LECTURE NOTES

UNIT–I:
Signals and their representation

Measuring Systems, Performance Characteristics, – Static characteristics – Dynamic Characteristics – Errors in Measurement – Gross Errors – Systematic Errors – Statistical analysis of random errors – Signal and their representation – Standard test, periodic, aperiodic, modulated signal – Sampled data pulse modulation and pulse code modulation.

INTRODUCTION

Determining a quantity or variable using a physical means is called the measurement and the means by which the quantity is determined is called Measuring Instruments. Thus, an instrument may be defined as a device for determining the value or magnitude of a quantity or variable. An instrument enables a person to determine the value of an unknown quantity.

Some of terms are used in the measurement work, which are defined below.

1. **True Value:** It is the average of an infinite number of measured values.
2. **Error:** It is the difference between the measured value and the true value.
3. **Index Scale:** It is the set of marks or divisions.
4. **Index Number:** It is the number of divisions moved.

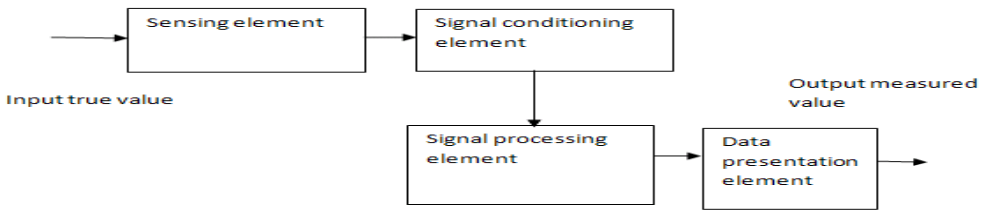
The essential requirements of a measuring instrument are:

(i) It should not alter the circuit conditions,

(ii) Power consumed by it should be small.

MEASUREMENT SYSTEMS

All measurement systems can be thought of being made of one or more of these blocks of Figure1. At the input we have the input element to be measured, temperature, displacement, .etc. that affecting the sensing element. Actually, sensing element is the process of continuous energy conversion from one form depending on what we want to measure, e.g, from mechanical form, optical form to finally electrical form, and then the electrical form get finally transform further to digital form before the output. The signal at the sensing element is in the form of voltage and current. The real pressure or real temperature which it exist at the sensing element, then the sensing element brings it generally to some sort of electrical form or electrical parameters like resistors, capacitor changes or in the form of voltage and current which have be further manipulated by electrical circuit called signal conditioning element.(sometimes amplifier or conversion from resistor to voltage). Then further signal processing goes on to remove noise and make it linear or something like that. Some of it can be analog and some of it can be digital. Finally, it goes to the data presentation element where data is utilized so it can be recorded, or can be displayed or can be controlled.



Here is the weight measurement system as an example. The input is the true weight which is sensed by a mechanical member that called load cell. The load cell converts the input (weight) to the strain that sensed by another member called strain gauge. The latter converts it to a resistance form then we feed it to the electrical circuit called wheat-stone. The wheatstone converts the resistance to a low voltage level. The low voltage goes to the amplifier. Finally, the amplified signal passes to the digital signal processing and from there to a microcontroller where some digital processing is done. Finally, the data may be displayed along with a unit. Here is how real measurement system looks like. It is basically cascaded of several blocks including the sensor, signal conditioning, plus some computer elements.

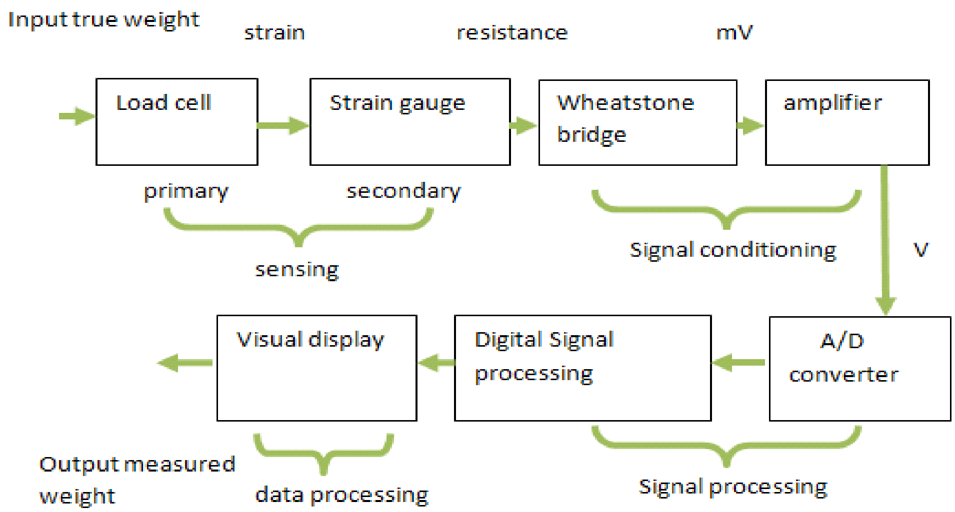


Figure 2: Weight measurement system

Sensing is actually extremely important in automation from various points of view:

a. Product quality control, because quality control is actually accessed by a sensor element.

b. Manufacturing process control. All the process control are closed feedback control. So the critical element is the feedback element. The performance of the control system is critical to the sensor.

c. Process monitoring and supervision. All these can be done plus providing energy efficient, obtaining set-point, form all these we need sensors.

d. Manufacturing automation, this how the manufacturing automation systems can be put together using programmable logic controllers. Then you find that they use various kinds of sensors. So sensors are extremely important in automation. They will give a value or information about physical quantity we need to know how to characterize the behavior of this device called a sensor in instrument, we need to understand instrument characteristics.

The performance characteristics may be broadly divided into two types, namely, 'statistics' and 'dynamic' characteristics. Static characteristics where the performance criteria for the measurement of quantities that remain constant. Or vary only quite slowly. Dynamic characteristics on the other hand, shows the relationship between the system input and output when the measured quantity is varying rapidly.

Calibration

The procedure that involves a comparison of the particular instrument with either a primary standard or a secondary standard with a higher accuracy than the instrument to be calibrated.

From Figure 3, the measurand 1 is considered to be the true value, while the measurand 2 is not only from the sensor instrument under calibration but it can be a result of other factors. It may be a result of temperature. For example, in the case of the weight measurement, the strain gauge,(the resistance change) is not only a function of the weight, it also a function of temperature. Because every resistance has some temperature coefficient. There are some noise can be induced from a power supply or from some power lines especially in the industry environment, there are plenty of noise sources. This signal (noise) can affect the measurement. When you want to characterize the instrument, you have to characterize it to respond these kinds on inputs.

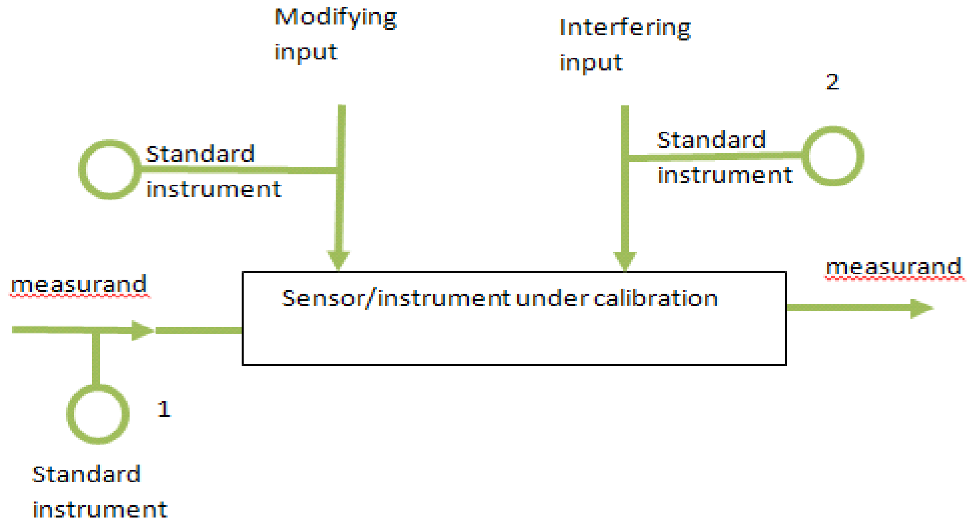


Figure 3 Calibration of an instrument

Essentially, we try to measure the measurand, the output of the instrument, and the modifying input like the temperature. Then we establish the characteristics of the instrument. Since the instrument must happen constructed to be unaffected by modifying input, so the most important thing is to see how the instrument characteristics depend on the measurand. There are different standards of instrument. The instrument can be calibrated against laboratory standard. The laboratory standard instrument can also be calibrated from time to time against other standard like the secondary standard which is special instrument that exists in some testing houses. So from time to time you should send the instrument to test houses and get calibrated. On the other hand, the guest house instruments again have to be calibrated against very accurate national standards. So in this way you have what can be called change of standards of increasing accuracy and add different levels you always calibrate according with respected instrument.

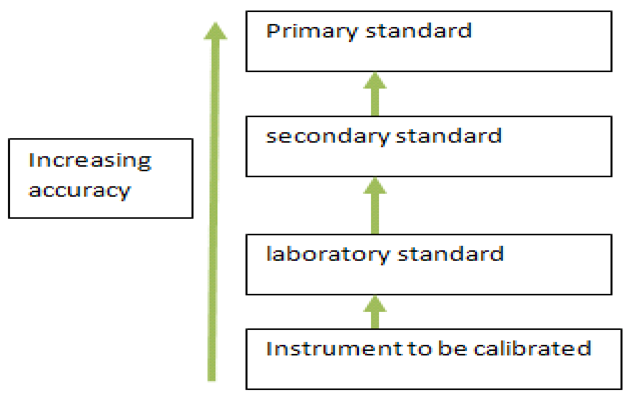


Figure 4: increasing accuracy with different levels of standards

Span
If in a measurement instrument the highest point of calibration is x2 units and the lowest point is x1 units, then the instrument range is x2 units and the instrument span is x2-x1.

