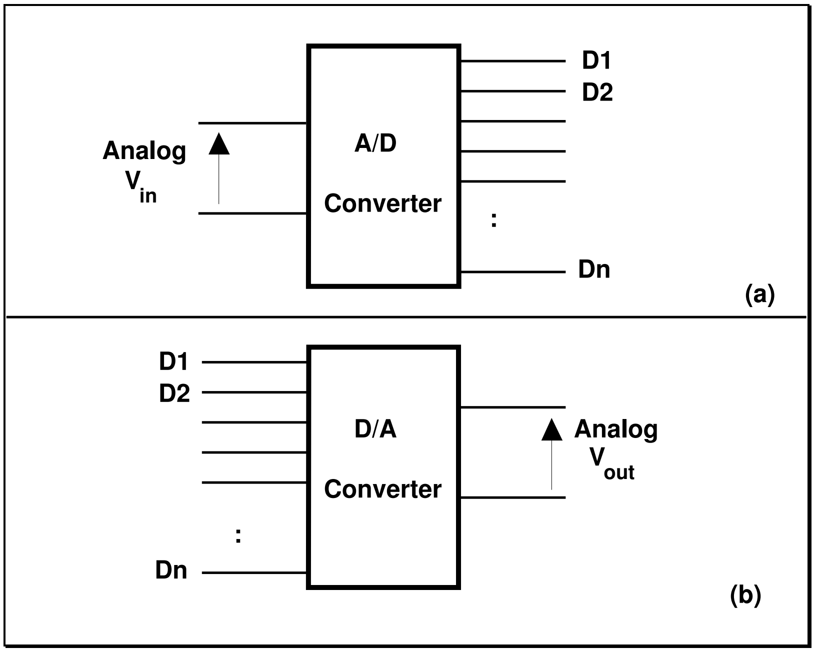


Figure 7.5 - (a) Schematic Representation of A/D Converter (b) Schematic Representation of D/A Converter

Figure 7.5 (a) shows the A/D device which provides a digital number output that is an approximation of the analog input. The accuracy of the approximation depends upon the resolution of the A/D converter, which in turn depends upon the number of output bits. For an A/D device of "n" bit resolution (as shown in Figure 7.5 (a)), there are 2^n possible combinations of output. If the analog input voltage range is V_i, then the A/D can only change when the input voltage changes by a value greater than or equal to:

$$\frac{V_i}{2^n} \qquad ...(1)$$

Incremental voltage changes of less than this size are not reflected in the binary output and so, information is lost. For example, an 8-bit A/D with an analog input range of 10 volts can only detect a change of voltage greater than or equal to:

$$\frac{10}{2^8} = 39.0625\text{mV}$$

If we were seeking to develop a control system that could respond to changes of 10 mV, then clearly an 8-bit A/D would not be appropriate in this instance.

Figure 7.5 (b) shows a D/A converter that provides an analog voltage whose value is dependent on the binary number at the input. The resolution of the D/A is also defined by the number of input bits into the system. The analog output of the D/A is therefore quantised and, for an n-bit converter, with an output voltage range of V_o, the output can only change in increments of:

$$\frac{V_o}{2^n} \qquad \text{...(2)}$$

The loss of data involved in the A/D and D/A conversion processes is referred to as "quantisation error". Quantisation error can be minimised by increasing the bit resolution of the A/D and D/A devices, but it can never be eliminated. Hence, if one were to take an analog signal, pass it through an A/D and then through the "dual" process, D/A, the original signal could not be recovered. The objective, therefore, in selecting A/D and D/A devices is to ensure that the approximated signals entering and leaving the computer are sufficiently accurate to ensure stable control.

D/A conversion is easier to realise than A/D conversion and is actually used within some A/D circuits. For this reason, we shall firstly examine the D/A process. Figure 7.6 shows the traditional way of achieving a conversion from binary input voltages to an analog output voltage. The technique is referred to as an "R-2R Resistive Ladder Network" and forms one side of the inputs to an operational amplifier, connected up to act as an inverting amplifier (See Figure 3.45 (b)).

In Figure 7.6, the digital inputs to the D/A device are each connected to a voltage controlled switch (either a BJT or FET). When a low input is provided to the switch, then the corresponding component of V_{REF} (from the resistive ladder) is switched to the non-inverting terminal of the operational amplifier (which is connected to ground), thus making no contribution to the output. A high digital input, on the other hand provides a voltage to the inverting terminal of the operational amplifier. The resistive ladder is designed to provide an input weighting corresponding to the order of the bit at the digital input. In the simple, 3-bit converter of Figure 7.6, the weighting ascribed to the digital signal input to switch 1 is twice that of switch 2, which is correspondingly, twice that of switch one. This provides the required relationship between the binary digital inputs and the analog outputs.

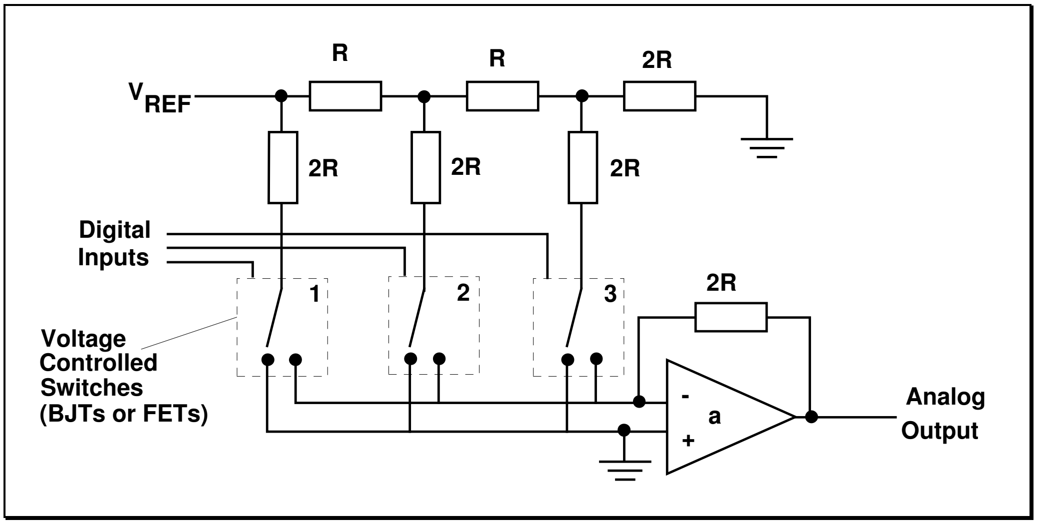


Figure 7.6 - Schematic of D/A Converter Operation (3-Bit)

The problem with the D/A technique shown in Figure 7.6 (which is widely used) is that the voltages representing the digital inputs may not all be the same. With TTL, for example, the voltage representing a binary one can be anywhere between 2.4 and 5.0 volts. Ordinarily, the variation in digital voltages is of no consequence, provided that it is within a given range - however, in D/A conversion, digital voltage variations can cause problems because they are unfortunately amplified into an analog output value. The problem is resolved by the voltage reference circuitry in the D/A converter, which performs appropriate level shifting to compensate for variations in input voltages. The voltage reference circuit also controls the range of analog output values which can be set (within limits) for a given range of digital inputs.

Analog to digital converters can be designed in a range of different ways. The fastest of the A/D converters are the so-called "flash converters" or "parallel converters". In order to understand the operation of these devices, one needs to understand the operation of a "comparator". A comparator is simply a specialised form of operational amplifier that provides a digital output voltage based upon the difference between two, analog input voltages. The device is shown in Figure 7.7.

The flash-converter A/D is composed of a number of comparators, each of which is given a different reference voltage, via a resistive ladder network and the incoming voltage. The comparators effectively then sort the voltage into different levels, depending upon the reference levels with which they work. As for the D/A circuit, reference levels are set up so that they effectively form a base-2 number system. The result is shown in Figure 7.8.

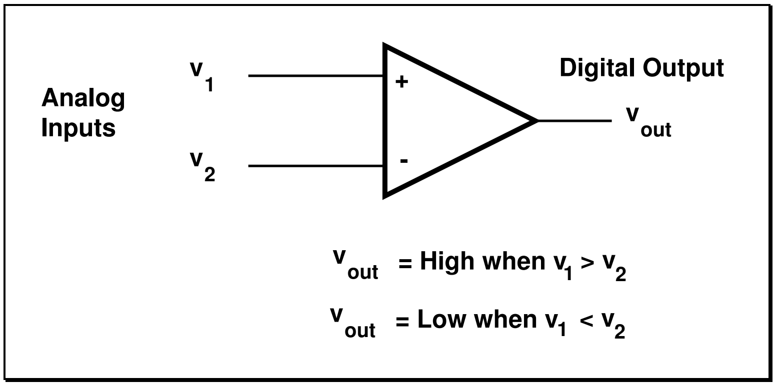


Figure 7.7 - Comparator Circuit Formed from an Operational Amplifier

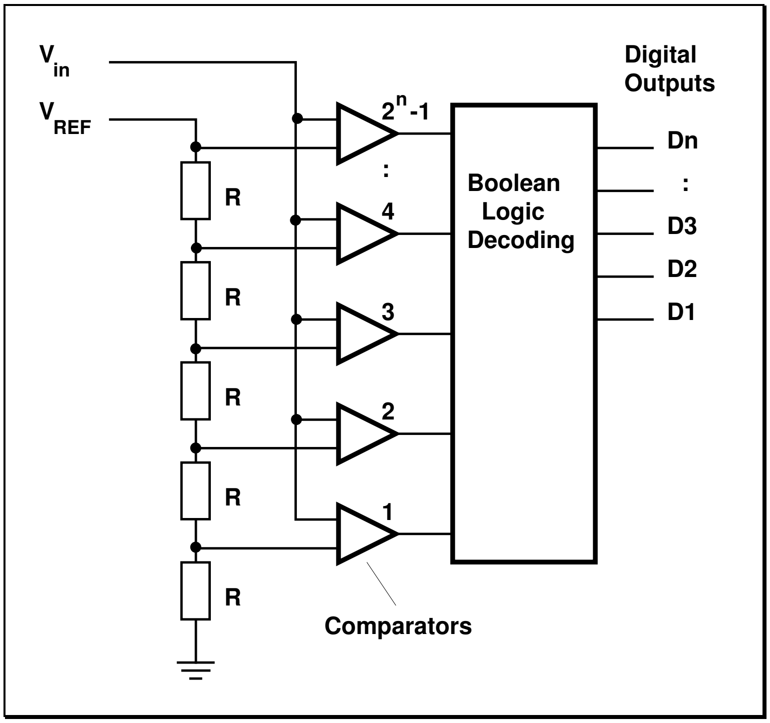


Figure 7.8 - Flash or Parallel A/D Converter

The problem with the flash-converter is that for n-bits of resolution it requires:

$2^n - 1$

comparators, which makes it unwieldy for any high levels of resolution (eg > 4 bits).

There are two, more commonly used A/D conversion circuits that are far more practical than the flash converter, despite the fact that they are not as fast. These are:

- Integrating Converters
- Successive Approximation Converters.

The operation of both these devices is relatively easy to understand. The integrating A/D converter is composed of two operational amplifiers (one configured as an integrator and the other as a comparator), a counter circuit and some relatively simple Boolean control logic. The device is shown schematically in Figure 7.9.

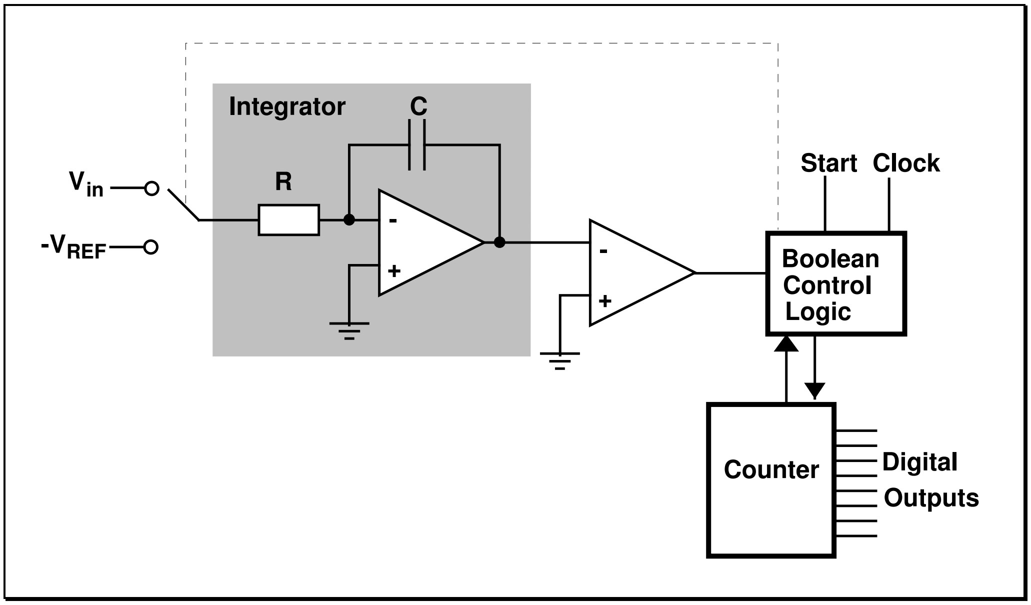


Figure 7.9 - Schematic of Dual-Slope, Integrating A/D Converter

In order to start the A/D conversion process in Figure 7.9, the Boolean control logic switches the analog input signal voltage to the integrator and at the same time switches on the counter. The output voltage of the integrator is the integral of the incoming analog waveform and increases with time. When the counter has reached a fixed number of counts, the integration process is halted by switching the input of the integrator to the negative reference voltage and resetting the counter. Given the negative reference voltage, the output of the integrator decreases until it reaches zero volts, which triggers the comparator and switches off the counter. The resulting digital output of the counter is then proportional to the analog voltage. The results of a typical conversion are shown in Figure 7.10. The dual-slope process is used to make the conversion independent of component tolerances.