Accuracy
It refers to the closeness of an instrument reading to the true value of the quantity or variable under measurement. This term describes the algebraic difference between the indicated value and the true value or theoretical value of the measurand. In practice, accuracy is usually expressed as a percentage of full scale output or percent of reading or digits for digital readouts.One of the most important parameter called accuracy. Usually, "accuracy is expressed as accurate to within x percent" of reading/span. It means that true value within ±x percent of instrument reading/span at all calibration points of the scale. When a temperature transducer with an error of ±1% of reading indicates 100oC, then the true temperature is between 99oC and 101oC.

Precision: It refers to the measure of the reproducibility or repeatability of the measurement. It is a measure of the degree to which successive measurements differ from one another. If a number of readings of a voltmeter are taken, then the expected value of 1.0 volt is not obtained on every occasion. A range of values such as 0.99, 1.01, 1.00, 1.02, 0.98, etc. are obtained about the expected value. The effect is termed as a lack of ***repeatability*** in the instrument.

Linearity
The calibration curve of a real instrument is typically not a exactly straight line. But still is very useful to imagine the system real one. It is very easily to interpret the true value. If you have an instrument sensitivity 10mV/oC, and if it gives 25mV signal, then you know that the temperature is 2.5 oC. So you can get just by dividing by a number. From that point of view, it is very attractive to express the characteristics of a linear one, but it is not line. Therefore, why you mention a line which can be used for inducing the true value from a reading. For ease of use, it is desirable that the reading of an instrument is considered Linearly related to the quantity being measured. The linearity specification indicates the deviation of the Calibration curve from a good fit straight line of it. How do you obtain the straight line? It can be obtained in various ways: (non-)linear method; this method is defined as the maximum deviation of an output reading from the good fit straight line and may be expressed as a percentage of full scale or reading. The true characteristics of the instrument is indicated by the curve shown in Figure5. So we can approximate it by the straight line. Therefore, the true value will be within two limits shown in Figure 5.

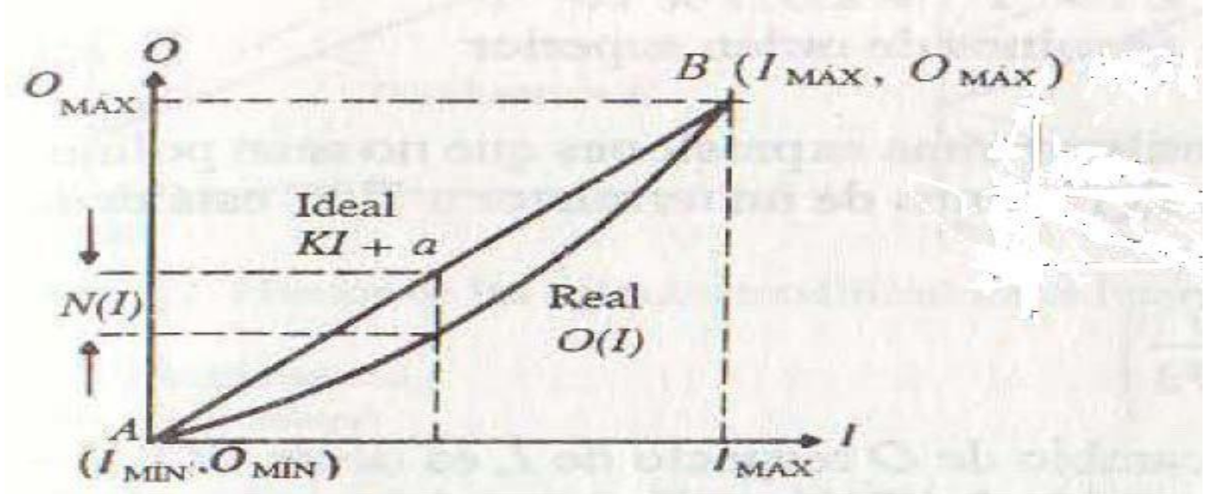


Figure 5: Non-linear method for obtaining a straight line

Actually linearity specification is only linear specification in the sense that it indicates deviation from linearity.

Sensitivity

The slope of an static calibration curve evaluated at an input value is the static sensitivity.

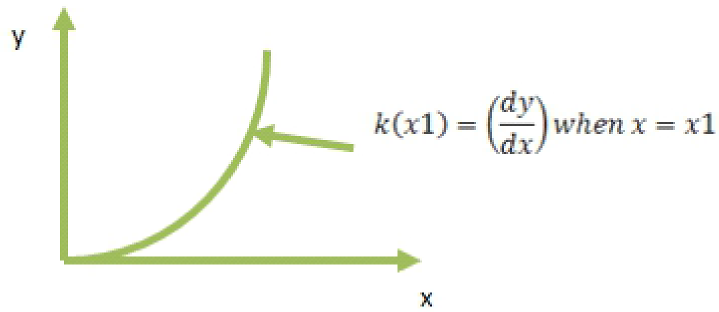


Figure 6: The slope of an static calibration curve

If you have a calibration curve, then you get a straight line. In case you have a linear characteristic, then you will have a single sensitivity. However, if you have a very non-linear one, sometime you do various things. In the case of three sensitivity figure, one sensitivity figure you apply along the line A, another sensitivity figure you apply at the region B which is the average slope of the line , and a third sensitivity figure you apply at the range of C.

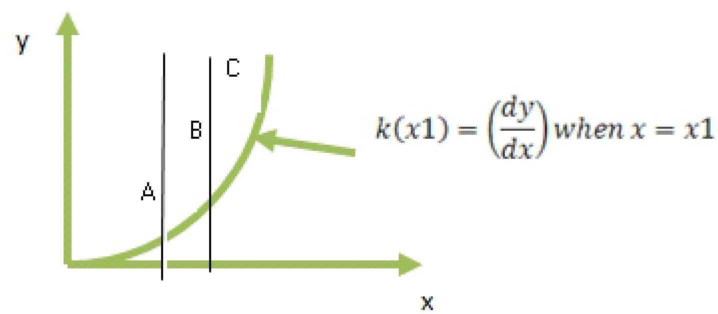


Figure7: Non-linear calibration characteristics curve

Repeatability or Precision

The repeatability of an instrument is the degree of closeness with which a measurable quantity may be repeatedly measured. It is defined as the maximum measure of variation in the measured data for a particular input value given by standard deviation δ .

Resolution

The measurement resolution of an instrument defines the smallest change in measured quantity that causes a detectable change in its output. For example, in a temperature transducer, if 0.2 $_o$C is the smallest temperature change that observed, then the measurement resolution is 0.2 $_o$C.

Dead Zone

Dead zone is the largest value of a measured variable for which the instrument output stays zero. It occurs due to factors such as static friction in a mechanical measurement system.

Hysteresis

Hysteresis error refers to the difference between responses to increasing and decreasing sequence of inputs. It can occur due to gear backlash in mechanism, magnetic hysteresis or due to elastic hysteresis.

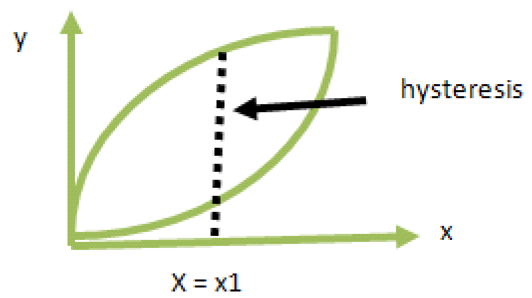


Figure 8: hysteresis

Bias/offset

It is the constant component of error that may be assumed to exist over the full range.

Sensitivity/Gain error

It is the component of error which is assumed to be proportional to the reading.

Correction

Instruments often provide facilities to correct for these error using signal conditioning circuitry.

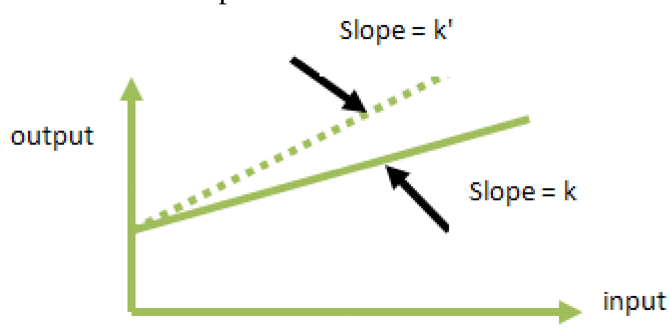


Figure 9: Bias and Gain error

now the errors that we have you know the, so we have actually typically an instrument is suppose to have a, suppose to have calibration curve. But, the reading that it has may not exactly match with the calibration curve it is if you, if you, if you read out an ammeter then it has some scale fixed. But, if you send exactly one ampere current then the, then the needle may not stand at one ampere, so this is the error. Now, the error is typically you know characterize as in into two different kinds, so since the instrument is actually assume to be linear instrument.

So, it is assume that the error can be of two types the first type is called bias of offset which is a constant error, which is, which is going to stay throughout the range. So, may be at half at when you have a reading of when you have an actual current of 2 amperes you reading shows 2.5. When you have 3 ampere it showed 3.5, when you have 10 ampere it shows 10.5, so you have a 0.5 ampere of bias. If you see ammeters normal ammeters you will find that such biases can be corrected by you know screwdrivers there are, there are, there are sometimes zero adjusts.

Similarly, there can be see there can be again error, so you have a sensitivity while we have a nominal sensitivity which is indicated by the scale and your actual instrument sensitivity may actually deviate from that and then you have a sensitivity or gain error. The error in reading due to this gain error is going to be proportional to the ready, so if you have if you have measuring 10 degree centigrade.

Then the error due to gain error is going to be half of if what you measure due to 20 degree when you when you measure 20 degree centigrade. So, we assume the errors are of two kinds and these typically in typical sensors and instruments very often they can be corrected by electronic signal conditioning means.

Drift

so sometimes what happens is that even if you correct even if you correct at any at some point of time during calibration even if you correct for the bias of the gain error you have drifts in the bias on the gain. So, again such bias and gain errors can develop due to you know variations in temperature variations in time or some other conditions. So, the rate at which it these will develop are characterize by a performance characteristic called drift, so typically drift is characterized for temperature and time.

The calibration of an instrument is usually performed under controlled conditions. As variations occur in these conditions and also with passage of time, the instrument characteristics change. Usually, typical factors for which drift is characterized are temperature and time.

Dynamic Characteristics:

It describes the ways in which an instrument or measurement system responds to sudden changes to the input. In general, the dynamic response of the measurement system is expressed in the form of a differential equation. For any dynamic system, the order of the ***differential equation*** which describes the system is called the ***Order of the System***.

(*i*) Zero-order System: It has an ideal dynamic performance, because the output is proportional to the input for all frequencies and there is no amplitude or phase distortion. A linear potentiometer is an example of a zero-order element.

(*ii*) First-order System: A first-order instrument or system is characterized by a linear differential equation. The temperature transducer is an example of first-order measuring devices, since this is characterized by a single parameter, i.e., *time constant*, *T*.

The *differential equation* for the first-order system is given by

$$x(t) = y + T \cdot \frac{dy}{dx}$$

where, $x =$ Input

$x(t) =$ Time function of the input

$y =$ Output.

(*iii*) Second-order System: This type of system is characterized by the second-order differential equation. The example of the second-order system is the mass-spring system of the measurement of the force. The second-order instrument or system is defined by the equation

$$a_2 \frac{d^2 y}{dt^2} + a_1 \frac{dy}{dt} + a_0 y = b_0 x$$

The solution of the *Equation* is given as

$$\frac{1}{\omega_n^2}\frac{d^2y}{dt^2}+\frac{2\xi}{\omega_n}\frac{dy}{dt}+y=Kx$$

where, $K=\frac{b_0}{a_0}=$ *Static frequency*

$$\omega_n=\sqrt{\frac{a_0}{a_1}}=\textit{Natural frequency}$$

$$\xi=\frac{a_1}{2\sqrt{a_0a_2}}=\textit{Damping ratio}$$

Thus, the second-order system is characterized by the two parameters —
the *natural frequency*, f_n or the *angular frequency*, $\omega_n\ (=2\pi f_n)$, and the
damping ratio, ξ.

In the second-order system, the *natural frequency* is the index of *speed of response*, whereas the *damping ratio* is a measure of the *system stability*. The second-order instrument is more common than first-order types.
The dynamic characteristics of an instrument or the measurement system are as follows:

(*i*) *Respond Time*, (*ii*) *Fidelity*,
(*iii*) *Measuring lag*, and (*iv*) *Dynamic error*.

(*i*) Respond Time: It is an important parameter to describe the dynamic response of an instrument. It characterizes the instrument to a step change in the measurand (input). It includes *rise time*, *delay time* and *time constant*.
(*ii*) Fidelity: It is defined as the degree of the measurement system. It indicates changes in the measurand without any dynamic error.
(*iii*) Measuring Lag: It is the retardation or delay in the response of a measurement system to changes in the measurand.
(*iv*) Dynamic Error: It is the difference between the true value of the quantity under measurement changing with time and the measured value of the quantity. It also referred to as ***Measurement error***.

ERRORS IN MEASUREMENTS

Error is the difference between *the true value* and *the measured value* of a quantity such as displacement, pressure, temperature, and the like. No electronic component of instrument is perfectly accurate. All have some
error or inaccuracy. The measurements cannot be made with perfect accuracy. It is important to find out how different errors have entered into the measurement and what the accuracy is. A study of error is a first step in finding ways to reduce them. Such study also determines the accuracy of the final result. *Error* is inevitable in any measurement. Well-designed electronic instrumentation systems limit the error to a value that is acceptable in terms of the accuracies required in an engineering analysis or the control of a process.

Limiting or Guarantee Errors

The design, *the materials used* and *the workmanship* are the important factors for the accuracy and precision of an instrument. In most of the instruments the accuracy is guaranteed by the manufacturer for the quality of the instrument to be within a certain percentage of full scale reading. The values of the circuit components, like *resistors*, *capacitors* and *inductors*, are mentioned within a certain percentage
of the rated values specified by the manufacturer. The limits of these deviations from the specified value are defined as ***Limiting errors***.

The magnitude of a quantity *Q* having a specified value *Q*1 and a *limiting error* ± _*Q* must have a value between the limits (*Q*1 – _*Q*) and (*Q*1 + _*Q*) or *Q* = (*Q*1 ± *dQ*).
For example, the specified value of a resistor is 560 _ with a limiting error of ±10 _. So, the value of the resistor will be between the limits (100 ± 10) ohm. An error of ± 2 ampere is negligible for the 1000 ampere of current measured, but the same error is not tolerable for the measurement of current of 10 ampere Thus, the quality of measurement is obtained by the *relative error*, *i.e.* the

ratio of limiting error δQ to the *true value Q* of the quantity under measurement.
The *relative error*, ε_r is given by

$$\varepsilon_r = \frac{\text{Absolute error}}{\text{True value}} = \frac{\delta Q}{Q} = \frac{\varepsilon_0}{Q}$$

When the error of an instrument is known, the effect of this error can be computed when combined with other errors. The signs of *relative errors* are given and must be preserved in the calculation.

Types of Errors
There is no measurement with perfect accuracy, but it is important to find out what accuracy actually is and how the different errors are present into the measurement.
The aim of study of errors is to find out the ways to minimize them. Errors may be introduced from different sources. Errors are usually classified as follows:
1. *Gross Errors*,
2. *Systematic Errors*, and
3. *Random Errors*.

1. *Gross Errors*: These errors are largely due to human errors in reading of instruments, incorrect adjustment and improper application of instruments, and computational mistakes. Complete elimination of such errors is probably impossible. The common gross error is the improper use of an instrument for measurement. For example, a well calibrated voltmeter can give an error in reading when connected across a high resistance circuit. The same voltmeter will give more accurate reading when connected in a low resistance circuit. It means the voltmeter has a "loading effect" on the circuit, altering the characteristics by the measurement process.

2. *Systematic Errors*: These errors are shortcomings of instruments, such as defective or worn parts, and effects of the environment on the equipment or the user.
This type of error is usually divided into two different categories:
(*i*) Instrumental Errors,
(*ii*) Environmental Errors.
(*i*) Instrumental Errors: These errors are defined as shortcomings of the instrument. Instrumental errors are inherent in measuring instruments due to their mechanical structure. For example, in the deflection type instrument friction in bearings of various moving components, irregular spring tension, stretching of the spring or reduction in tension due to improper handling or overloading of the instrument may cause incorrect readings, which will result in errors. Other instrumental errors may be due to calibration; improper zero setting, variation in the air gap, etc. Instrumental errors may be avoided by following methods:
(*a*) Selecting a suitable instrument for the particular measurement;
(*b*) Applying correction factors after determining the amount of instrumental error;
(*c*) Calibrating the instrument against a standard instrument.
(*ii*) Environmental Errors: These errors are due to external conditions surrounding the instruments which affect the measurements. The surrounding conditions may be the changes in temperature, humidity, barometric pressure, or of magnetic or electrostatic fields. Thus, a change in ambient temperature at which the instrument is used causes a change in the elastic properties of the spring in a moving-coil

mechanism and so causes an error in the reading of the instrument. To reduce the effects of external conditions surrounding the instruments the corrective measures are to be taken as follows:
(*a*) *To provide air-conditioning*,
(*b*) *Certain components in the instrument should be completely closed i.e., hermetically sealed*, and
(*c*) *To provide magnetic and electrostatic shields*.
Systematic errors can also be subdivided into:
(*a*) *Static errors*, and
(*b*) *Dynamic errors*.
Static errors are caused by limitations of the measuring device or the physical laws governing its behaviour. A static error is introduced in a micrometer when excessive pressure is applied in twisting or rotating the shaft.
Dynamic errors caused by the instrument do not respond fast enough to follow the changes in a measured variable.
3. *Random Errors*: These errors are those errors which are due to unknown causes and they occur even when all systematic errors have been taken care of. This error cannot be corrected by any method of calibration or other known methods of control. Few random errors usually occur in well-designed experiments, but they become important in high-accuracy work. For example, a voltmeter with accurately calibrated is being used in ideal environmental conditions to read voltage of an electric circuitry system. It will be found that the readings vary slightly over the period of observation. This variation cannot be corrected by any method of calibration or other known method of control and it cannot be explained without minute investigation. The only way to offset these errors is by increasing the number of readings and using statistical methods in order to obtain the best approximation of the true value of the quantity under measurement.