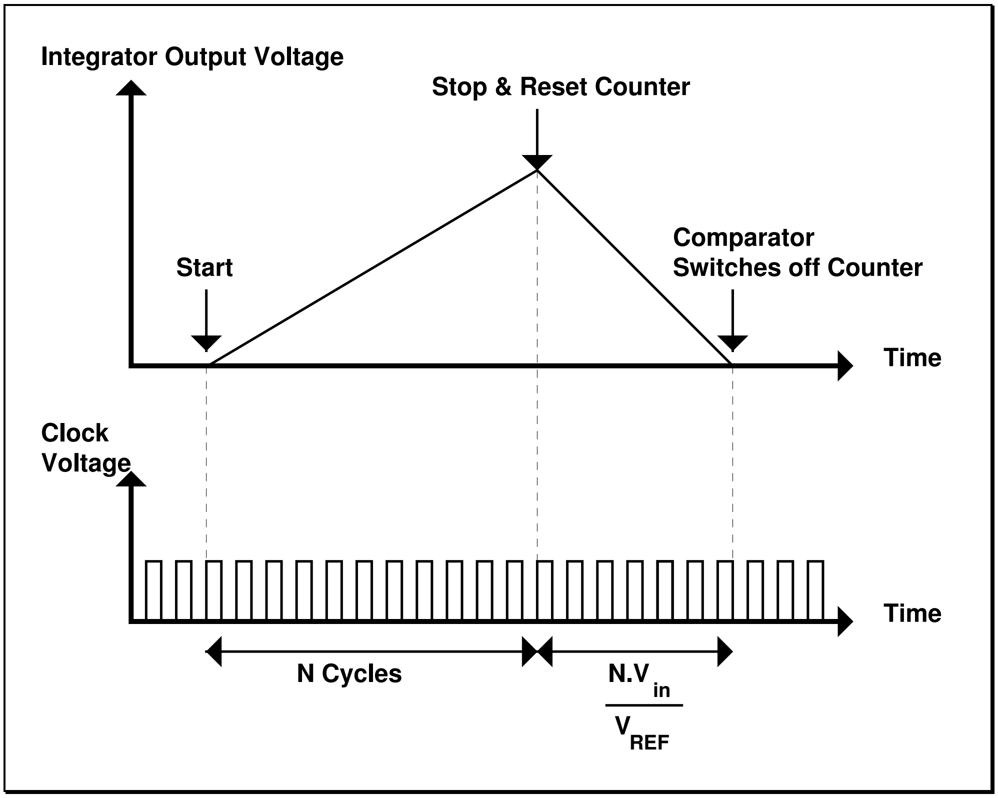


Figure 7.10 - Operation of Dual-Slope Integrating A/D Converter

The problem with the integrating A/D converters is that their conversion time is dependent upon the magnitude of the input voltage. This can be observed by studying Figure 7.10. The maximum height of the integrator output is determined by the size of the input voltage (since the slope of the waveform is proportional to the input voltage). The negative reference voltage is fixed and so the down-going edge of the integrator output voltage always has the same slope. Hence, the higher the input voltage the longer the time taken to obtain a conversion to a digital representation, since the number of counts increases with input voltage.

The successive approximation A/D converter is another device that is conceptually easy to understand. It is a simple, closed-loop system in which a digital number, stored in a register, is converted to analog (via an internal D/A converter) and then compared with the incoming analog system (via a comparator). When the two voltages match, the value stored in the register is an accurate representation of the incoming analog signal. The successive approximation A/D is shown schematically in Figure 7.11. The conversion control logic is implemented in the form of a state machine.

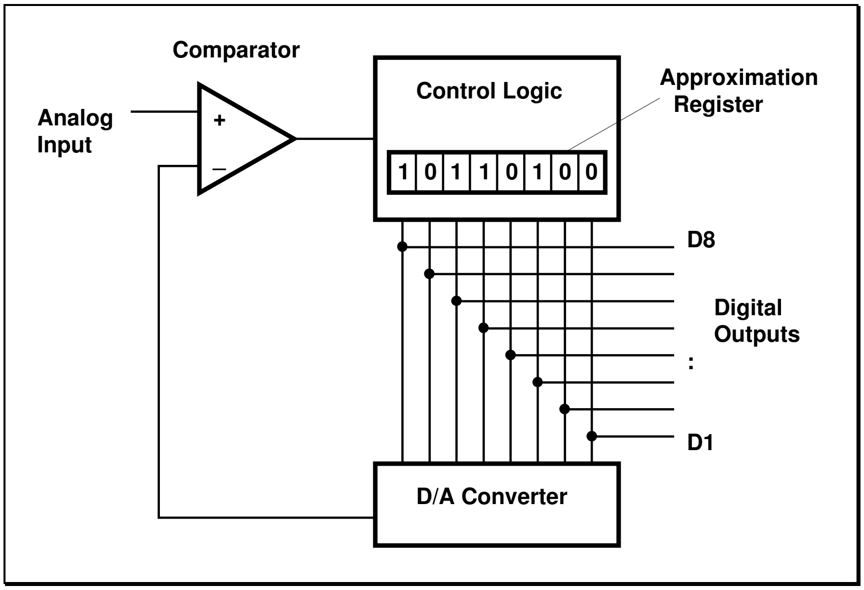


Figure 7.11 - Schematic of 8-Bit Successive Approximation A/D Converter

In a successive approximation A/D, the conversion process begins when the most significant bit in the approximation register is set (while the other bits are reset). The analog output (from the D/A), resulting from this number is compared with the input signal. If the input voltage is greater than the voltage represented by the digital number, then the bit remains set to one, otherwise it is reset to zero. The second most significant bit is then set and the same test carried out. The process continues until all the bits in the approximation register have been tested, and thereafter, the register contains the digital equivalent of the analog input.

The conversion time in a successive approximation A/D is only dependent on the number of bits in the approximation register and is therefore constant for any particular device, regardless of the input voltage level. This is a major advantage over the integrating A/D and is one reason why successive approximation A/Ds are amongst the most widely used A/D devices.

It is clear that all A/D devices require a period of time in which to convert a signal from analog to digital representation. We already know that as a result of the limited number of bits used in digital representation that the conversion of data is accompanied by a loss of information, through quantisation error. There is however, another form of error introduced by the A/D conversion process, because of the time delay in converting from analog to digital. This is referred to as a "sampling error" because, in effect, when we use A/D devices, we are really only sampling an analog waveform.

Sampling is normally carried out at uniform time intervals (regardless of A/D type), but the sampling frequency is of crucial importance, as demonstrated in Figure 7.12.

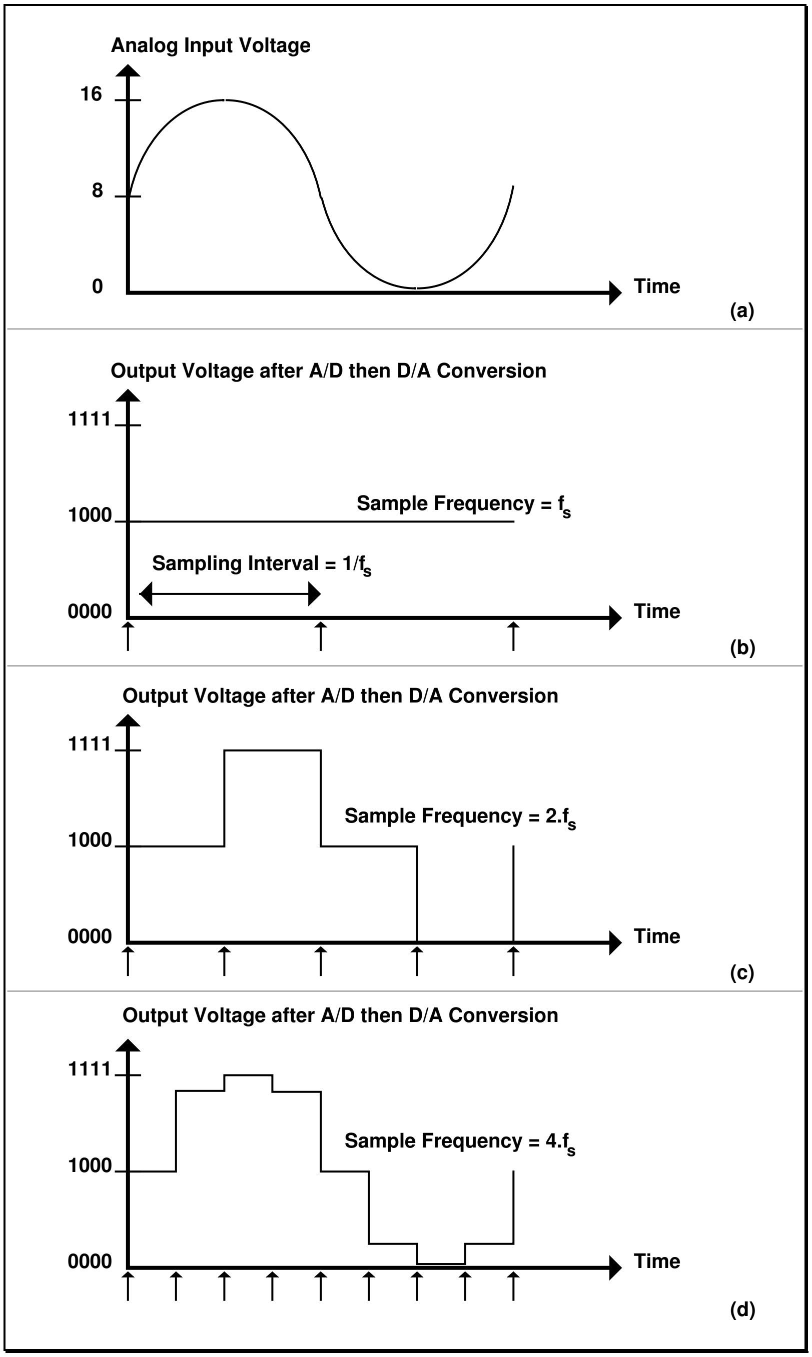


Figure 7.12 - Results after A/D then D/A Conversion after Sampling Waveform (a) at a Range of Frequencies - (b) f_s (c) $2 \cdot f_s$ (d) $4 \cdot f_s$

Figure 7.12 (a) shows a voltage waveform which is to be converted into a digital form by a 4-bit A/D converter and then converted back to analog form by a 4-bit D/A converter. Figures 7.12 (b), (c) and (d) show the results for a range of different sampling frequencies. The sampling frequency in Figure 7.12 (d) is twice that in Figure 7.12 (c) which is, in turn, twice that in Figure 7.12 (b). The higher the sampling frequency, the greater the accuracy of the approximation of the original waveform. Since A/D converters require a finite time for conversion, we must ensure that the conversion speed is sufficient to enable an adequate sampling rate to be achieved, otherwise, the results may be meaningless.

The so-called "Nyquist Sampling Theorem" establishes the basic rule-of-thumb for sampling of waveforms. The theorem is based on the fact that any waveform is the sum of a number of sinusoidal components of differing frequency and amplitude (the Fourier Theorem). The Nyquist Sampling Theorem states that the minimum sampling frequency must be at least twice that of the highest frequency component of the original waveform, in order for the waveform to be adequately represented in a digital form. In practice, sampling frequencies are selected to be three to four times the highest frequency component of the incoming analog waveform.

Another problem that arises in the A/D conversion process, as a result of the time taken to complete a conversion, is that the analog input waveform may vary during the conversion period. This means that even though a processor may initiate the analog to digital conversion cycle on an A/D chip at uniform time intervals, the digital outputs may not be uniformly spaced in time. This is particularly true for integrating A/D devices whose conversion times are dependent on the voltage magnitude. The solution to this problem is realised by a device that samples the analog input waveform, at uniform time intervals, and holds that voltage level until the next sample is taken. In other words, these devices sample the incoming waveform, and provide an output voltage equal to the input voltage, until triggered by an external signal (eg: from a processor). After triggering has occurred, the devices hold their output voltage at a constant level until they return to sampling mode. A/D conversion can then take place during the hold period. These devices are known as "sample and hold" devices and can either be designed individually or purchased as a commercial item.

The basic structure of the sample and hold device is shown in Figure 7.13. It is composed of a capacitor (for holding a captured voltage), a transistor switch (for switching between hold and sample modes) and two operational amplifiers (configured as voltage followers). The first operational amplifier provides a high-input/low-output impedance circuit to drive the switch. The second operational amplifier provides a high-input impedance to minimise the charge drain on the capacitor. When the sample line on the device is asserted, the output voltage approximately equals the input voltage, as does the voltage across the capacitor. When the sample line is reset (ie: hold is asserted) the capacitor is disconnected from the input and retains the previously attained voltage and charge, thus providing the "hold" function of the device.

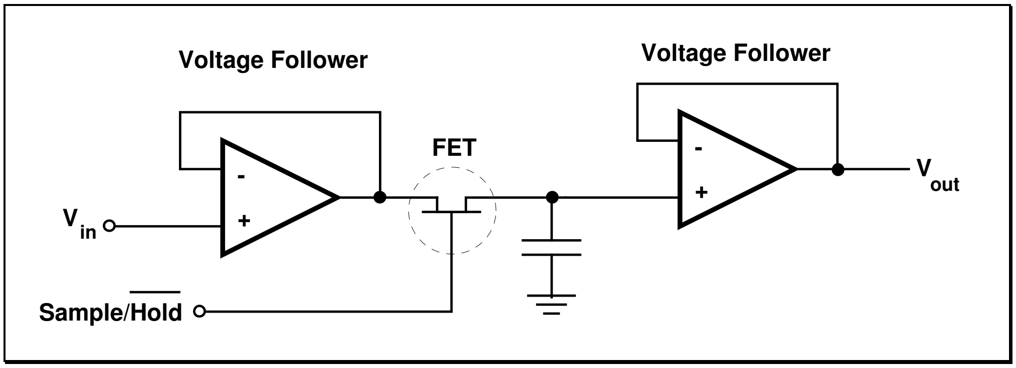


Figure 7.13 - Schematic of Sample and Hold Circuit

A number of factors determine the effectiveness of the sample and hold circuit. The primary one is the capacitor's ability to retain charge. The rate at which the capacitor discharges (thereby diminishing the stored voltage) is referred to as the droop rate of the circuit. The lower the droop rate, the better the holding characteristics of the circuit. The other major factor affecting the circuit is the switching speed of the transistor which ultimately determines how quickly a sample can be captured for holding.

In addition to the above-mentioned reasons for using sample and hold circuits, there are several common applications in which sample and hold circuits can be used to precede the A/D conversion phase. These are:

- When the input signal has a high frequency component that needs to be captured and digitised within a time-frame shorter than the conversion time of an A/D device.
- When the input signal is not continuous or continuously present. The latter situation may occur with transients or when there are several signals feeding into the A/D through a switch (multiplexer)
- When the cost of a high-speed A/D is unacceptably high for a given commercial application and a lower speed device can be substituted in tandem with a sample and hold circuit.

The last application is of interest to us because it highlights one of the major issues associated with A/D conversion - that is, component cost. A/D devices are relatively expensive in digital circuit terms and their cost is dependent upon conversion speed, resolution, etc. Although the cost of such devices is gradually decreasing, there may be instances in which it is necessary to share one A/D amongst several incoming signals. In digital circuit terms, the switching circuit that performs this function is called a multiplexer and is a commonly available IC component.

Figure 7.14 shows an application with eight incoming analog signals which share one A/D converter via sample and hold circuits and an "8 to 1" multiplexer.