STATISTICAL ANALYSIS

A statistical analysis of measurement data allows an analytical determination of the uncertainty of the final test result. The result of a certain measurement method may be predicted on the basis of sample data without having detailed information on all the disturbing factors. To make statistical methods and interpretation meaningful, a large number of measurements are usually required. And also, *systematic errors* should be small compared with *residual* or *random* errors, because statistical treatment of data cannot remove a fixed bias contained in all the measurements.

Arithmetic Mean

The most probable value of a measured variable is the arithmetic mean of the number of readings taken. The best approximation will be made when the number of readings of the same quantity is very large. Theoretically, an infinite number of readings would give the best result, although in practice, only a finite number of measurements can be made.
For a data set, the mean is the sum of the observations divided by the number of observations. It identifies the central location of the data, sometimes referred to in English as the average. The mean is calculated using the following formula

$$M = \frac{\Sigma(X)}{N}$$

Where Σ = Sum of
X = Individual data points
N = Sample size (number of data points)

Mean Deviation

(i) Mean deviation for ungrouped data:

For *n* observation $x1, x2, ..., xn$, the **mean deviation about their mean** $\bar{x}$ is given by

$$\text{M.D}\ (\bar{x}) = \frac{\sum |x_i - \bar{x}|}{n}$$

Variance and Standard deviation

The mean, mode, median, and trimmed mean do a nice job in telling where the center of the data set is, but often we are interested in more. For example, a pharmaceutical engineer develops a new drug that regulates iron in the blood. Suppose she finds out that the average sugar content after taking the medication is the optimal level. This does not mean that the drug is effective. There is a possibility that half of the patients have dangerously low sugar content while the other half has dangerously high content. Instead of the drug being an effective regulator, it is a deadly poison. What the pharmacist needs is a measure of how far the data is spread apart. This is what the variance and standard deviation do. First we show the formulas for these measurements. Then we will go through the steps on how to use the formulas.

We define the *variance* to be

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x-\bar{x})^2$$

and the *standard deviation* to be

$$s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x-\bar{x})^2}$$

Variance and Standard Deviation: Step by Step

1. Calculate the mean, x.
2. Write a table that subtracts the mean from each observed value.
3. Square each of the differences.
4. Add this column.
5. Divide by n -1 where n is the number of items in the sample This is the *variance*.
6. To get the *standard deviation* we take the square root of the variance.

STANDARD TEST INPUT SIGNALS

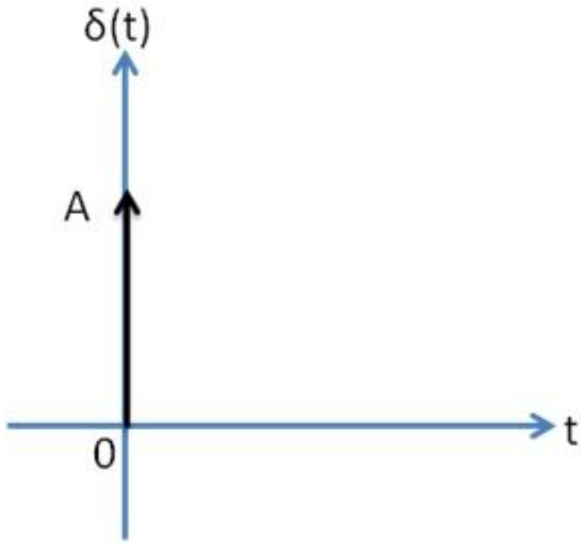
For the analysis point of view, the signals, which are most commonly used as reference inputs, are defined as standard test inputs. The performance of a system can be evaluated with respect to these test signals. Based on the information obtained the design of control system is carried out. The commonly used test signals are1. Step Input signals.2. Ramp Input Signals.3. Parabolic Input Signals. The transient response may be experimental or oscillatory in nature 4. Impulse input signal

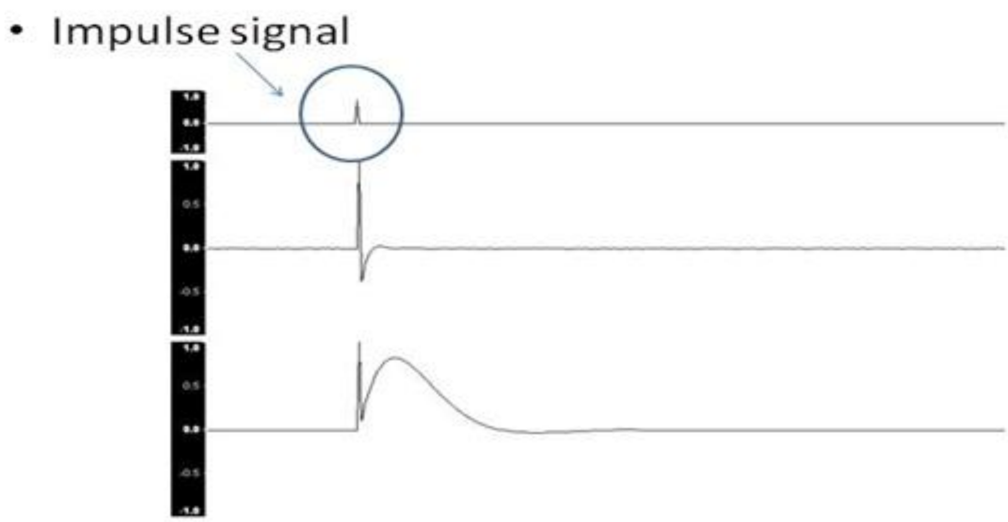
1. Impulse Signal

Impulse response in control system imitates sudden shock quality of actual input signal. Impulse is the output of system when given by small input. Impulse response emphasis on change in the system in reaction to some external change. It is the reply of the system to the direct delta input.

$$\delta(t)=\begin{cases} A & t=0 \\ 0 & t\neq 0 \end{cases}$$

When A=1 then the impulse signal is called Unit impulse signal.



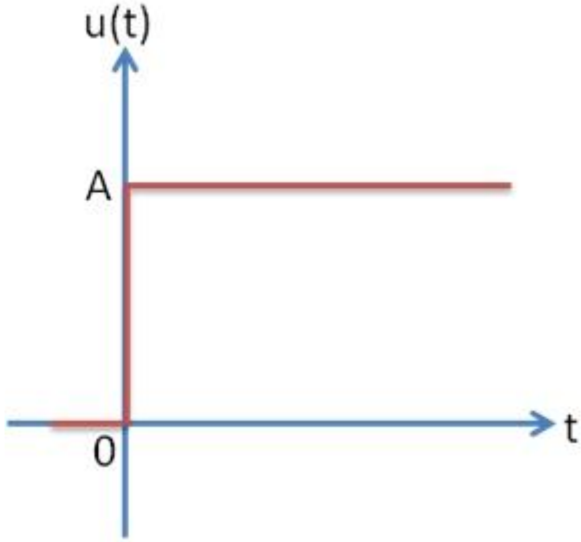


2. <u>Step Signal</u>

The step signal defines the sudden change in properties of actual signal. It is being used to see the transient response of system as it gives you the idea about how the system reply to interruption and somehow the system stability.

$$u(t) = \begin{cases} A & t \geq 0 \\ 0 & t < 0 \end{cases}$$

When A=1, the step is called unit step signal.

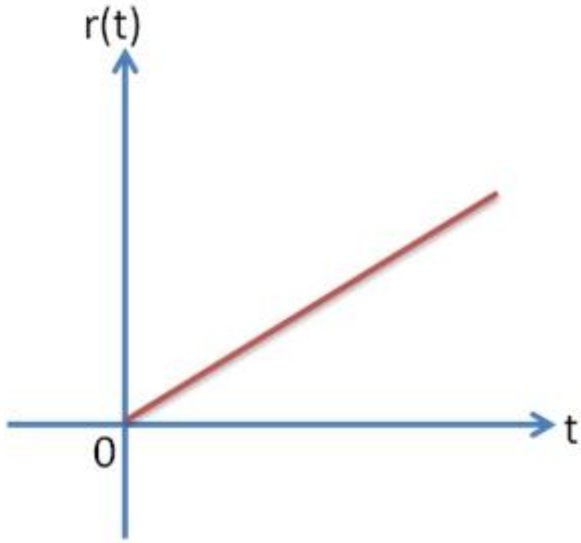


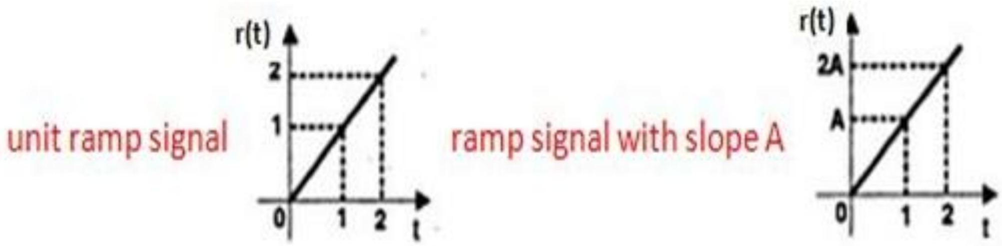
3. Ramp Signal

The ramp signal tells you the constant velocity attribute of actual input signal. It is being used to determine the behaviour of system with the velocity factor.

$$r(t) = \begin{cases} At & t \geq 0 \\ 0 & t < 0 \end{cases}$$

When A=1, ramp signal is called unit ramp signal.



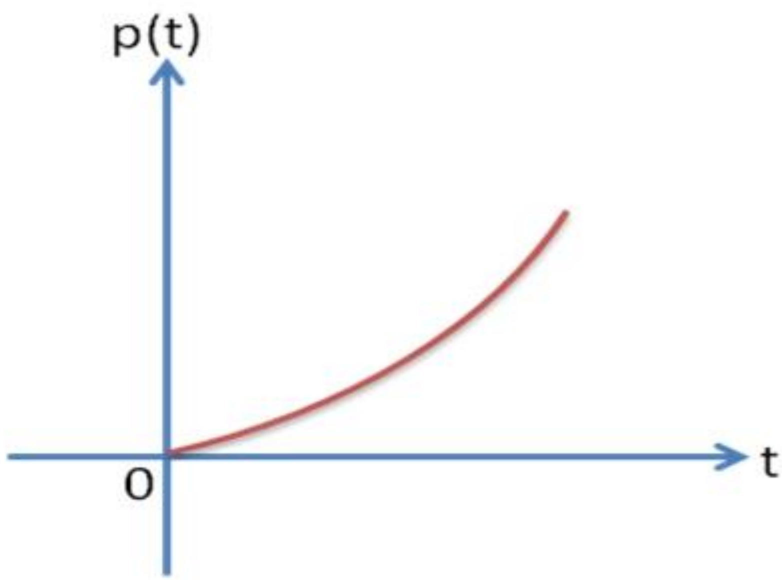


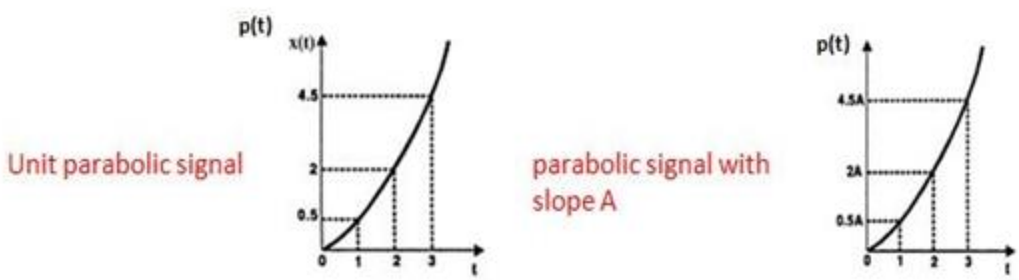
4. <u>Parabolic Signal</u>

Parabolic signal gives the constant acceleration distinction of actual input signal. It gives the idea about how the system will respond along with acceleration.

$$p(t) = \begin{cases} \dfrac{At^2}{2} & t \geq 0 \\ 0 & t < 0 \end{cases}$$

When A=1, the parabolic signal is called unit parabolic signal.





Periodic and aperiodic signals

A CT signal $x(t)$ is said to be *periodic* if it satisfies the following property:

$$x(t) = x(t + T_0),$$

at all time t and for some positive constant $T0$. The smallest positive value of $T0$ that satisfies the periodicity condition, Eq. , is referred to as the *fundamental period* of $x(t)$.
Likewise, a DT signal $x[k]$ is said to be *periodic* if it satisfies

$$x[k] = x[k + K_0]$$

at all time k and for some positive constant $K0$. The smallest positive value of $K0$ that satisfies the periodicity condition, Eq. is referred to as the fundamental period of $x[k]$. A signal that is not periodic is called an *aperiodic* or *non-periodic* signal. Figure shows examples of both periodic and aperiodic

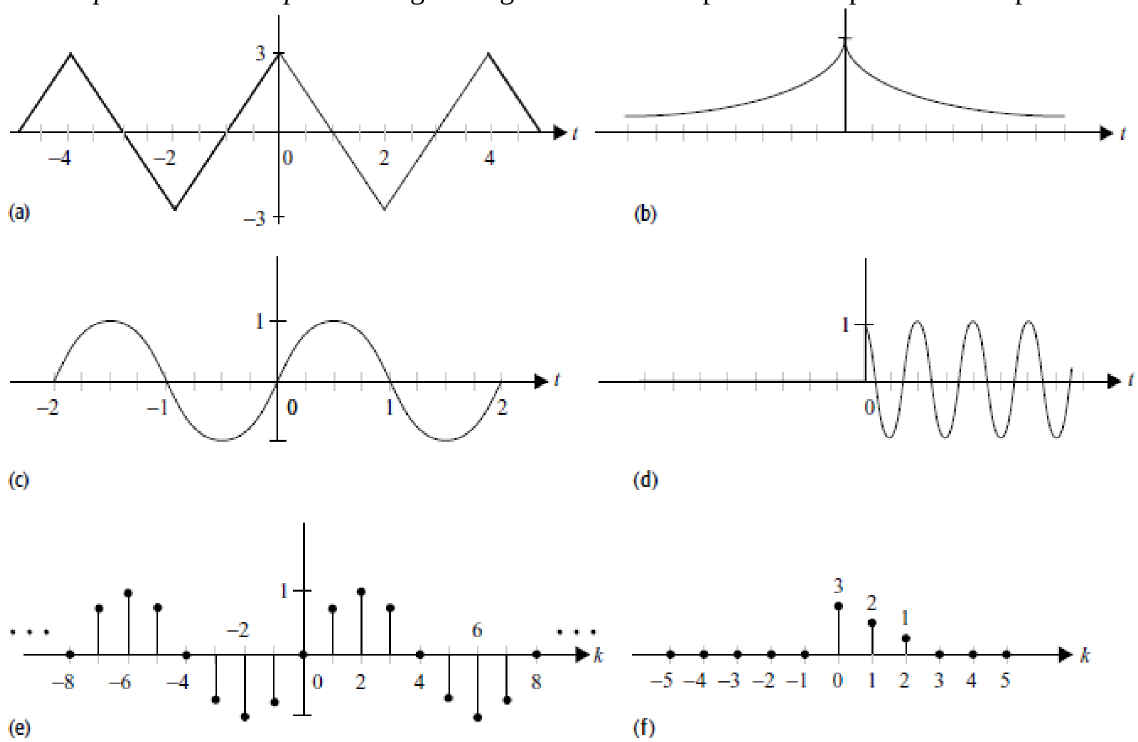


Fig. 1.6. Examples of periodic ((a), (c), and (e)) and aperiodic ((b), (d), and (f)) signals. The line plots (a) and (c) represent CT periodic signals with fundamental periods T_0 of 1.5 and 3, while the stem plot (e) represents a DT periodic signal with fundamental period $K_0 = 10$.

signals. The reciprocal of the fundamental period of a signal is called the *fundamental frequency*. Mathematically, the fundamental frequency is expressed as follows

$$f_0 = \frac{1}{T_0}, \text{ for CT signals,} \quad \text{or} \quad f_0 = \frac{1}{K_0}, \text{ for DT signals,}$$

where $T0$ and $K0$ are, respectively, the fundamental periods of the CT and DT signals. The frequency of a signal provides useful information regarding how fast the signal changes its amplitude. The unit of frequency is *cycles per second* (c/s) or *hertz* (Hz). Sometimes, we also use *radians per second* as a unit of frequency. Since there are 2_ radians (or 360◦) in one cycle, a frequency of *f*0 hertz is equivalent to 2_*f*0 radians per second. If radians per second is used as a unit of frequency, the frequency is referred to as the *angular frequency* and is given by

$$\omega_0 = \frac{2\pi}{T_0}, \text{ for CT signals,} \quad \text{or} \quad \Omega_0 = \frac{2\pi}{K_0}, \text{ for DT signals.}$$

A familiar example of a periodic signal is a sinusoidal function represented mathematically by the following expression:

$$x(t) = A\sin(\omega_0 t + \theta).$$

The sinusoidal signal $x(t)$ has a fundamental period $T_0 = 2\pi/\omega_0$ as we prove next. Substituting t by $t + T_0$ in the sinusoidal function, yields

$$x(t + T_0) = A\sin(\omega_0 t + \omega_0 T_0 + \theta).$$

Since

$$x(t) = A\sin(\omega_0 t + \theta) = A\sin(\omega_0 t + 2m\pi + \theta), \text{ for } m = 0, \pm 1, \pm 2, \ldots,$$

the above two expressions are equal iff $\omega_0 T_0 = 2m\pi$. Selecting $m = 1$, the fundamental period is given by $T_0 = 2\pi/\omega_0$.

The sinusoidal signal $x(t)$ can also be expressed as a function of a complex exponential. Using the Euler identity,

$$e^{j(\omega_0 t + \theta)} = \cos(\omega_0 t + \theta) + j\sin(\omega_0 t + \theta),$$

we observe that the sinusoidal signal $x(t)$ is the imaginary component of a complex exponential. By noting that both the imaginary and real components of an exponential function are periodic with fundamental period $T_0 = 2\pi/\omega_0$, it can be shown that the complex exponential $x(t) = \exp[j(\omega_0 t + \theta)]$ is also a periodic signal with the same fundamental period of $T_0 = 2\pi/\omega_0$.

Although all CT sinusoidals are periodic, their DT counterparts $x[k]$ = $A\sin(\Omega_0 k + \theta)$ may not always be periodic. An arbitrary DT sinusoidal sequence x[k] = $A\sin(\Omega_0 k + \theta)$ is periodic iff $\Omega_0/2\pi$ is a rational number.

Modulated signal – Sampled data pulse modulation and pulse code modulation.

Pulse modulaton

Digital Transmission is the transmittal of digital signals between two or more points in a communications system. The signals can be binary or any other form of discrete-level digital pulses. Digital pulses can not be propagated through a wireless transmission system such as earth's atmosphere or free space.