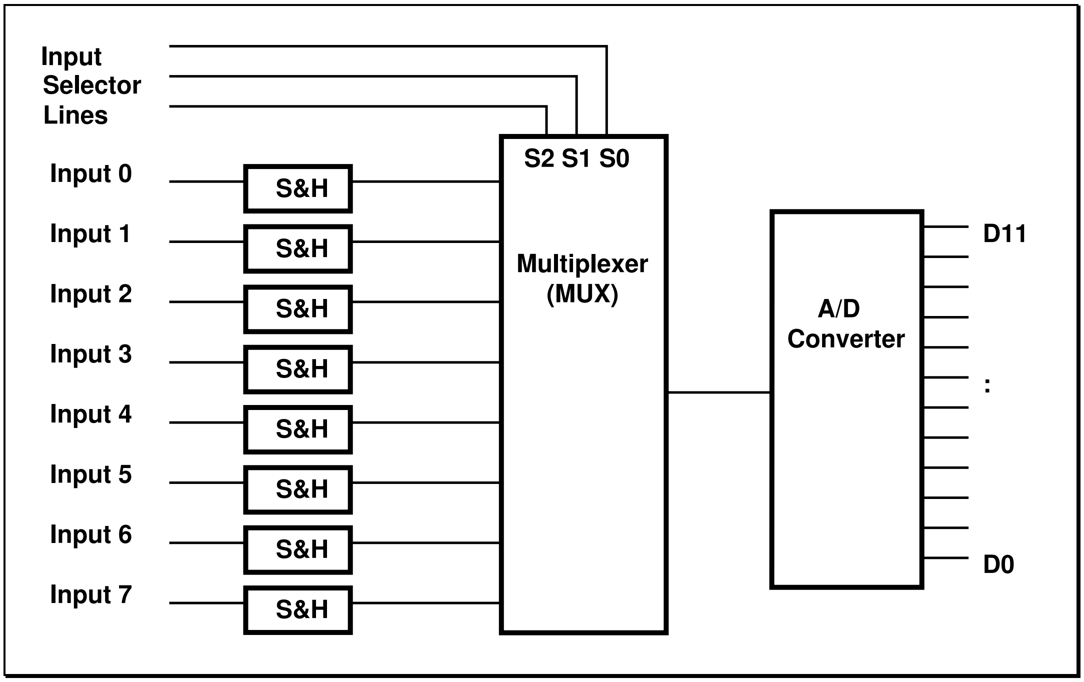


Figure 7.14 - Using Sample and Hold Circuits with a Multiplexer to Share an A/D Converter Between Eight Analog Input Lines

The operation of the multiplexer is relatively easy to understand. A number of digital input-selector lines are used to connect the required input to the output of the circuit. For example, in the system shown in Figure 7.14, setting S2, S1 and S0 to one, zero and one, respectively, would connect analog input line 5 to the A/D converter and so on.

The concept of sharing an A/D converter between a number of inputs really only has marginal benefits and normally one has to determine whether it is better to use one high performance A/D or several lower performance A/Ds to accomplish the task. If conversion speed is an issue, and one A/D is handling many input lines then its sampling performance per input line is reduced accordingly. However, the major benefit of sharing an A/D is when the device has a very high output resolution which cannot be readily achieved by several low-resolution chips in parallel.

7.4 Signal Conditioning, Protection and Isolation

7.4.1 Introduction

The concepts behind signal conditioning, protection and isolation are almost as broad as the entire field of Electrical Engineering and it is not practical to provide a detailed coverage of all the points involved. The purpose of the sections that follow is to overview some basic devices that can be considered for the various tasks. However, before we do so, it is important to clarify what we mean by the terms encompassed in this section.

The basic problem that we need to address really relates back to Figure 7.2, which is our global diagram for interfacing. The general case of interfacing signals from the real world to digital circuits really requires that we take analog signals in various energy forms and convert them to digital voltages of a size compatible with the inputs to the computer (or more specifically, programmable parallel interface). In section 7.3, we looked at the process of converting analog voltages into a digital form. In section 7.5 we will look at converting signals from other energy forms into a voltage representation. However, in this section we examine the process of taking voltages and converting them into an appropriate size and shape that is compatible for digital circuits, and conversely, taking outputs from digital circuits and converting them to a form suitable for driving external systems. We also consider the need to shelter digital circuits, and the users of those circuits, from dangerously high voltage and current levels that may exist in external systems. The following definitions are used in this book:

- ***Signal Conditioning:***
  Conversion of voltages from their raw form (either analog or digital) to a form suitable for use in another (analog or digital) circuit

- ***Protection:***
  The design of circuits that can be used to prevent spuriously high signal levels from damaging other circuits or human operators

- ***Isolation:***
  The design of circuits, based on devices that transfer electrical energy via some non-electrical intermediary form (magnetic circuits, optical circuits) and facilitate energy transfer without direct electrical connection.

7.4.2 Signal Conditioning Circuits

Signal conditioning circuits can be broadly classified into several groups, based upon the type of incoming signals and required outgoing signals. These include:

(i) Waveform Correction Circuits:
  - Schmitt Trigger
  - Debouncing Circuits

(ii) Scaling Circuits:
  - Analog Amplifiers
  - Digital (PWM) Amplifiers
  - Transformers

(iii) Filtering Circuits:
  - Analog (R-L-C circuits)
  - Digital (Switched-Capacitance Filters)

We shall examine each of these circuits briefly, in turn.

(i) Waveform Correction Circuits

Practical digital circuits operate at a very high speed and require inputs that are close to the "ideal" rectangular waveform shape in order to function correctly. Many incoming waveforms, which are thought to be digital, may not in fact change at a sufficiently high speed to be considered thus by a TTL or MOS digital circuit. Sometimes, it is also desirable to have an incoming analog signal converted to a digital signal by some "simple" circuit. The circuit that achieves these ends is known as a "Schmitt Trigger" and is a commonly available component. The circuit symbol for the Schmitt Trigger is a triangle with a hysteresis loop drawn within. This is shown in Figure 7.15 (a). The operation of the circuit is shown schematically in Figures 7.15 (b) and (c).

The circuit works with two distinct threshold or "trip" levels, known as the upper trip point (UTP) and the lower trip point (LTP). When the input voltage exceeds the UTP, the output of the trigger sharply rises, as shown in Figures 7.15 (b) and (c). Similarly, when the input waveform drops below the LTP, the output of the trigger drops sharply.

The fact that the LTP is lower than the UTP means that the Schmitt Trigger can also be used to improve incoming digital signals that contain some noise. Noise can sometimes vary a digital signal below or above the required digital logic level, thereby creating an erroneous bit. However, when the Schmitt Trigger is used to improve the signal, any noise induced during a "high" period in the waveform would have to reduce the signal below the LTP before any change in output would occur. Similarly, any noise induced during a low period in the waveform would have to raise the signal above the UTP before a change in output would occur.

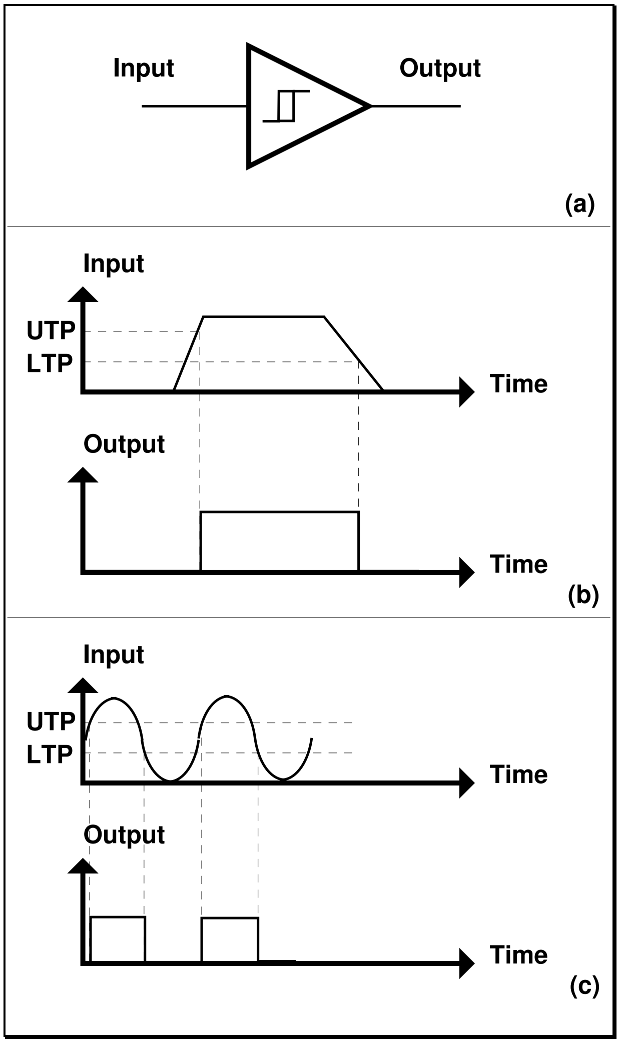


Figure 7.15 - (a) The Schmitt Trigger Circuit Symbol
(a) "Squaring Up" Slowing Changing Input Signals
(c) Producing Digital Signals from an Offset Sinusoidal Input Waveform

The Schmitt Trigger circuit can improve circuits with a small amount of noise but there are some inputs that contain so much noise that they need to be processed by other circuits before being fed into a digital circuit. The most common problem is with mechanical switches that all suffer from a property known as "bounce". Bounce, as its name implies, means that switches do not open and close (break and make) cleanly, but rather oscillate from one state to another for a few milliseconds before settling. Although a human user would not notice this short time-frame phenomenon by observing an ohm-meter, the results of switch bounce can lead to incorrect data flowing into a digital circuit. As a result, all switches need to be "debounced" before information from them can be processed.

Debouncing can be achieved by simply waiting several milliseconds before reading a switch state or in simpler digital circuits, through a simple "debouncing circuit". A debouncing circuit can be achieved via a Schmitt Trigger based system with an RC network or, preferably,. through an R-S Flip-Flop arrangement, made from two NAND gates, such as the one shown in Figure 7.16.

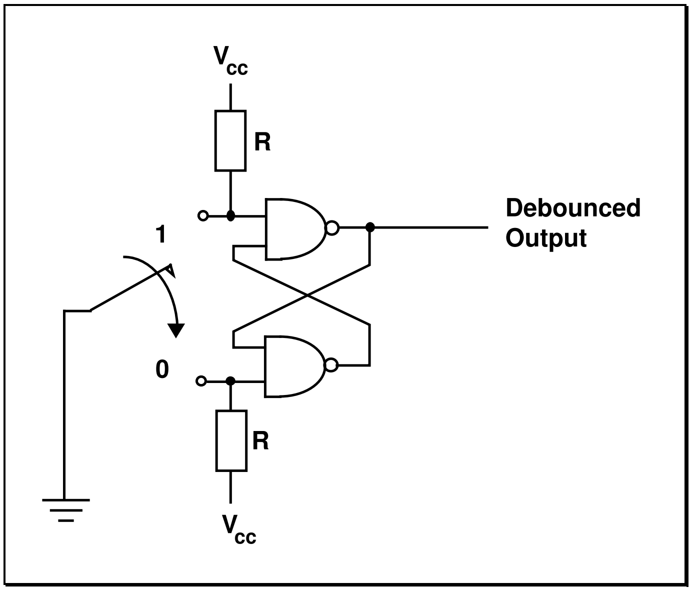


Figure 7.16 - Switch Debouncing Circuits Based on R-S Flip-Flops

(ii) Scaling Circuits

One of the most common requirements in interfacing devices is the scaling of voltages. There are two techniques by which voltages can be scaled. These are:

- Amplification
- Transformation.

Amplification can be either analog (linear transistor circuits such as operational amplifiers) or digital (PWM switching of transistors). Both analog and digital amplification can be used on time-invariant or time-variant, alternating and direct current signals and are the most versatile methods for scaling. The simplest techniques are realised through the use of operational amplifier circuits, which have been discussed in detail in 3.3 and 3.5, but in recent years, the more energy-efficient (and complex), PWM digital techniques have been applied for power circuits. The basic concept of analog and digital amplification is shown in Figure 7.17.

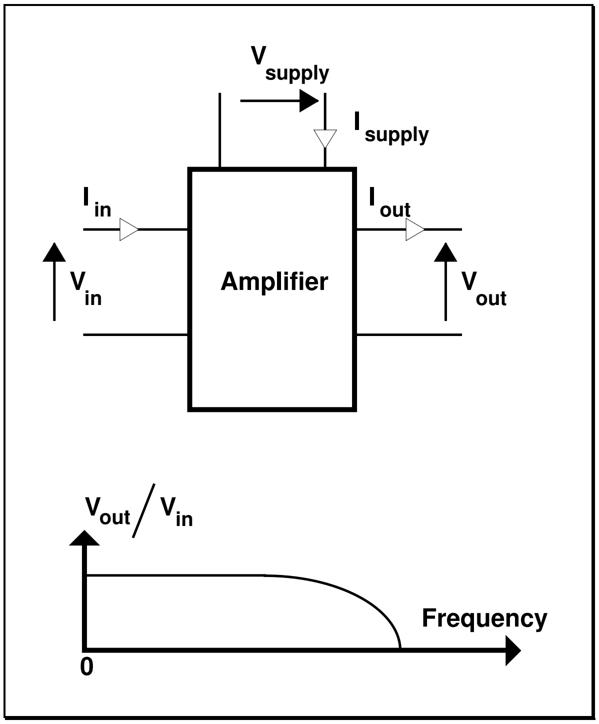


Figure 7.17 - The Concept of Amplification

Looking at Figure 7.17, and applying the concept of conservation of energy, we know that:

$$v_{in} \cdot i_{in} + v_{supply} \cdot i_{supply} = v_{out} \cdot i_{out} + P_{amp} \qquad ...(3)$$

where:

P_{amp} is the power converted to heat in the amplifier.

The important point to note about equation (3) is that the output signal power of the amplifier ($v_{out} \cdot i_{out}$) can be larger than the input signal power ($v_{in} \cdot i_{in}$) because additional power is provided by the amplifier supply rails. This means that the amplifier can be used to convert small signals into larger signals.

We also know that the output voltage and current from an amplifier are a scaled version of the input voltage and current, but that the relationship between input and output is also frequency dependent as a result of the physical attributes of the amplifier, as shown in Figure 7.17 and discussed in detail in 3.4 for the analog amplifier.

The transformer is the other obvious device that we can use for scaling of analog signals. The operation of the transformer is shown conceptually in Figure 7.18

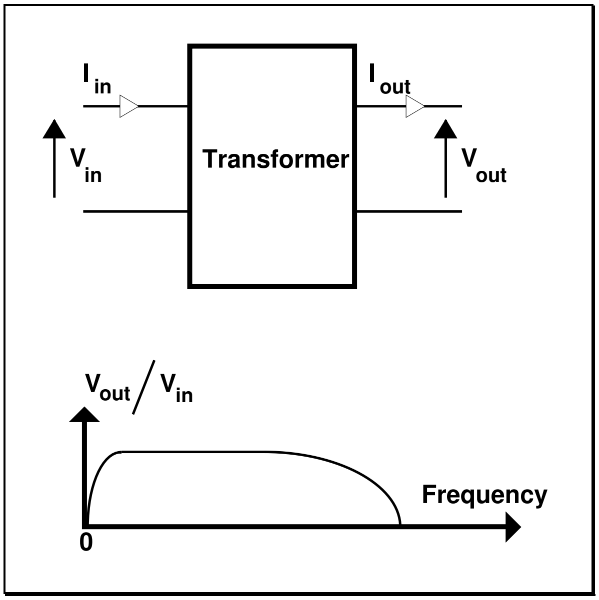


Figure 7.18 - The Transformer in Concept

Looking at Figure 7.18, we can deduce the following relationships between input and output power in a transformer:

$$v_{in} \cdot i_{in} = v_{out} \cdot i_{out} + P_{trans} \qquad ...(4)$$

where:

P_{trans} is the power consumed by the transformer through hysteresis, eddy-current and copper losses.

Equation (4) highlights the major difference between the transformer and the amplifier. The transformer is a passive device and if we use it to scale up voltage, then the corresponding output current decreases and vice versa. The other drawback of the transformer is that its frequency response is somewhat limited. Firstly, no transformation occurs at zero frequency and secondly, the upper frequency limits of common transformers are typically much lower than those of common amplifiers. The other limitation of the transformer is its physically large size. However, the transformer has the major advantage of providing complete electrical isolation between the input and output.

While analog amplification and transformation are both suitable for scaling inputs to digital circuits, the problem of converting the small signals produced by digital circuits into large voltages and currents for driving high powered systems presents a number of additional problems.

Analog transistor amplifiers can be used at high power levels and a number of high-powered operational amplifiers can also be purchased. The problem with all analog amplifier circuits is that they are really using transistors as a variable resistance that controls energy flow from the supply rails to external circuits. The bulk of the difference between the signal and supply energy input and the signal energy output is dissipated in the variable resistance provided by the transistors (when in linear operation). This means that a great deal of energy is converted into unwanted heat in the analog amplification process. The energy wastage is irrelevant in small circuits but can be of significance in large systems. The heat generated by the energy wastage is relevant in all circuits because it has to be compensated for by cooling mechanisms such as fins or electric fans.

A more modern, digital form of signal amplification that is used to provide a variable output voltage is called pulse-width modulation or PWM. PWM circuits are available as commercial ICs and are really sequential circuits that can generate rectangular output waveforms with varying duty cycles (ie: varying "on" to "off" ratios). The outputs are then used to switch power transistors on and off.

The average value of the PWM output can thus be varied between zero and the supply rail voltage. This is shown in Figure 7.19, which illustrates the average output voltage for a range of PWM output waveforms.

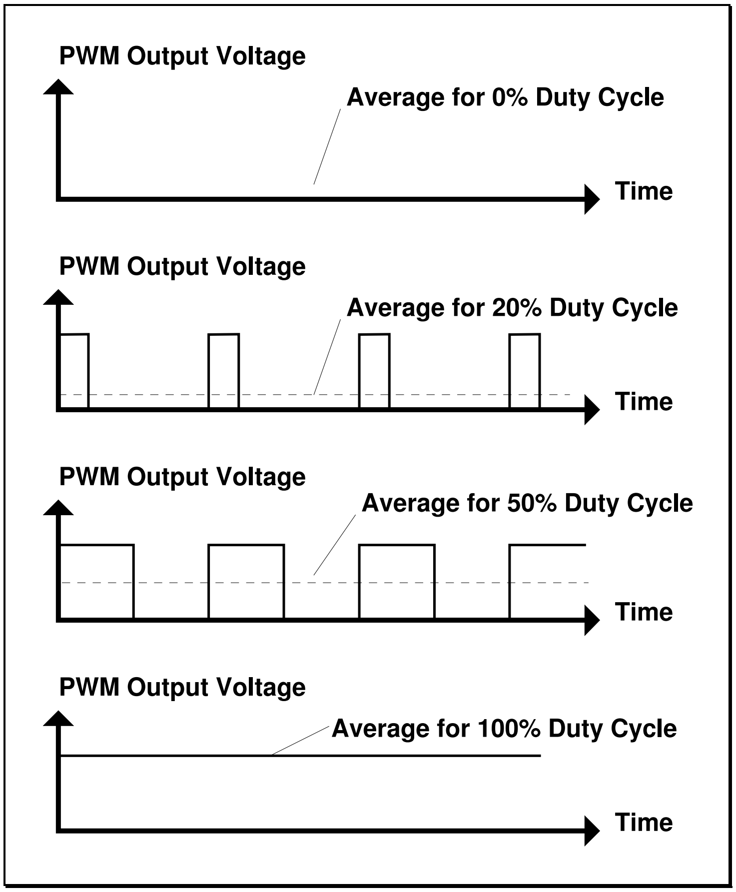


Figure 7.19 - PWM Output for a Range of Different Duty-Cycles

Although we normally talk about varying the average voltage with PWM circuits, the reality is that the output is always a rectangular waveform, whose average value can be periodically varied. It is not a smooth output. Many systems are however tolerant of this type of digital waveform, including d.c. motors, where the technique is used. This is because the motors have an inductance whose natural tendency is to smooth (choke) the waveform. In other applications, smoothing filters may need to be added to make the outputs more suitable.

PWM circuits are only low power digital circuits and cannot be directly used to drive any high power circuits. Instead, the output of a PWM circuit is normally used to drive the base or gate of high power transistors which, in turn, switch between high and low supply-rail voltages and can thereby provide substantial power outputs. The net effect is the creation of a digital amplifier circuit.

The benefits of PWM based amplification over analog amplification can be substantial. Firstly, the transistors in a PWM based amplifier do not function in their linear region as variable resistors. They are only used in their "on" and "off" modes. This means that the power dissipation in the amplifier is substantially reduced and that power amplifier circuits based on PWM techniques can be much smaller than those based on analog circuits. The disadvantage is that we are always dealing with rectangular waveforms and this means that we sometimes need to develop more sophisticated control systems when we wish to drive systems based on, for example, sinusoidal voltages as in the case of three-phase motors. PWM circuits also have a limited bandwidth in which they can operate. This is governed by the switching speed of the transistors within the PWM and, more importantly the switching speed of the power transistors in the amplification section of PWM based drives.

The most common application of PWM based amplification is in switch-mode power supplies, commonly found in computers. The typical arrangement is shown in Figure 7.20.

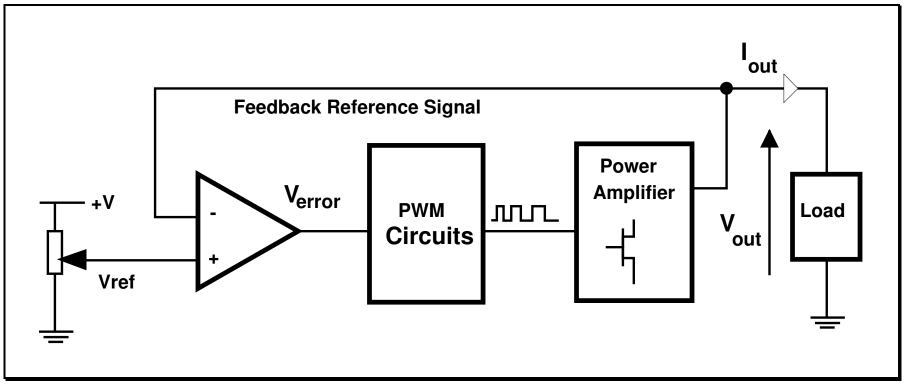


Figure 7.20 - Switch Mode Power Supplies Based Upon Amplification of PWM Signals

(iii) Filtering Circuits

Fourier analysis tells us that any waveform can be represented as the sum of a number of sinusoidal components of differing frequency and amplitude. Although we normally deal with systems in the time domain, there are instances where we also need to examine the frequency domain of signals. Mathematically, the conversion between the time-domain representation of a signal, x(t), and its frequency domain (spectrum) representation, X(f), are defined by the dual Fourier transforms:

$$X(f) = \int_{-\infty}^{+\infty} x(t) \cdot e^{-j2\pi ft} dt \qquad ...(5)$$

and

$$x(t) = \int_{-\infty}^{+\infty} X(f) \cdot e^{j2\pi ft} df \qquad ...(6)$$

Equation (5) gives us the frequency spectrum of any time domain waveform, which, in general, is composed of a number of wanted components and also unwanted components that need to be removed. The process of removing unwanted components at various frequencies is called filtering. The spectrum of an incoming signal is shown in Figure 7.21, including wanted and unwanted components. The unwanted components can be filtered out by using a range of different filters, generically known as:

- Low-Pass (which only pass frequencies from zero to an upper limit)
- Band-Pass (which only pass frequencies between two frequency limits)
- Notch (which pass frequencies between zero and an upper limit but not between two intermediary points).

The design of filter circuits is a specialised field in its own right and is a complex process, which will not be detailed in this book, particularly since there are many classic text books in the field. However, suffice to say that there are three techniques for filtering waveforms that have been obtained from some external system. These are:

- Develop an analog circuit, composed of resistors, capacitors and inductors, that has the appropriate frequency domain characteristic to achieve the desired result. There are many text books written on how this can be achieved

- Develop a digital filter circuit using modern filtering techniques such as switched-capacitance filters (which are available as standard components). This is relatively easy to do but has performance limitations.

- Use an algorithmic approach, where the unfiltered signal is fed into a computer system, and an algorithm developed to obtain the frequency spectrum, and then only use the wanted components.

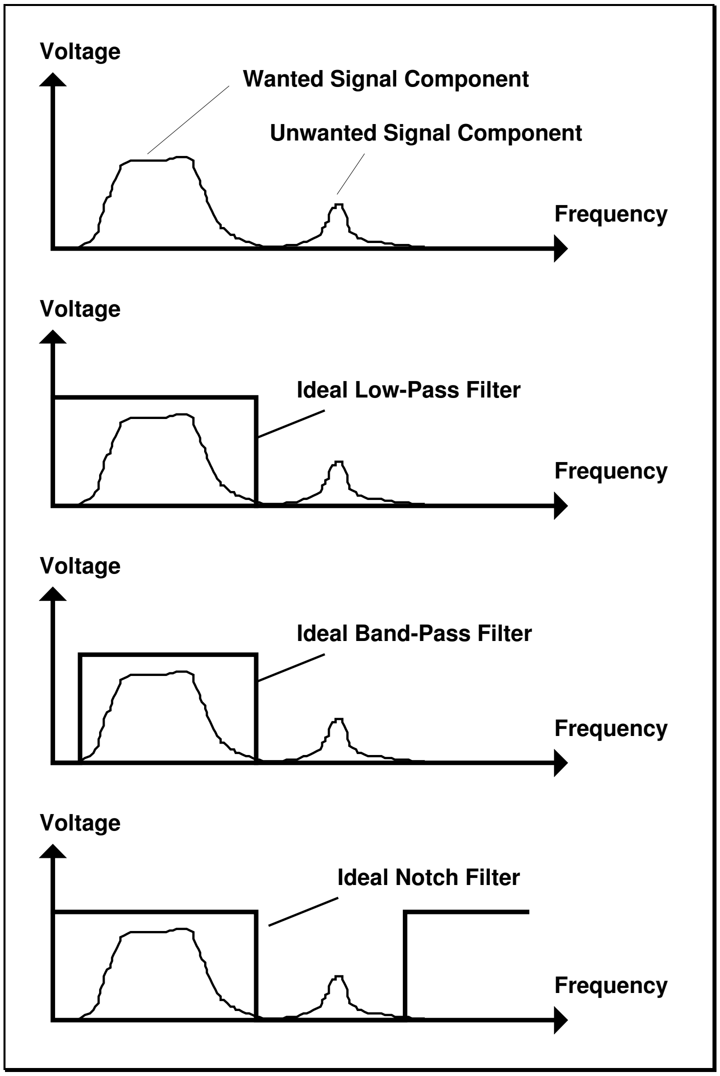


Figure 7.21 - Filtering Out Unwanted Signals with Low-Pass, Band-Pass and Notch Filters

Of the three possible approaches to filtering, the second is perhaps the most practical because it can be based upon a commercial solution that only needs minor tailoring in order to get the desired effect. The other two approaches require a sound understanding of network theory and Fourier analysis techniques and are not really suitable for novices in the field. Moreover, even if one does understand the basic theory of analog filter design, fabrication can be a problem if elements such as inductors need to be included. Inductors can take up a significant amount of space and, in any case, normally need to be specially wound for filter applications.

As a final point, it is interesting to note that all electrical circuits exhibit filtering characteristics because they have resistance, capacitance and inductance, even when these are parasitic entities and not actual (lumped) components in the circuit. The most common characteristic is therefore the low-pass characteristic, which occurs with all forms of amplifier and was typified in Figure 7.17. The transformer is a circuit element which exhibits a band-pass characteristic, because it will pass neither d.c. nor high frequency signals.

7.4.3 Protection Circuits

Industrial systems, composed of mechanical and electronic devices can be very hostile to the circuits within a computer system are concerned. Digital circuits in a computer system are designed to work at low (digital) voltages and can only source or drain very small currents (in the order of milli-Amps). The outside world on the other hand is filled with large analog voltages and currents which could very easily damage the internal circuitry of a computer system. The signals fed back to a computer from an external system are also unpredictable for a number of reasons:

- The external equipment may be susceptible to large voltage or current spikes (for example, an induction motor current is abnormally large when the motor is started direct on line)

- The external equipment may accidentally be connected to other high voltage systems, as a result of cables on moving equipment accidentally disconnecting themselves

- Terminals on external equipment may be short-circuited as a result of human error or conducting materials falling across terminals.

As far as feedback signals into the computer system are concerned, simply scaling the levels to a suitable size and conditioning them into an appropriate shape may not be adequate for protecting the computer circuits from damage. There are situations in which external devices could be subjected to unexpected energy surges which ultimately lead to voltage spikes at the input of the computer system. If this is a possibility then there are some simple protection measures that can be employed.

There are several devices available for protecting digital circuits from high voltages and currents by disconnecting them from the source. They include:

- Zener Diodes
- Thyristors
- Relays.

The first two devices have been covered in detail in Chapter 3 and so we will now examine the operation of the relay, which is a relatively simple device to understand.

Essentially, the relay is a spring-loaded switch that is pulled open or closed by a force generated by a magnetic field which is, in turn, generated by the flow of current in a coil. Relays come in an enormous range of sizes and can be triggered by coil currents as low as micro-amps or as high as amps. Relays also come in two different configurations, referred to as "normally-open" and "normally-closed". A normally-open relay is one where the switch is in its open-circuit position until current flows in the coil. A normally-closed relay is one where the switch is in its short-circuit position until current flows in the coil. The relay is shown schematically in Figure 7.22.