Alex H. Reeves developed the first digital transmission system in 1937 at the Paris Laboratories of AT & T for the purpose of carrying digitally encoded analog signals, such as the human voice, over metallic wire cables between telephone offices.

Advantages & disadvantages of Digital Transmission

Advantages

--Noise immunity

--Multiplexing(Time domain)

--Regeneration

--Simple to evaluate and measure

Disadvantages

--Requires more bandwidth

--Additional encoding (A/D) and decoding (D/A) circuitry

Pulse Modulation

-- *Pulse modulation* consists essentially *of* sampling analog information signals and then converting those samples into discrete pulses and transporting the pulses from a source to a destination over a physical transmission medium.

--The four predominant methods *of* pulse modulation:

1) pulse width modulation (PWM)

2) pulse position modulation (PPM)

3) pulse amplitude modulation (PAM)

4) pulse code modulation (PCM).

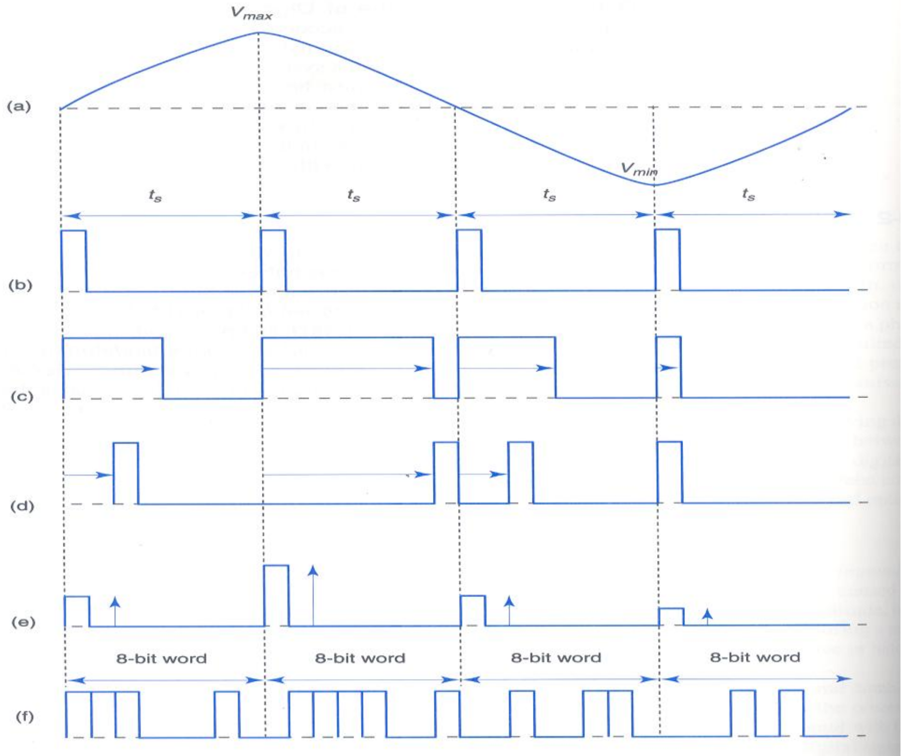


FIGURE 10-1 Pulse modulation: (a) analog signal; (b) sample pulse; (c) PWM; (d) PPM; (e) PAM; (f) PCM

<u>Pulse Width Modulation</u>

--PWM is sometimes called *pulse duration modulation* (PDM) or *pulse length modulation* (PLM), as the width (active portion *of* the duty cycle) *of* a constant amplitude pulse is varied proportional to the amplitude *of* the analog signal at the time the signal is sampled.

--The maximum analog signal amplitude produces the widest pulse, and the minimum analog signal amplitude produces the narrowest pulse. Note, however, that all pulses have the same amplitude.

<u>Pulse Position Modulation</u>

--With PPM, the position *of* a constant-width pulse within a prescribed time slot is varied according to the amplitude *of* the sample *of* the analog signal.

--The higher the amplitude *of* the sample, the farther to the right the pulse is positioned within the prescribed time slot. The highest amplitude sample produces a pulse to the far right, and the lowest amplitude sample produces a pulse to the far left.

<u>Pulse Amplitude Modulation</u>

--With PAM, the amplitude of a constant width, constant-position pulse is varied according to the amplitude *of* the sample *of* the analog signal.

--The amplitude *of* a pulse coincides with the amplitude *of* the analog signal.

--PAM waveforms resemble the original analog signal more than the waveforms for PWM or PPM.

Pulse Code Modulation

--With PCM, the analog signal is sampled and then converted to a serial n-bit binary code for transmission.

--Each code has the same number *of* bits and requires the same length *of* time for transmission

Pulse Modulation

--PAM is used as an intermediate form *of* modulation with PSK, QAM, and PCM, although it is seldom used by itself.

--PWM and PPM are used in special-purpose communications systems mainly for the military but are seldom used for commercial digital transmission systems.

--PCM is by far the most prevalent form *of* pulse modulation and will be discussed in more detail.

Pulse Code Modulation

--PCM is the preferred method *of* communications within the public switched telephone network because with PCM it is easy to combine digitized voice and digital data into a single, high-speed digital signal and propagate it over either metallic or optical fiber cables.

--With PCM, the pulses are of fixed length and fixed amplitude.

--PCM is a binary system where a pulse or lack of a pulse within a prescribed time slot represents either a logic 1 or a logic 0 condition.

--PWM, PPM, and PAM are digital but seldom binary, as a pulse does not represent a single binary digit (bit).

PCM system Block Diagram

--The band pass filter limits the frequency of the analog input signal to the standard voice-band frequency range of 300 Hz to 3000 Hz.

--The sample- and- hold circuit periodically samples the analog input signal and converts those samples to a multilevel PAM signal.

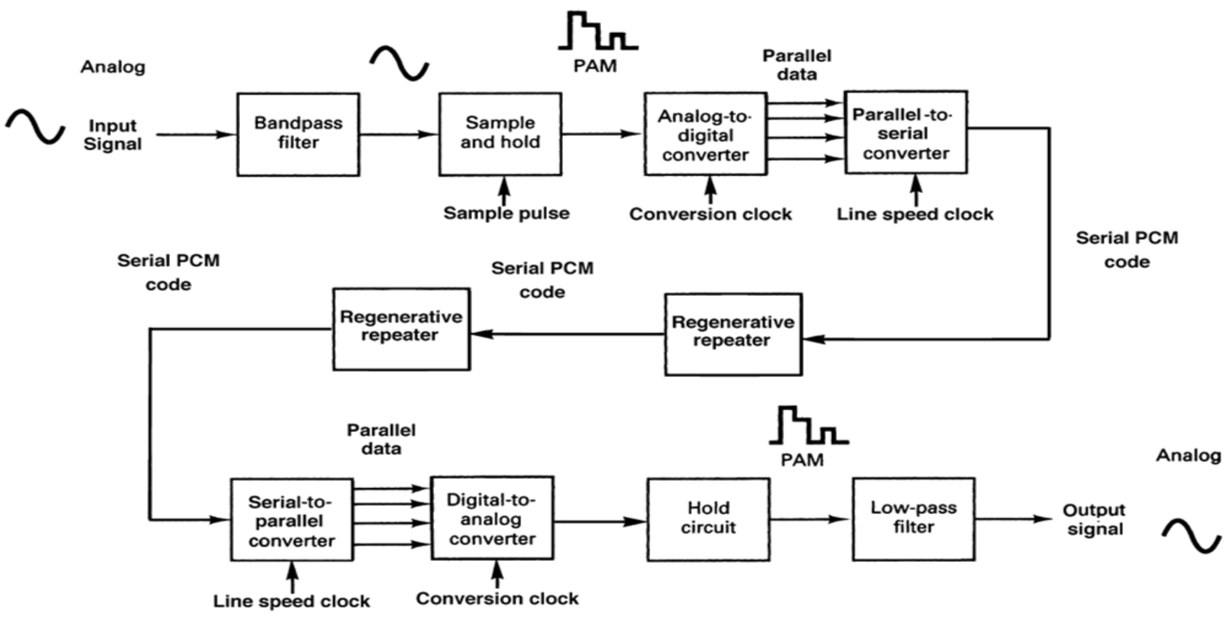
--The analog-to-digital converter (ADC) converts the PAM samples to parallel PCM codes, which are converted to serial binary data in the parallel-to-serial converter and then outputted onto the transmission linear serial digital pulses.

--The transmission line repeaters are placed at prescribed distances to regenerate the digital pulses.

--In the receiver, the serial-to-parallel converter converts serial pulses received from the transmission line to parallel PCM codes.

--The digital-to-analog converter (DAC) converts the parallel PCM codes to multilevel PAM signals.

--The hold circuit is basically a low pass filter that converts the PAM signals back to its original analog form.



The block diagram of a single-channel, simplex (one-way only) PCM system.

PCM Sampling:

--The function of a sampling circuit in a PCM transmitter is to periodically sample the continually changing analog input voltage and convert those samples to a series of constant- amplitude pulses that can more easily be converted to binary PCM code.

--A sample-and-hold circuit is a nonlinear device (mixer) with two inputs: the sampling pulse and the analog input signal.

--For the ADC to accurately convert a voltage to a binary code, the voltage must be relatively constant so that the ADC can complete the conversion before the voltage level changes. If not, the ADC would be continually attempting to follow the changes and may never stabilize on any PCM code.

--Essentially, there are two basic techniques used to perform the sampling function

1) natural sampling

2) flat-top sampling

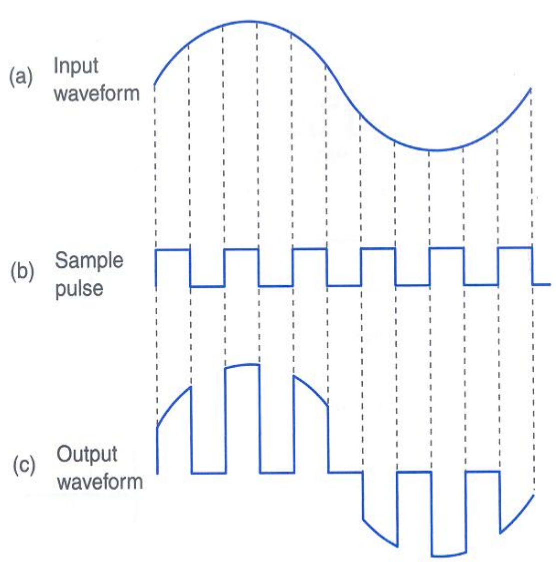


FIGURE 10-3 Natural sampling: (a) input analog signal; (b) sample pulse; (c) sampled output

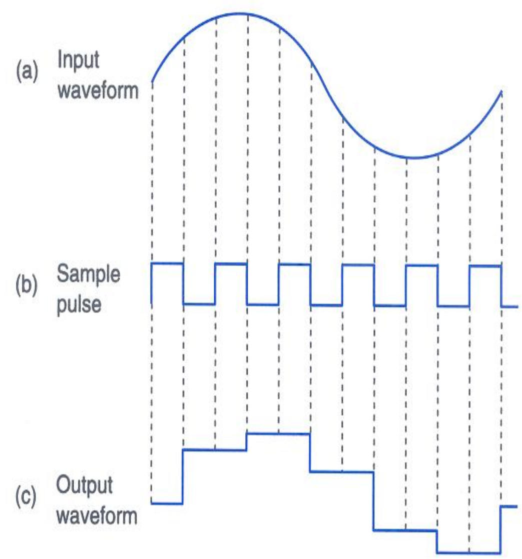


FIGURE 10-4 Flat-top sampling: (a) input analog signal; (b) sample pulse; (c) sampled output

--Natural sampling is when tops of the sample pulses retain their natural shape during the sample interval, making it difficult for an ADC to convert the sample to a PCM code.

--The most common method used for sampling voice signals in PCM systems is *flat- top sampling,* which is accomplished in a *sample-and-hold circuit.*

-- The purpose of a sample-and-hold circuit is to periodically sample the continually changing analog input voltage and convert those samples to a series of constant-amplitude PAM voltage levels.

Sampling Rate

--The *Nyquist sampling theorem* establishes the *minimum Nyquist sampling rate (fs)* that can be used for a given PCM system.

--For a sample to be reproduced accurately in a PCM receiver, each cycle of the analog input signal *(fa)* must be sampled at least twice.

--Consequently, the minimum sampling rate is equal to twice the highest audio input frequency.

--Mathematically, the minimum Nyquist sampling rate is:

$$fs \geq 2fa$$

--If *fs* is less than two times *fa* an impairment called *alias or foldover distortion* occurs.

Quantization and the Folded Binary Code:

Quantization

--*Quantization* is the process of converting an infinite number of possibilities to a finite number of conditions.

--Analog signals contain an infinite number of amplitude possibilities.

--Converting an analog signal to a PCM code with a limited number of combinations requires quantization.

Folded Binary Code

--With quantization, the total voltage range is subdivided into a smaller number of sub-ranges.

--The PCM code shown in Table 10-2 is a three-bit sign- magnitude code with eight possible combinations (four positive and four negative).

--The leftmost bit is the sign bit (1 = + and 0 = -), and the two rightmost bits represent magnitude.

-- This type of code is called *a folded binary code* because the codes on the bottom half of the table are a mirror image of the codes on the top half, except for the sign bit.

Table 10-2 Three-Bit PCM Code

8 Sub ranges

Sign	Magnitude		Decimal value	Quantization range
1	1	1	+3	+2.5 V to +3.5 V
1	1	0	+2	+1.5 V to +2.5 V
1	0	1	+1	+0.5 V to +1.5 V
1	0	0	+0	0 V to +0.5 V
0	0	0	−0	0 V to −0.5 V
0	0	1	−1	−0.5 V to −1.5 V
0	1	0	−2	−1.5 V to −2.5 V
0	1	1	−3	−2.5 V to −3.5 V

Quantization

--With a folded binary code, each voltage level has one code assigned to it except zero volts, which has two codes, 100 (+0) and 000 (-0).

--The magnitude difference between adjacent steps is called the *quantization interval* or *quantum.*

--For the code shown in Table 10-2, the quantization interval is 1 V.

--If the magnitude of the sample exceeds the highest quantization interval, *overload distortion* (also called *peak limiting)* occurs.

--Assigning PCM codes to absolute magnitudes is called quantizing.

--The magnitude of a quantum is also called the *resolution.*

--The resolution is equal to the voltage of the *minimum step size,* which is equal to the voltage of the *least significant bit* (V_{lsb}) of the PCM code.

--The smaller the magnitude of a quantum, the better (smaller) the resolution and the more accurately the quantized signal will resemble the original analog sample.

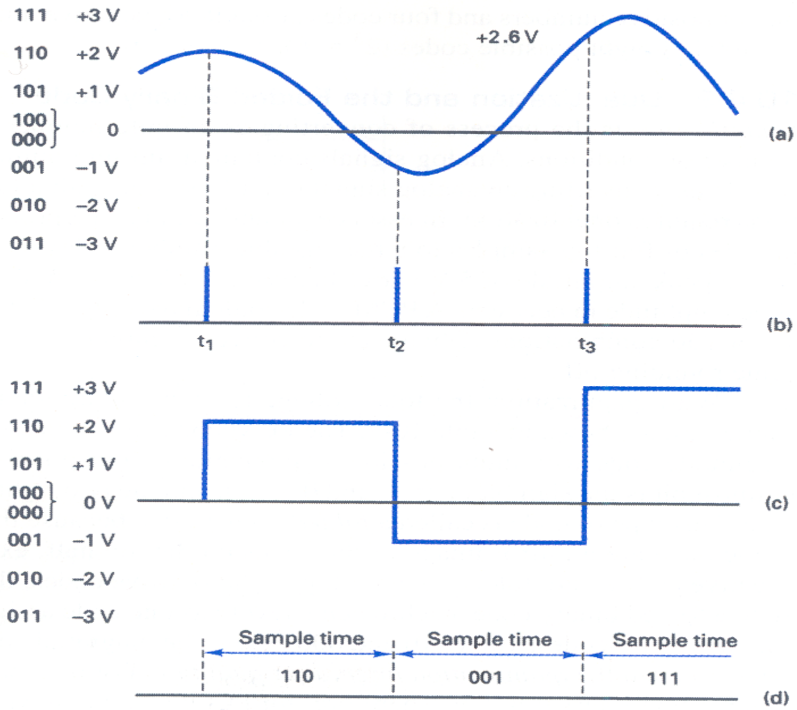


FIGURE 10-8 (a) Analog input signal; (b) sample pulse; (c) PAM signal; (d) PCM code

--For a sample, the voltage at *t3* is approximately +2.6 V. The folded PCM code is

$$\frac{\textbf{sample voltage}}{\textbf{resolution}} = \frac{\mathbf{2.6}}{\mathbf{1}} = \mathbf{2.6}$$

--There is no PCM code for +2.6; therefore, the magnitude of the sample is rounded off to the nearest valid code, which is 111, or +3 V.

--The rounding-off process results in a quantization error of 0.4 V.

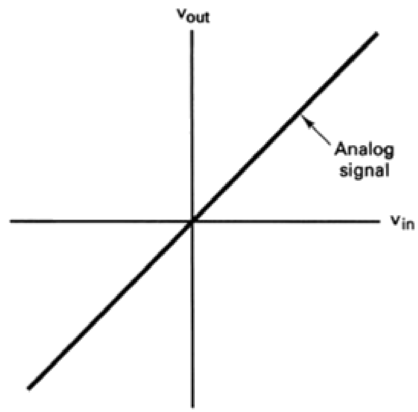
--The likelihood of a sample voltage being equal to one of the eight quantization levels is remote.

--Therefore, as shown in the figure, each sample voltage is rounded off (quantized) to the closest available level and then converted to its corresponding PCM code.

--The rounded off error is called the called the *quantization error (Qe).*

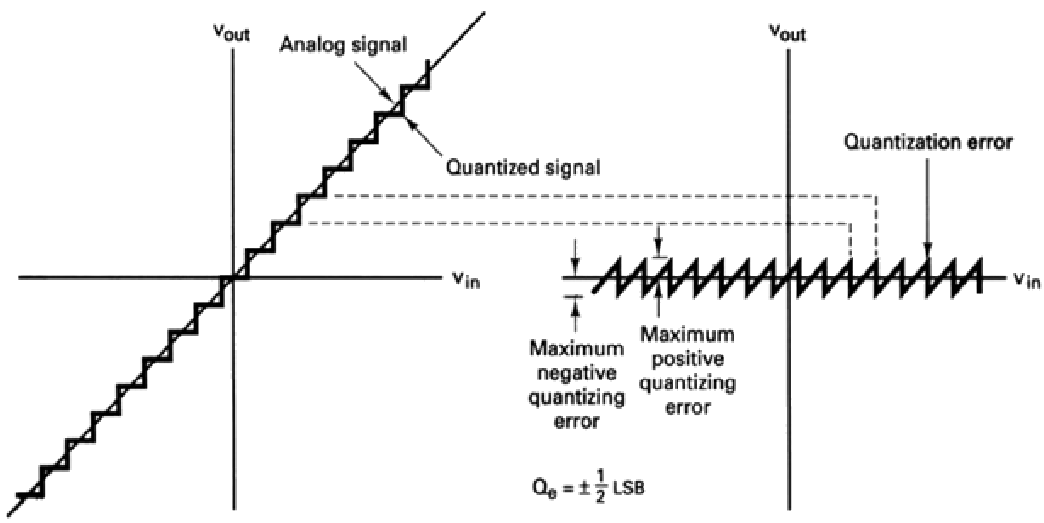
--To determine the PCM code for a particular sample voltage, simply divide the voltage by the resolution, convert the quotient to an n-bit binary code, and then add the sign bit.