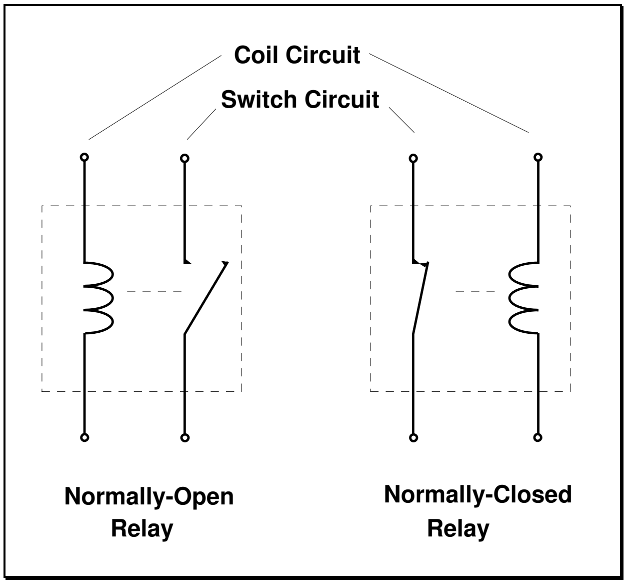


Figure 7.22 - Relay Configurations

Relays can operate relatively quickly in mechanical terms, typically changing state in the order of milli-seconds. This switching speed is generally fast enough to prevent injury to human users but may not be fast enough to protect delicate circuits. In these instances, it is necessary to revert to semiconductor based switches that can provide response times in the order of nanoseconds or microseconds rather than milli-seconds.

In Chapter 3, it was noted that Zener diodes can be used to regulate voltage levels. If the breakdown voltages on these diodes are judiciously chosen, then these diodes will restrict voltage inputs to computer circuits to a maximum acceptable level, while still allowing normal signals to pass through. Back to back Zener diodes enable protection in a.c. circuits. The configuration is shown in Figure 7.23, where an additional safety feature, a (normally-closed) relay, is incorporated into the Zener diode branch of the circuit. The relay is "slow-acting" in comparison to the Zener diode but, when a current flows through the Zener diode branch for a sufficient period, the relay will isolate the offending current source in order to minimise the potential for damage.

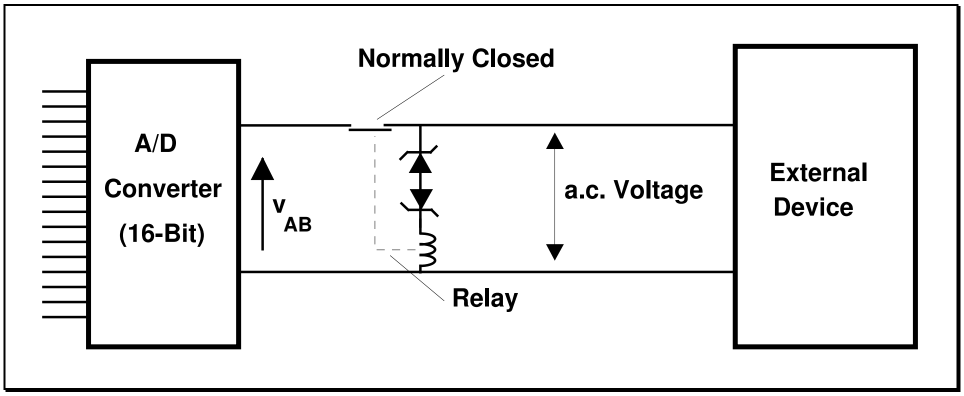


Figure 7.23 - Protection Using Zener Diodes and Relays

The semiconductor alternative to the sort of relay circuit shown in Figure 7.23 is the Silicon Controlled Rectifier (SCR) based circuit known as the crowbar. The crowbar circuit has already been discussed in 3.7.2 and is shown in Figure 3.53.

High currents in mechatronic circuits can arise from short-circuit conditions or changing mechanical conditions such as excessive loads on motors. Most over-current conditions can be handled through the use of relays that are triggered once currents exceed a certain level. Sometimes, the current flowing through a particular branch of a circuit can be too high for a given relay. In those instances, a scaled version of the current may need to be passed through the relay coil and this is normally achieved by using a transformer with a suitable turns ratio.

So far we have only looked at protecting the inputs of low power circuits. However, there may be instances where we need to protect the outputs. This typically occurs when we need to drive some external device from a low-power circuit. The common solution is to use transistor based amplification, but sometimes simpler, relay solutions can achieve the same result. Consider the situation where the programmable parallel interface (PPI) from a computer system needs to perform a function such as turning on a lamp that draws several Amperes of current. A simple circuit, composed of an external power supply that is turned on by a relay with a triggering current of a few micro-Amperes of current can be used to drive the lamp as shown in Figure 7.24. The triggering current for the relay can be set up by inserting an appropriate current-limiting resistor R_L in series with the coil. The current through the external device can be adjusted by designing an appropriate power supply or by varying an appropriate series resistance.

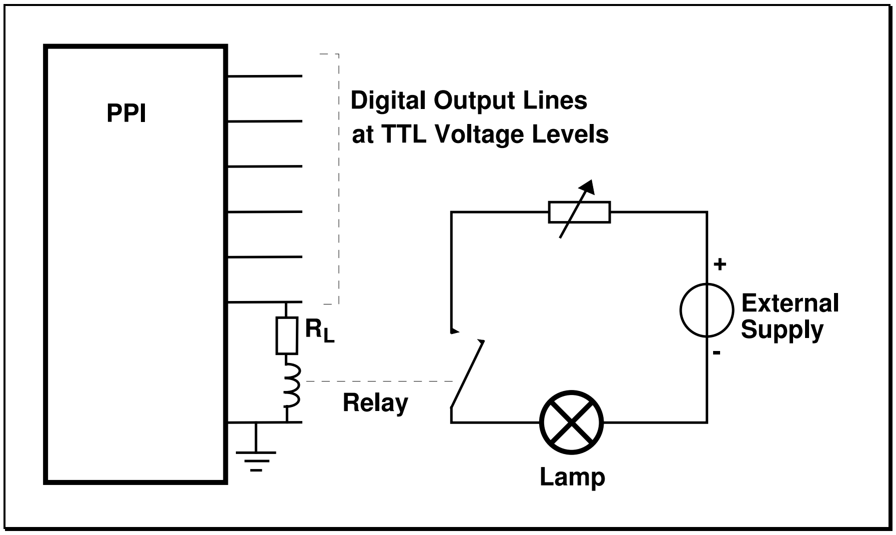


Figure 7.24 - Driving Simple High Current Circuits from Digital Outputs by Using Relays

The circuit of Figure 7.24 is really a crude electro-mechanical version of an amplifier and can also be used in situations where a low current output is incapable of driving a large fan-out of other circuits. Providing that the relay switching speed is adequate, this simple type of circuit protects the low-power digital circuit from damaging itself by providing an excessively high output current.

7.4.4 Isolation Circuits

A common problem with extracting voltages from external systems is that voltages are relative quantities. For example, if we had two terminals on an external device (A and B) and we wanted to feed the voltage across those terminals (V_{AB}) back to the computer through an A/D converter, then we may need to provide special isolation circuits, even if V_{AB} is relatively small. The problem is clearly illustrated by the example in Figure 7.25.

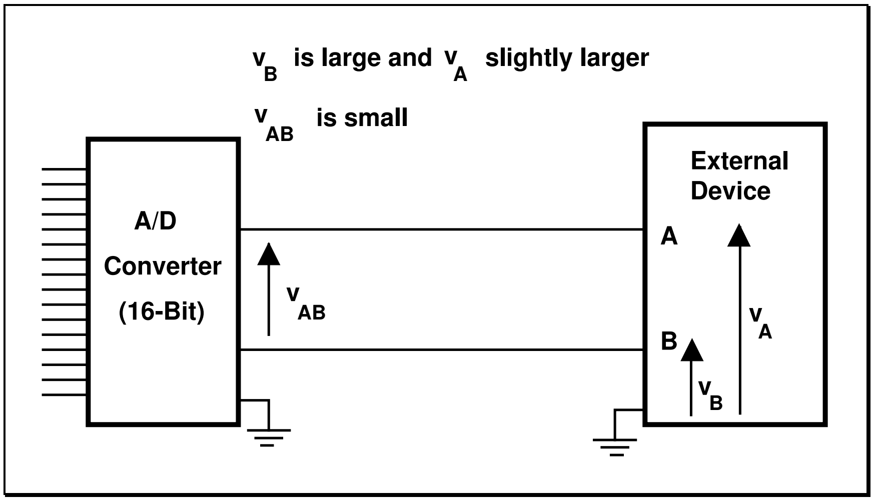


Figure 7.25 - The Problem of Measuring Small Voltage Differences Between Two Voltages Which are Both High with Respect to Earth

In Figure 7.25, it is evident that if V_B is large (say 500 volts) and V_A is larger (501 volts) then even though the difference between the two is small, the value of either, with respect to earth is high. In these situations it is not the voltage V_{AB} which is of concern. It is the fact that the voltage on either of the two lines (with respect to the ground pin that will exist on the A/D) will be very high. In other words two pins, separated by a few millimetres, may have a very large voltage across them. This is clearly dangerous and highlights the need for isolation.

It would be very convenient if terminal B was simply floating on the external device and we could simply ground it without effecting the voltage across A and B. However, in general we do not have this luxury and grounding a high voltage can be extremely dangerous. For this reason, we need to use devices that can isolate the computer circuitry from the high voltages and still provide the required signal. If the voltages are low-frequency a.c., then a transformer inserted between the two devices will perform the necessary task, by allowing us to ground the side closest to the A/D device. This is shown in Figure 7.26.

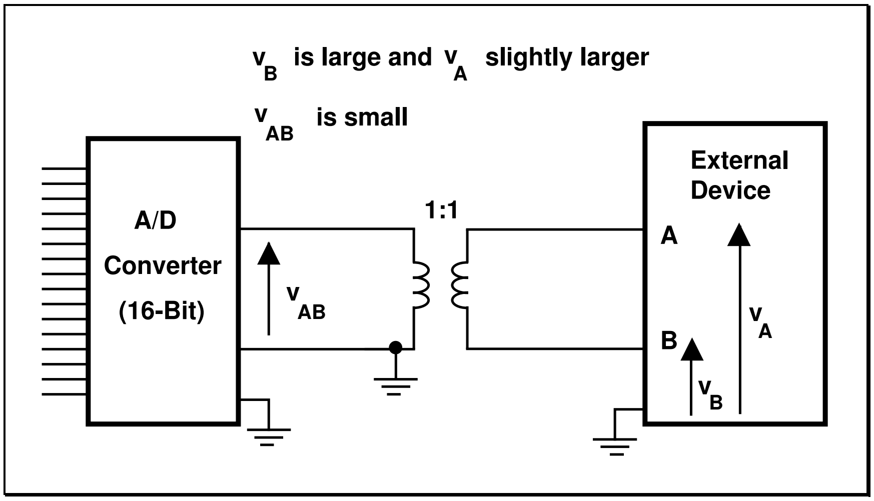


Figure 7.26 - Isolation Using Transformers

A similar problem exists when we need to measure the current passing through a conductor. Theoretically, we could simply insert a small resistance into the line and measure the voltage across it - however, that would simply create the same problem as in Figure 7.25. A good solution for a.c. waveforms is to use a current transformer (which is simply a voltage transformer with a small resistance across one side). This is shown in Figure 7.27 where, as an example, we wish to measure the current through a motor circuit. The current in the motor circuit is transformed to the primary side, and since the A/D has a very high input resistance, the current predominantly passes through the small resistor, R. The voltage across the resistor is then proportional to the current in the motor. Moreover, the apparent resistance in the motor circuit is reduced by the square of the turns ratio, thereby minimising the effects of the current monitoring process.

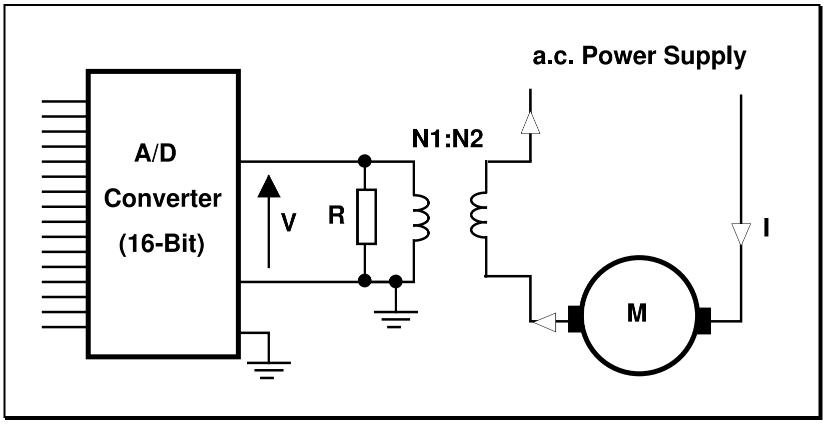


Figure 7.27 - Using a Current Transformer to Isolate and Monitor Current

Transformers are convenient devices to use if space permits because they not only provide isolation but also scaling of voltage levels through their turns-ratio. However, transformers do not work with time-invariant. signals and generally attenuate high frequency signals so they clearly have a limited application range. Other devices such as opto-couplers (opto-isolators) provide another mechanism for isolation that can be used for both a.c. and d.c. signals. The operation of the opto-coupler is shown schematically in Figure 7.28 (a).

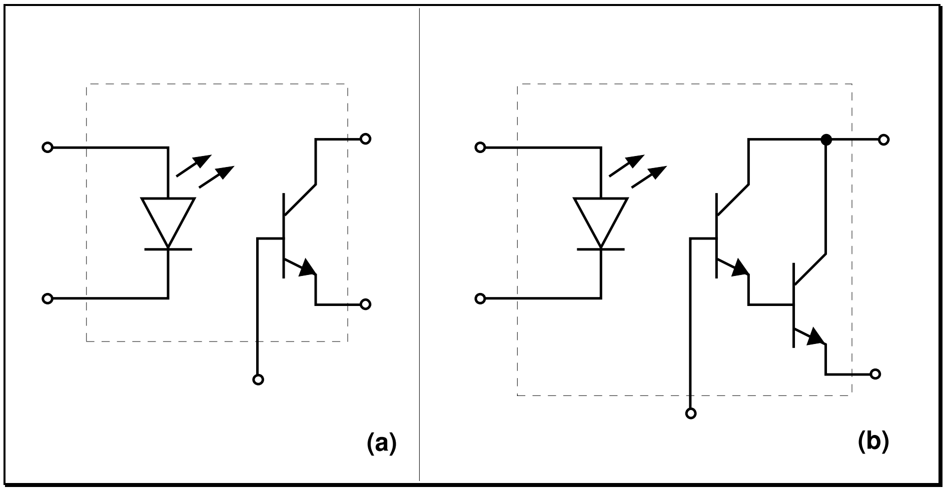


Figure 7.28 - Opto-Couplers (Isolators)
(a) Simple Opto-Isolator
(b) Darlington-Pair High Gain Opto-Isolator

The simple opto-isolator is composed of a Light-Emitting-Diode (LED) and a light-sensitive phototransistor. The output of the transistor is dependent upon the intensity of light emitted from the LED which is, in turn, dependent upon the amount of current flowing through it. The transfer gain can be improved significantly (by a factor of 15 or more) by using a Darlington-Pair transistor configuration as in Figure 7.28 (b).

Opto-isolators sometimes suffer from linearity problems and so it is necessary to ensure that non-linearity is compensated for by processing algorithms. However, the opto-isolators do fulfil an important role in interfacing because they are physically smaller than transformers and are capable of isolating both a.c. and d.c. signals.

7.5 Energy Conversion - Transducers

The purpose of a transducer is to convert one energy form into another. In terms of interfacing industrial systems to digital circuits (and computers), the purpose of transducers is to convert some raw energy form, such as pressure, temperature, etc. into voltage. Depending on the type of transducer, the output can either be analog (which is still the more common form) or digital. The voltage then has to be scaled and, if analog, converted into a digital form for interaction with other digital circuits.

We could also extend the idea of transducers and say that an electric motor, relay or solenoid are transducers because they convert electrical energy into some form of mechanical energy, normally to provide movement. In these instances, the transducers are more specifically "actuators". These devices are generally electromagnetic in nature and are vitally important in the creation of mechatronic systems. Electromagnetic actuator devices will be covered in some detail in Chapter 9.

Limiting our discussions here primarily to input transducers, we can say that there are an enormous number of commercially available devices to meet the requirements of modern interfacing tasks and the range is continually expanding. It is well nigh impossible to cover the range of available devices and there is little point in doing so, since their characteristics are ultimately dependent on the specific devices in question. In this section therefore, we will only review a few of the basic transducers.

There are a number of characteristics that need to be ascertained before transducers can be selected for an application. These include:

(i) Threshold
(ii) Resolution
(iii) Input Range
(iv) Linearity
(v) Frequency Response
(vi) Monotonicity
(vii) Hysteresis
(viii) Repeatability
(ix) Slew-Rate
(x) Stability.
.

Characteristics (i) to (iii) are the ones which most people would intuitively check. The input threshold defines the minimum input energy that the transducer will detect. The resolution defines the minimum change in input energy that will be reflected in the output of the device and the input range determines the minimum and maximum input energy levels for the transducer.

Characteristics (iv) to (vi) really define the real-time performance of the transducer. Linearity in transducers signifies that the output energy is linearly related to the input energy and simplifies processing of data. However, not all transducers are linear and many have a limited range of linearity, after which they tend to saturate. As with any other transformation device, the frequency response is important. Most mechanical and electronic systems tend to act as low-pass filters that attenuate high-frequency components. This is a reflection of the difficulty involved in changing any physical energy at a very high rate. However, when selecting transducers it is necessary to ensure that they can at least cope with the range of input frequencies expected. Sometimes it is necessary to measure or estimate the highest frequency input before selecting a transducer. Monotonicity is a characteristic that can have significance in control applications. In a monotonic device, an increasing input always leads to an increasing or static output. Similarly, a decreasing input always leads to a decreasing or static output. Devices which are not monotonic are difficult to deal with in a control sense.

Characteristics (vii) to (x) relate to both the short and long-term accuracy of the transducer. Hysteresis, which leads to differences between forward and reverse characteristics of many physical devices is also common in transducers. This can lead to two different outputs for the same input, depending on whether the input is on the rise or fall when the output is recorded. Repeatability is really a tolerance or accuracy factor. It defines the maximum error between successive output values for a given input value. Slew-rate is a performance factor and some might argue that it is related to the frequency response. Slew-rate is the maximum rate of rise of output and so defines how quickly the transducer can respond to inputs such as step changes. Stability relates to the ability of the transducer to maintain its characteristics and accuracy over a period of time. Most electronic components have characteristics that vary with age - doping levels in semiconductors change, resistance characteristics change and so on. Mechanical components wear with use and so the working lifespan of a transducer needs to be considered during the selection process.

There are too many transducers to enable a complete coverage of this subject and so, what follows is an overview of some of the more common devices:

(a) Switches

Few would consider the simple mechanical switch to be a transducer and yet, upon reflection, it is evident that this device converts mechanical energy into electrical voltages. More specifically, switches in all their various forms (including key-pads and keyboards) are the fundamental interface between digital circuits and human users. Basic mechanical switch operation is self-evident and the problems in converting mechanical switch movements into sensible voltages (by debouncing) have already been covered in 7.4.2.

(b) Light Emitting Diodes

Light Emitting Diodes or LEDs are another family of simple devices that many would overlook when discussing transducers. However, they are an invaluable part of many control and monitoring systems because they provide a very simple mechanism for output of information to human users. In essence, the purpose of the LED is to convert electrical energy into light in the visible spectrum. LEDs are also used in conjunction with phototransistors in order to generate opto-couplers as described in 7.4.4.

LEDs are available in a range of different colors, including red, green yellow, orange and white. The common circuit symbol for all LEDs is the same, regardless of output color and is the simple diode symbol with two additional arrows, as shown in Figure 7.28. The characteristics of the diode are essentially similar to those of the normal p-n junction diode, discussed in Chapter 3, with the only major differences being that the LED's forward breakdown voltage is typically higher and reverse breakdown voltage lower than that of a normal diode.

LED's are extremely useful because they can be driven by most common digital circuits, normally with an open-collector gate, that can provide the necessary forward current to cause illumination.

(c) Potentiometers

Potentiometers are another family of simple transducers that are commonly used in interfacing and control circuits. A potentiometer is really nothing more than a variable resistor, whose resistance changes with the movement of a central arm known as a wiper. The potentiometer therefore translates rotational position into a variable voltage. Potentiometers are used at the human-electronic interface, where they normally appear as control knobs, and also in older servo motor control systems, where they were coupled to the shaft of the motor to indicate the rotational orientation (position) of the motor. The potentiometer is shown schematically in Figure 7.29.

A fixed d.c. or a.c. voltage can be applied across the two ends of the potentiometer and the output voltage taken between any one end and the wiper arm. The resolution of the potentiometer depends on the number of turns of wire forming the total resistance. High quality potentiometers can become relatively expensive devices, particularly when they are used in high power circuits for accurate position detection. High quality potentiometers are also designed to provide minimal resistance variation with temperature. Potentiometers were primarily used in analog systems and most modern, computer-based control systems achieve accurate user input via other mechanisms such as keyboard or mouse input.

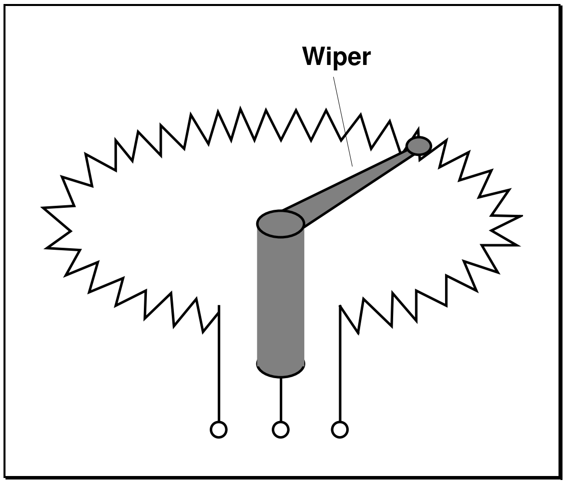


Figure 7.29 - Schematic of Potentiometer

(d) Shaft Encoders and Resolvers

One of the most common requirements of transducers is to convert a rotational or linear position into a voltage signal that can be used in a digital control system. Transducers which achieve this objective are known as resolvers or encoders. The term "resolver" tends to relate to a family of analog devices, whereas encoders are digital devices. By far the larger of the position feedback problems is the conversion of rotational position to voltage because this is a basic requirement of all servo drive systems and is therefore used in robots and CNC machines.

A simple, linear, position feedback system can be created by using a linear version of the potentiometer shown in Figure 7.29. The movement of a wiper arm, along one axis, creates an output voltage which is directly proportional to position. This is analogous to the orientation sensing of the traditional potentiometer.

A more sophisticated form of analog position feedback system for shaft rotation is the so-called synchro-resolver, which was used for some years in NC and CNC machine tool systems. The system works by having one set of electrically-energised rotational coils (known as the armature) connected to the shaft of a rotating device and one stationary set of coils, within which the rotor spins. The sinusoidal and cosinusoidal output voltage waveforms (generated in accordance with Faraday's Law) are then used to detect relative position. The problem with these devices is their cost and complexity but the accuracy can be significantly better than the potentiometer technique.

The most common form of rotational position detection system is the digital shaft encoder. There are essentially two types of shaft encoder - those that provide absolute position and those that provide incremental position. Both of them work on the same principle, which is based on a rotating disk with transparent slots. A set of LEDs is placed on one side of the disk and a set of phototransistors on the other. Each time a transparent slot passes by a LED, the corresponding phototransistor on the other side of the disk generates a pulse. The basic scheme is shown schematically for the incremental position encoder of Figure 7.30.

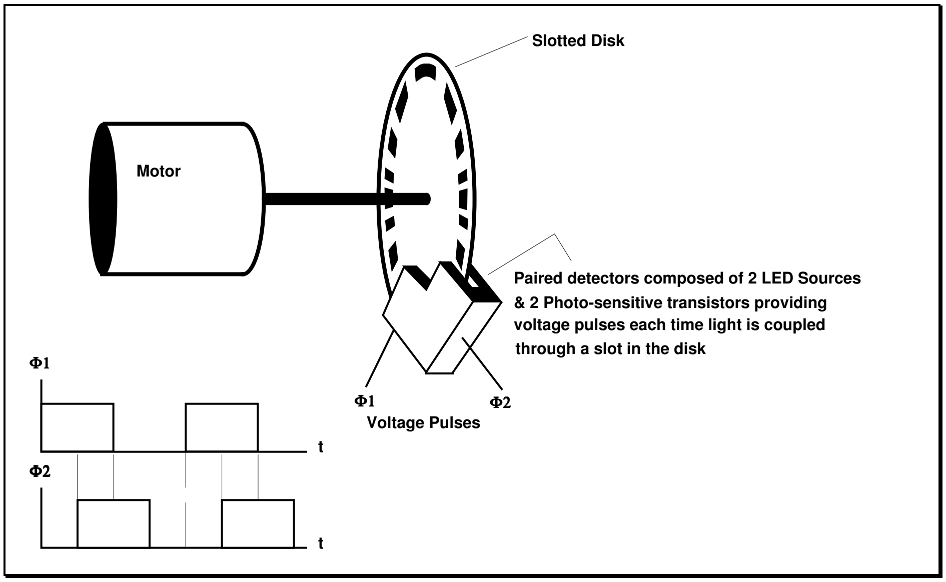


Figure 7.30 - Incremental Position Encoder

In Figure 7.30, two LED-Phototransistor pairs are placed 90° apart in space. The rotational movement of the disk causes an identical pulse train to be generated by each phototransistor - however, the pulse trains are different in phase because of the physical separation of the detectors. Using a counter circuit to count the pulses provides an indication of incremental movement. the phase sequencing between the two outputs provides an indication of direction.

Absolute position encoders are generally less accurate than incremental encoders. In an absolute position encoder, each disk has "n" circular tracks of slots, from the outside to the central hub. An equivalent number of LED-phototransistor pairs are place along the same radial position, on the disk. At any instant, the reading from the in-line phototransistors is an "n-bit" number representing absolute position.

(e) Pressure Transducers

Pressure is a relatively difficult phenomenon to measure directly and the general trend in quantifying force per unit area is to measure the displacement that the force creates. There are a number of different types of pressure transducer available. The majority of these are two-part systems. The first part converts pressure into displacement and the second part converts displacement into a voltage signal which can be used as a feedback element. There are a range of different mechanisms that can be used to create a displacement as a result of a given pressure. These include pistons, Bourdon tubes, diaphragms, etc. Strain gauges, discussed in (f), below, can also be used in some instances.

Of the different types of pressure transducers, those based on the Bourdon tube offer the widest pressure operating range and a degree of accuracy in the order of half of one percent. However, one of the most versatile types of pressure transducers is based upon the piezo-electric crystal (composed of either natural quartz, Barium Titanate ceramics, etc), whose molecular structure creates a diaphragm. The application of a force across two faces of the crystal will generate a potential difference across another two faces, as a result of the electrostatic charge developed. Transducers based on piezo-electric crystal tend to be based upon a shear of the crystal, rather than a simple compression, because the latter causes unacceptably high charges to be built up. Despite their wide operating range, the accuracy of piezo-electric transducers is somewhat lower than other systems and is typically in the order of one percent.

(f) Strain Gauges

The strain gauge operates on the principle that the electrical resistance of a material is dependent upon its geometry. Therefore, if the geometry is altered through a deformation resulting from an applied stress, then the electrical resistance changes accordingly:

$$\frac{\Delta R}{R} \propto \frac{\Delta L}{L} \qquad ...(7)$$

where:

R is the original electrical resistance of the material
L is the original length of the material
The constant of proportionality is called the strain-gauge-factor.

The resistive material, which is the basis of the strain gauge is physically attached to the device under test by some mechanism (such as glue, etc.) and the material needs to be temperature matched to the device under test, so that the thermal expansion and contraction of the device is not recorded by the gauge.

Strain-gauge resistive materials can also be connected in groups of four as a bridge (two vertical resistance elements and two horizontal resistance elements) to eliminate temperature sensitivity and facilitate measurement in compressive and tensile loads.

(g) Temperature Transducers

The most commonly used transducers for measurement of temperature are thermo-couples. These analog devices are based on the junction of two dissimilar metals. One side of the junction is held at a reference temperature and the other is subjected to the test temperature. A voltage is developed across the junction as a result of the temperature difference. There are two major problems with thermo-couples. One is that they are non-linear devices and the other is that they only provide a very low output voltage. Thermo-couples are very widely available as commercial devices, some incorporating amplification stages and so on. In particular, it is possible to purchase devices which incorporate data transmission facilities to enable results to be sent over long distances via a low level network.

(h) Proximity and Level Sensors

A spectrum of devices has been designed to sense parameters such as proximity and level, through widely varying techniques. Proximity can be detected by micro-switches (as in many CNC machines and robots) or by inductive and capacitive sensors, where the moving device has suitable electromagnetic or electrostatic properties (permeance and permittivity). Proximity can also be detected by a light-beam (light source plus photo detector) or ultrasonic beam interruption system. Level sensing depends upon the properties of the medium. Simple level sensing systems can be composed of float-balls coupled to position transducers (such as potentiometers) and more sophisticated systems can be based on light-beam interruption (for translucent fluids) or pressure sensing techniques.

The discussion on transducers could go on indefinitely and there are many commercial organisations regularly issuing catalogues of devices and specifications which, in the final analysis, are far more useful in a practical sense than the general descriptions that can be included in any text book. The solution to most interfacing problems requiring transducers normally begins with a search of such catalogues and the descriptions provided above should only serve as a general introduction into this enormous field.

7.6 Attenuation Problems

One of the most awkward problems that needs to be overcome in the interfacing process is the tyranny of distance. In particular, we often view the conductor (wire), between two nodes, as an ideal element in an electrical circuit as is shown in Figure 7.31. We assume that the wire has no resistance to the flow of current and that therefore, the signal emanating from node A is the same as the one reaching node B.

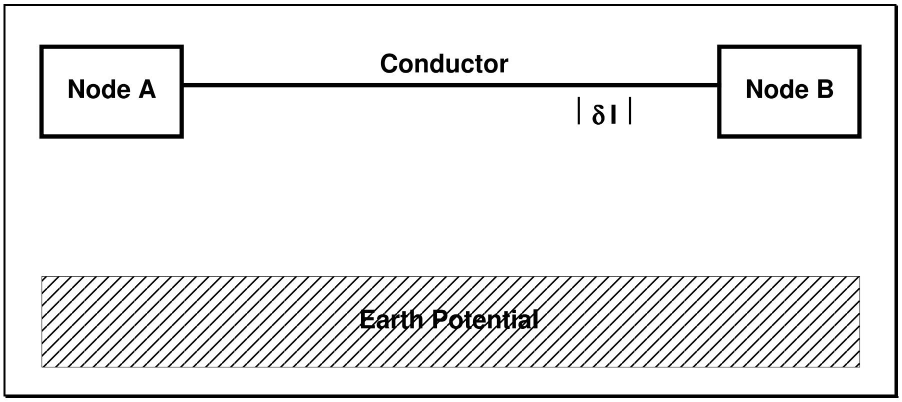


Figure 7.31 - Idealised Point to Point Link

In order to understand why this isn't the case over long distances, we need to look at a more appropriate model of the conducting material. Firstly, we look at an infinitely small section of the wire (δl in length) and examine its physical properties.

There are no "lossless" conducting materials. All materials have some resistance to current flow, and energy is converted to heat within the conducting medium. So the circuit model for our infinitely small length (δl) of wire has a series resistor "R" to reflect the loss of energy at the receiving end of the conductor. Equally, the air between the conductor and earth is not a perfect insulator and therefore provides an alternative path in which current can flow through to earth. The conductance (inverse of resistance) of this alternate path to earth is "G" and reflects that current that does not appear at the receiving end of the conductor as a result of charge flow through the alternate path.

Since the conductor has current flowing through it, a magnetic field is produced around the conductor, and the resultant magnetic flux linkage of the infinitely small section of wire is represented by a series inductor "L". The conductor will also have a certain voltage (and net charge), with respect to earth, causing an electric field between the conductor and earth, thereby giving rise to a capacitance "C".

The series inductance and the shunt capacitance reflect energy storage and release within the conductor. Since both devices store and release energy at differing rates, the voltage at the "output" end of the infinitely small length of conductor will not generally be in phase with that at the "input" end. The circuit model for the entire conductor can be built up from these infinitely small sections and hence we could draw it as shown in Figure 7.32.

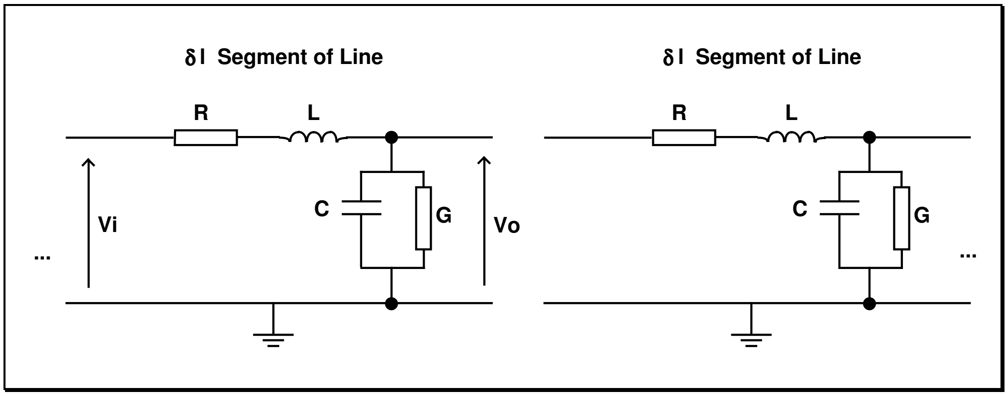


Figure 7.32 - Lumped Parameter Approximation of a Conductor

Figure 7.32 shows what is referred to as a "lumped parameter" approximation of the conductor because all the physical properties that, in reality, are distributed evenly along the line are represented by simple "lumped" circuit elements. Nevertheless, the approximate circuit provides us with some insight into what happens when signals pass through the conductor. Mathematical analysis (Fourier Series) tells us that any voltage waveform, regardless of its shape can be represented by the sum of a number of sinusoidal waveforms of differing frequency and amplitude. So, we can analyse any type of waveform on the conductor by assuming that it is made up of a number of sinusoidal components and using the phasor method to obtain a transfer ratio for the conductor.

For any of the individual sinusoidal components of the digital waveform, the ratio of output voltage (at the end of an infinitely short length of wire δl) over input voltage is obtained as follows:

The impedance of the parallel branch is given by:

$$Z_p = \frac{1}{j2\pi fC + G} \qquad ...(8)$$

The impedance of the series branch is given by:

$$Z_s = j2\pi fL + R \qquad ...(9)$$

The ratio of output voltage to input voltage is then obtained through voltage division and is given by the expression:

$$\frac{v_o}{v_i} = \frac{Z_p}{Z_p + Z_s} \qquad ...(10)$$

Substituting equations (8) and (9) into (10) gives us the complex number expression:

$$\frac{v_o}{v_i} = \frac{\dfrac{1}{j2\pi fC + G}}{(j2\pi fL + R) + (\dfrac{1}{j2\pi fC + G})} \qquad ...(11)$$

The magnitude ratio of output voltage over input voltage is given by:

$$\left|\frac{v_o}{v_i}\right| = \frac{1}{\sqrt{(2\pi fRC + 2\pi fLG)^2 + (1 + RG - 4\pi^2 f^2 LC)^2}} \qquad ...(12)$$

The phase difference of the output voltage with respect to the input voltage is given by:

$$\Phi_{v_o - v_i} = -\mathrm{Arc}\tan\frac{2\pi fRC + 2\pi fLG}{1 + RG - 4\pi^2 f^2 LC} \qquad ...(13)$$

Expressions (12) and (13) are both frequency dependent and therefore we can say that the phase and magnitude-ratio of output voltage to input voltage will be different for each sinusoidal component of the waveform on the conductor. Substituting some "limit" values into these expressions, we can observe that in the infinitely small section of conductor:

- Very high frequency ("f" tending to infinity) sinusoidal components will be attenuated (diminished) to zero
- Low frequency ("f" tending to zero) sinusoidal components will be attenuated by a factor of (1 + GR)

- Low frequency ("f" tending to zero) sinusoidal components will be slightly "shifted" in phase with respect to the input.

If we were to exaggerate this effect for, say a digital voltage waveform, then the output voltage at the end of the infinitely small section would look as the voltage waveform illustrated in Figure 7.33.

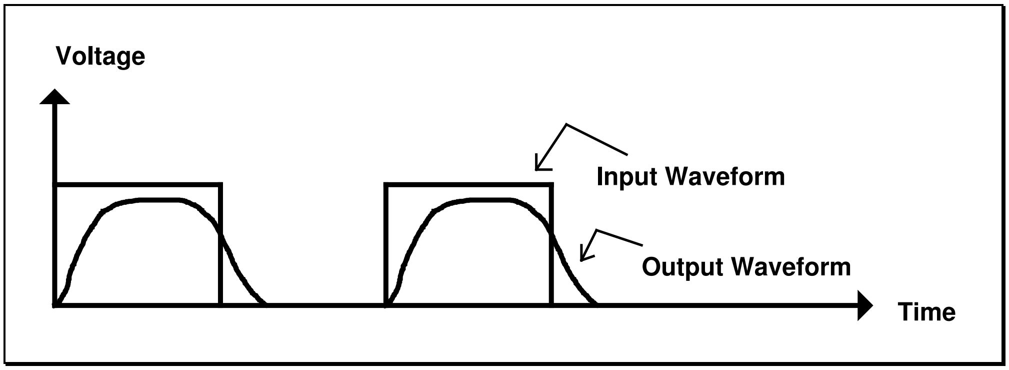


Figure 7.33 - Exaggerated Output Voltage at the end of δl Segment

The attenuation of high frequency sinusoidal components, in the output waveform, means that the edges lose their sharpness. The phase shift in low frequency components, means that the output waveform appears to "lag" behind the input waveform. As a result of attenuation in some of the sinusoidal components, the output waveform also appears attenuated in some areas. If we then say that since a conductor is composed entirely of these infinitely small sections, the distortion and attenuation of the output voltages, with respect to the input voltages, is increased with length. If the length of the transmission line is sufficiently large, then the output waveform will be attenuated and distorted to such an extent that it cannot be discerned from noise. This is shown schematically in Figure 7.34 for a digital signal, but the same is true for analog signals.

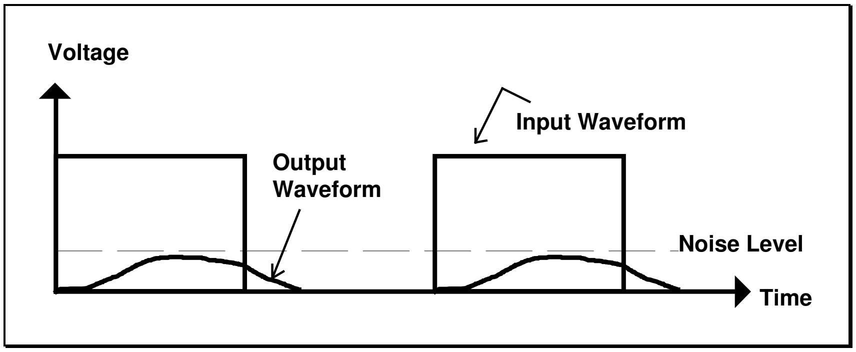


Figure 7.34 - Degenerative Effects of Long Conductors

Now that we know what happens to voltages over long lengths of conductor, we need to consider what happens in a typical interfacing problem. Figure 7.35 shows a relatively common scenario.

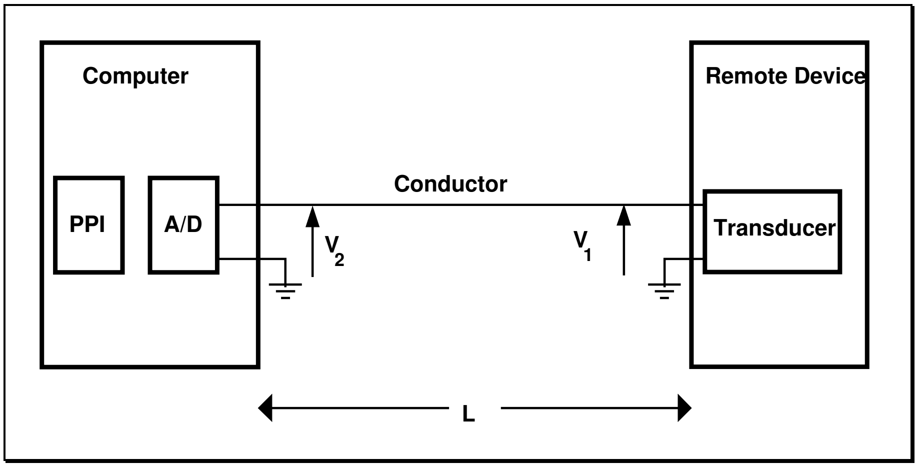


Figure 7.35 - A Common Interfacing Problem where the Distance "L" Between the Transducer and Control System is Large

In Figure 7.35, we have a transducer separated by a long distance, "L", from the control computer. The term "long" is difficult to quantify, because it ultimately depends upon the original size of the signal (v_1) and the characteristics of the conductor (resistance per unit length, etc.). However, if we were to place an order of magnitude figure of say 10 metres on the term "long" then it gives one some feel for the problem.

Transducers normally produce relatively low output voltages which may attenuate to noise level over a long distance. Further, from equation (12), we know that the conductor connecting the two nodes is effectively acting as a low-pass filter that attenuates high frequency components of the transducer output. If the computer needs the high frequency components for monitoring or control purposes, then clearly the conductor may be responsible for corrupting the signal acquisition process.

The obvious solution is to amplify the signal at the transducer end before it is transmitted down the conductor. This can be achieved with a simple operational amplifier circuit. However, all amplifiers need a low-voltage d.c. power supply which is not normally available at the remote end. The provision of such a power supply can be an issue, particularly if there is no power whatsoever at the transducer end that can be converted to low-voltage d.c. It may then be necessary to use a battery solution or solar cells, if sufficient light is available. The attenuation problem is particularly irritating if there are a number of widely-spaced transducers that each need to have their own power supply.

There is no simple solution to the attenuation problem, but one alternative is to use data communications techniques, introduced in 7.7. In fact, some commercial transducers come equipped with data communications transceivers that help overcome the problem of attenuation.

Another problem related to long distance interconnection of low voltage devices is the problem of electromagnetic interference or EMI. Long conductors are liable to have voltages induced in them as a result of magnetic fields produced by currents in other high power conductors. This is common in industrial environments, particularly where high-power induction machines or furnaces are switched on and off. The effect can be minimised by using conductors shielded by a conducting foil or copper-braid (thereby creating a Faraday Cage).

EMI is also a problem where two long conductors, both carrying small signals, are close to one another and in parallel - the current in one induces a voltage in the other. The phenomenon is really EMI, but is more specifically referred to as "cross-talk". It is resolved by removing the parallel path between conductors which is, in turn, achieved by twisting pairs of small-signal conductors around one another. There are many commercially available cables composed of twisted-pairs for minimising cross-talk. Normally, twisted-pair cables are also shielded with foil and/or copper braid and are bound together in some form of plastic sheath.

7.7 Data Communications

Data communications techniques are useful for interfacing purposes because they can provide a modular hardware structure which one can use to transmit information over relatively long distances. Data communications techniques are primarily used to transfer information from one computer to another, but in many instances they can also be used to transfer data to and from remote sensors, actuators and transducers. This is shown schematically in Figure 7.36.

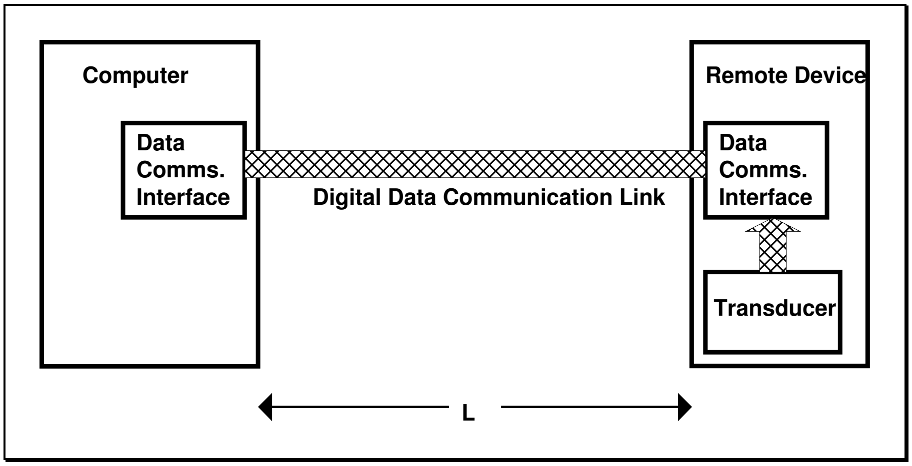


Figure 7.36 - Interfacing Via Data Communications Links (Point to Point)

There are basically two types of data communications links that can be established. They are:

(i) Point to point links
(ii) Networks:
  - Star
  - Bus
  - Ring

In a point to point link, only two nodes are generally involved and they are connected together via a number of conductors that form the data communications link or transmission medium. The point to point environment is shown in Figure 7.36. In a network, a number of nodes can be interconnected via a range of different topologies (star, bus and ring), with the bus topology currently being the most prevalent. The bus topology is shown in Figure 7.37.

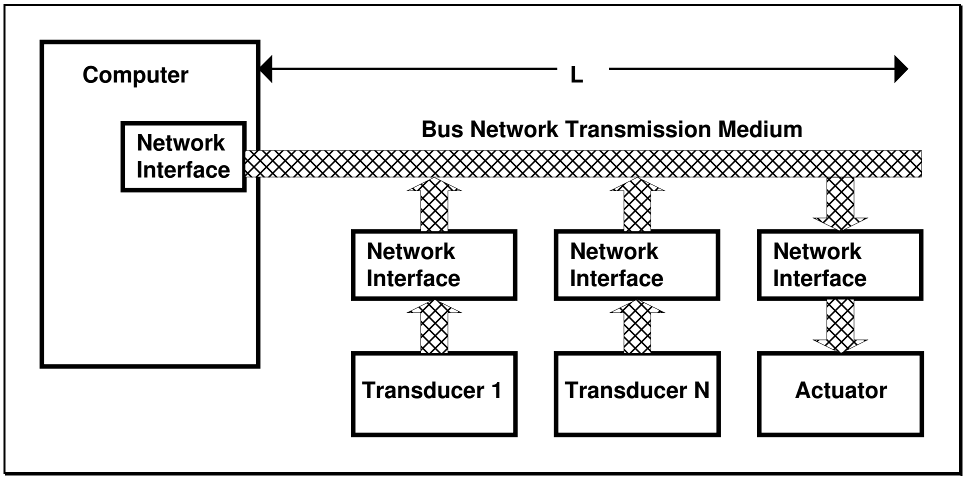


Figure 7.37 - Interfacing Devices Via a Bus Network

In both the point to point link and the network, data is transmitted in a digital format. The role of the data communications or network interface is to take the raw signal and convert it into the appropriate size and digital format for transmission. The problem however is that interfaces are not always available and this means that the system developer has to undertake this role by combining a number of commercially available components. In some cases, particularly for common transducers, the manufacturers have already pre-empted the need for networks (because they have realised that their transducers may be situated some distance away from the control or monitoring computer) and they have integrated the transducer with a suitable network interface.

Communications between devices (nodes) on a network or point to point link can occur in two ways:

- Parallel
- Serial.

External parallel communications is analogous to the data/address bus communications that occurs within a computer system and it also tends to exhibit the same limitations, particularly transmission distance. The most common parallel networking system available is the so-called General Purpose Instrumentation Bus (GPIB), otherwise known as the Hewlett-Packard Instrumentation Bus (HPIB) or as IEEE 488. The most common point to point parallel link is the computer to printer link, commonly referred to as a "Centronics" link. The Centronics link was originally designed for computer to printer communication but can theoretically be used for a range of different applications.

Serial communications is the more common form of communication outside the computer system because it minimises the number of conductors between points or nodes on a network. In a serial link, information passing out of the computer data bus is converted into a sequential pulse train via a clocked shift-register. The data can then be transmitted on a single conductor (plus reference conductor) or converted to light pulses by LEDs for transmission on an optic fibre cable. There are a plethora of standards and defacto standards for serial communications and for networks and this has tended to make the networking of devices such as transducers somewhat difficult.

In terms of interfacing computers to mechatronic systems, the other problem with point to point links and networks is that can be relatively slow for real-time processing of data. This is particularly true of serial systems.

The field of data communications is enormous and is the subject of the complementary text to this one (Data Communications and Networking for Manufacturing Industries), so it will not be pursued in any detail herein. The main objective of introducing it in this text is because of the number of commercially available transducers with proprietary network interfaces that can make the task of interfacing devices over long distances considerably easier. However, if one had to design a network interface from first principles, for each type of transducer, then it is doubtful that data communications would be the preferred option. In particular, in Figures 7.36 and 7.37, if we assumed that the transducers provided an analog output signal, then the data communications or network interface for the transducer would need to:

- Convert the signal to digital form and scale it to an appropriate size
- Convert the signal to the data communications form, with appropriate timing, etc.
- Amplify the signal to the required data communications level for transmission
- Read information from the host computer (via the data link), interpret it and adjust transducer parameters accordingly
- Respond to and obey the rules (communications protocol) specified for the link or network.

Similarly, the data communications or network interface at the computer end would need to:

- Interface to the internal system (Address/Data) bus
- Read incoming signals and generate outgoing signals based upon the application software running on the computer
- Respond to and obey the rules (communications protocol) specified for the link or network.

None of these tasks, in their own right, is trivial and the development of systems which combine all these factors requires a significant amount of skill, thereby making non-standard solutions difficult to generate.

7.8 Combining the Interfacing Stages

Now that we have examined the basic hardware elements in the interfacing process, we can return to the global diagram that has been featured throughout this book. This has been redrawn in Figure 7.38, where each of the basic stages has been labelled with a number in parentheses.

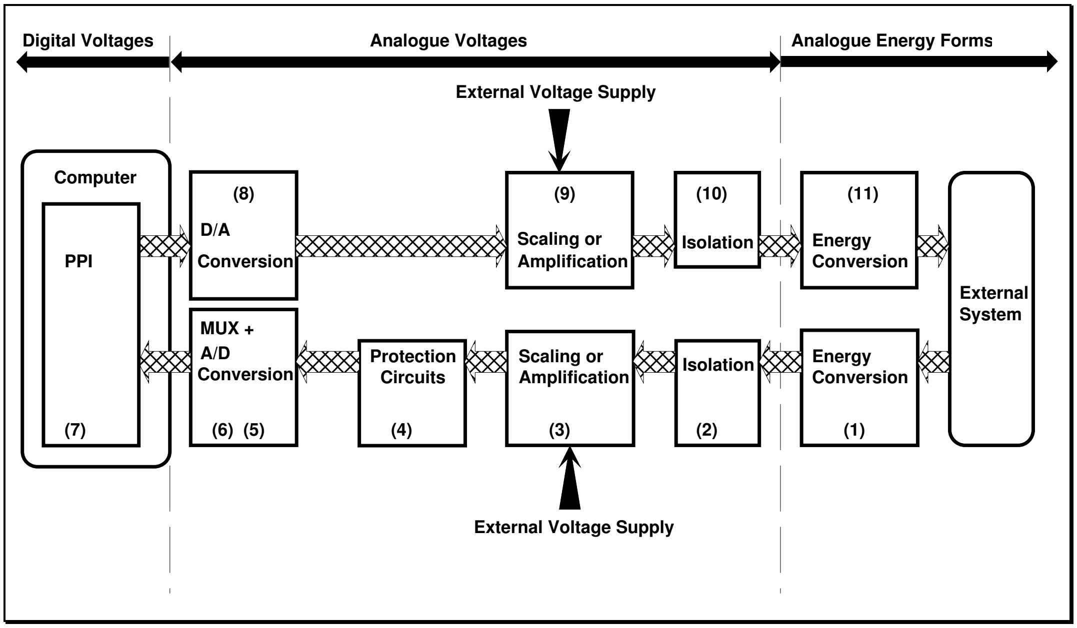


Figure 7.38 - The Basic Closed Loop for the Interfacing Process

With the hindsight that is available after having read the previous sections of this chapter, Table 7.1 summarises the basic options which are available for each of the interfacing stages labelled in Figure 7.37. Table 7.1 should provide a reasonable summary of the issues involved in the interfacing of digital systems to mechatronic systems. However, we have not addressed the problems that arise from the physical environment itself. These issues include chemical, vaporous, metallic (swarf), thermal, vibrational and electromagnetic problems common to a range of different industrial situations. A full discussion of these is outside the scope of this book. However, it is indeed fortunate that for many years, specialist industrial enclosures, disk-drive systems, key-boards, screens, etc. have been available for protecting digital computer controls in situations where they must be located in a harsh environment. The only limiting feature of industrial enclosures is that they are relatively expensive (generally costing more than the computer or digital control system itself) and their costing needs to be considered as part of the complete interfacing process.

Interfacing Stage	***Typical Devices***	***Issues / Selection Criteria***
Feedback		
(1) Energy Conversion	Transducers (Strain gauges, Thermo-couples, Current-Transformers, tacho-generators, encoders, etc.)	Energy rating; output voltage range; linearity; isolation of output voltages from inputs
(2) Isolation	①Transformers ②Opto-Isolators ③Operational Amplifiers	Frequency response; linearity; extent of electrical isolation
(3) Scaling	①Transformers ②Operational Amplifiers	Frequency response; linearity; scaling range
(4) Protection	①Zener Diodes ②Relays ③Thyristors	① Reverse diode breakdown voltage ② Switching speed and turn-on current for relays
(5) Multiplexing	Multiplexer Chips	Number of input channels; switching speed
(6) A/D Conversion	Analog to Digital Converter Chips	Number of output bits (resolution); conversion type and speed; cost, integration with PPI (PIA)
(7) PPI	Programmable Parallel Interface or Peripheral Interface Adaptor	Number of input and output ports; number of bits per port
Driving Force		
(8) D/A Conversion	Digital to Analog Conversion Chip	Number of input bits; integration with PPI (PIA)
(9) Scaling / Amplification	① Transformers for scaling ② Operational Amplifiers for Amplification ③ External circuits switched by relays activated by computer ④ Pulse-Width Modulator Amplifiers (PWMs)	① ② & ④ Scaling ratio; frequency response; buffering; power consumption; linearity; heat generation ② Configuration for current or voltage amplification ③ Relay turn-on current; switching speed ④ Duty cycle, Parasitic switching effects
(10) Isolation/ Buffering	①Transformers ②Opto-Isolators ③Operational Amplifiers	Linearity; frequency response; extent of system isolation
(11) Energy Conversion	Actuators - Motors, Solenoids, Relays, Speakers, Heating, Lighting, Chemical (Batteries), etc.	Depends on final system

Table 7.1 - Interfacing Options for Figure 7.38

7.9 Commercial Realities of Interface Design

Upon reading sections 7.1 to 7.8 of this text, one might come to the conclusion that the design of an interface between a complex digital circuit, such as a computer, and a complex external system is a difficult task. In a commercial sense, one also has to consider the developer's time and the technical resources required to design, build and debug an interfacing circuit. At the very least, voltmeters, ammeters, soldering and desoldering equipment, a high bandwidth oscilloscope, PAL programming board and personal computer may be required. One also needs to consider the external services that may be needed to produce a professional looking board - including items such as printed circuit board fabrication and design if in-house software is not available.

Most organisations ultimately conclude that this work can only be amortised and hence justified on two occasions:

- When many boards need to be produced
- When no commercial (off-the-shelf) packages are available.

In practice therefore, the solution to a one-off interfacing problem is generally to tailor a commercial interfacing board that most closely resembles the final product to be produced. This simplifies the task enormously because it enables designers to concentrate on the system at hand rather than the intricacies of debugging circuits that generate "glitches" at seemingly random times. However, the commercial interfacing board solution is also rather costly because boards are designed in a general-purpose manner, normally providing more functionality than may be required for any one application. Moreover, the selection of an appropriate board still requires a relatively sound understanding of all the basic design principles raised throughout this text.

There are an enormous range of commercial interfacing boards, including servo drive control boards for motors, closed-loop PID controllers, etc. Many of these boards plug directly into the back-plane bus of a personal computer or workstation and provide a building-block solution to industrial problems. Often, the complementary part of such boards is a range of transducers and actuators that enable many common industrial problems to be resolved without special hardware design.

The fundamental problem with designing hardware interfaces to computer systems from first principles is the complexity of the back-plane bus on the computer system. Mapping programmable parallel interfaces and A/D and D/A devices onto the address and data bus structure of a modern computer can be a difficult and time consuming task. To a large extent it is also a question of "reinventing the wheel". A half-way solution between a full commercial interfacing board with on-board processing of signals, etc. is a so-called "Input/Output (I/O)" or "A/D Board". Normally such boards plug directly into the back-plane bus of a computer system and provide analog and digital inputs and outputs. The application specific portions of the interface are then left to the designer.

Higher quality I/O boards resolve the problem of isolation by providing opto-isolated connections to the outside world and other useful hardware items such as counter-timers, etc. Low-end boards are generally only composed of a memory-mapped PPI device with A/D and D/A converters.

With these points in mind, one should generally approach interfacing problems by undertaking a study of commercial catalogues to ensure that one is not re-inventing the wheel (unless that is one's intention). One should also consider the possibility of the Programmable Logic Controller (PLC) as an alternative to either personal computers or specialised digital control boards. In many instances, the PLC can provide sequential control of simple systems and resolve many of the interfacing problems that are encountered in industry. Table 7.2 provides a rough-cut guide to the sorts of problems that typically arise and, for each, a logical sequence of steps to consider in resolving the problems. Note particularly how we use the personal computer as a low-cost tool for interfacing - in low volume interfacing applications, the cost of a personal computer is less than 50% of an engineer's salary for one week. This means that we no longer simply treat the PC as a complex device for computation but also, because it has a widely accepted back-plane bus structure, as a low-cost junction point for commercially available transducers and interface boards.

Problem	***Available Professionals***	***Possible Courses of Action***
Old sequential control machine (relay controlled) requires modernising of control	Mechanical / Manufacturing Engineers	Consider PLC control - evaluate range of sensors and actuators from PLC manufacturers data books
Need to develop 200-off low-cost, hand-held remote controllers (digital) to open and close safety doors on machinery	Electrical/Electronic Engineers	Consider development of a digital circuit from first principles - possibly using a low-cost processor and commercial transmitter / receiver combination
Need to control and monitor a chemical process in the laboratory environment	Mechanical /Manufacturing / Chemical Engineers	① Consider a personal computer (PC) based solution using a commercial I/O card and transducers ② Consider a PC based solution with commercially available data-acquisition and control hardware and software

Problem	*Available Professionals*	*Possible Courses of Action*
Need to develop a special-purpose robot control system with multiple axes	Electrical / Mechanical / Manufacturing Engineers	① Consider basing the system on a commercial (PC) computer and plug-in servo control cards with on-board processing and PID control. Only develop supervisory software. ② Consider basing the system on a commercial (PC) computer mother-board with commercially designed I/O facilities and design remainder of interface hardware and software from first principles
Need to develop 4-off high-bandwidth signal acquisition, processing and control systems that are coordinated by a host computer	Electrical / Electronic / Mechanical / Manufacturing Engineers	① Consider Development based on a Digital Signal Processor development kit plugged into the back-plane of a PC - After development, consider transferring the entire control and monitoring function to a stand-alone DSP device networked to a host computer system ② Consider using 4 commercial PC motherboards for the control and one full PC for host coordination - use a commercial network

Table 7.2 - Sample Interfacing Issues and Possible Courses of Action