Linear input-versus-output transfer curve



$$Q_e = \frac{\text{resolution}}{2}$$

Linear transfer function



Quantization

Quantization error *(Qe)*

1) For the PCM coding scheme shown in Figure 10-8, determine the quantized voltage, quantization error *(Qe)* and PCM code for the analog sample voltage of + 1.07 V.

A) To determine the quantized level, simply divide the sample voltage by resolution and then round the answer off to the nearest quantization level:

$$\frac{\mathbf{+1.07V}}{\mathbf{1V}} \mathbf{= 1.07 = 1}$$

The quantization error is the difference between the original sample voltage and the quantized level, or *Qe* = 1.07 -1 = 0.07

From Table 10-2, the PCM code for + 1 is 101.

Dynamic Range (DR): It determines the number of PCM bits transmitted per sample.

-- Dynamic range is the ratio of the largest possible magnitude to the smallest possible magnitude (other than zero) that can be decoded by the digital-to-analog converter in the receiver. Mathematically,

$$DR = \frac{V_{\max}}{V_{\min}} = \frac{V_{\max}}{\text{resolution}} = 2^n - 1 \qquad DR_{(dB)} = 20\log\left(2^n - 1\right) = 20 \log \frac{V_{max}}{V_{min}}$$

where DR = dynamic range (unitless)

V_{min} = the quantum value

V_{max} = the maximum voltage magnitude of the DACs

n = number of bits in a PCM code (excluding the sign bit)

For n > 4

$$DR = 2^n - 1 \approx 2^n$$

$$DR_{(dB)} \approx 20\log\left(2^n - 1\right) = 20n\log 2 \approx 6n$$

Signal-to-Quantization Noise Efficiency

--For linear codes, the magnitude change between any two successive codes is the same.

--Also, the magnitude of their quantization error is also same.

The maximum quantization noise is half the resolution. Therefore, the worst possible signal voltage- to-quantization noise voltage ratio (SQR) occurs when the input signal is at its minimum amplitude (101 or 001). Mathematically, the worst-case voltage SQR is

$$SQR = \frac{\text{resolution}}{Q_e} = \frac{V_{lsb}}{V_{lsb}/2} = 2 \qquad Q_e = \frac{\text{resolution}}{2}$$

For input signal minimum amplitude

SQR = minimum voltage / quantization noise

$$SQR_{(\min)} = \frac{V_{\min}}{Q_e} = \frac{\text{resolution}}{Q_e} = 2$$

For input signal maximum amplitude

SQR = maximum voltage / quantization noise

$$SQR_{(\max)} = \frac{V_{\max}}{Q_e}$$

SQR is not constant

$$SQR_{(dB)} = 10 \log \frac{v^2/R}{(q^2/12)/R}$$

where R = resistance (ohms)

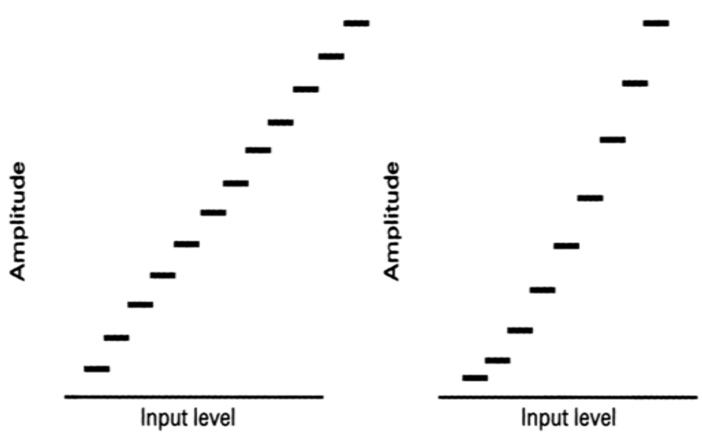
v = rms signal voltage (volts)

q = quantization interval (volts)

v^2/R = average signal power (watts)

$(q^2/12)/R$ = average quantization noise power (watts)

<u>Linear vs. Nonlinear PCM codes</u>



Linear Nonlinear

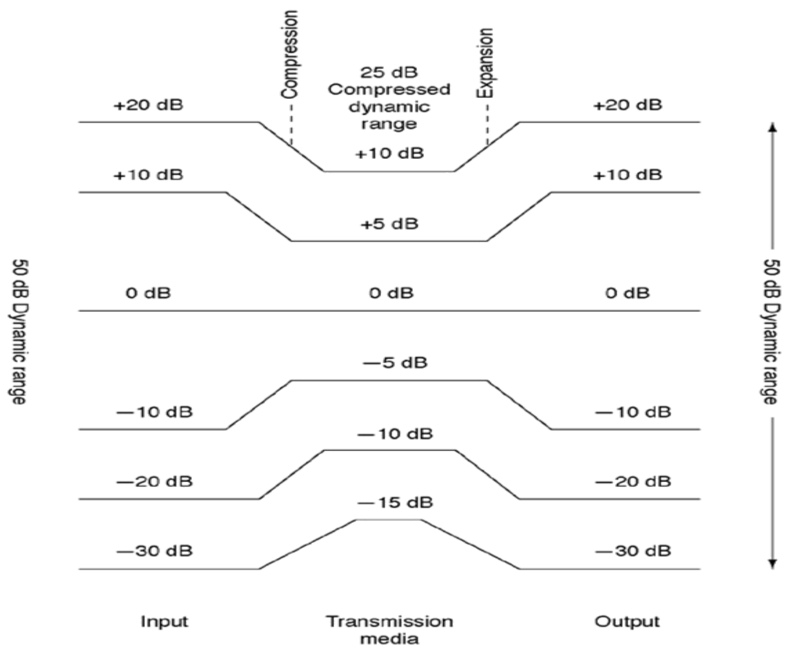
<u>Companding</u>

--Companding is the process of compressing and then expanding

--High amplitude analog signals are compressed prior to txn. and then expanded in the receiver

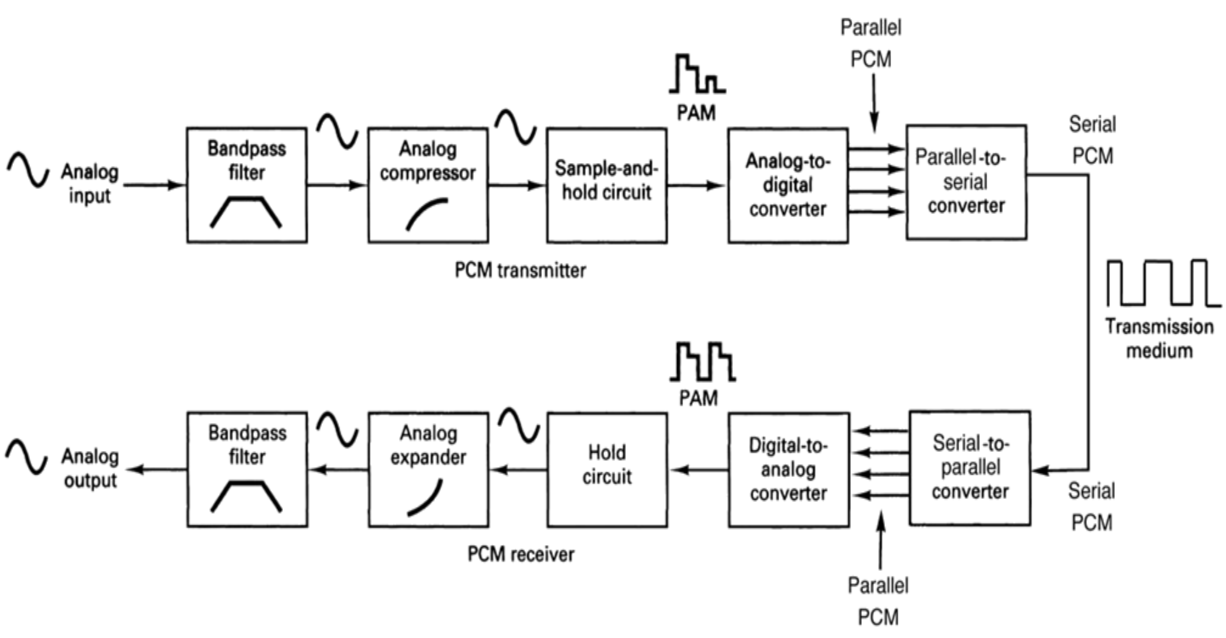
--Higher amplitude analog signals are compressed and Dynamic range is improved

--Early PCM systems used analog companding, where as modern systems use digital companding.



Basic companding process

Analog companding



PCM system with analog companding

--In the transmitter, the dynamic range of the analog signal is compressed, and then converted o a linear PCM code.

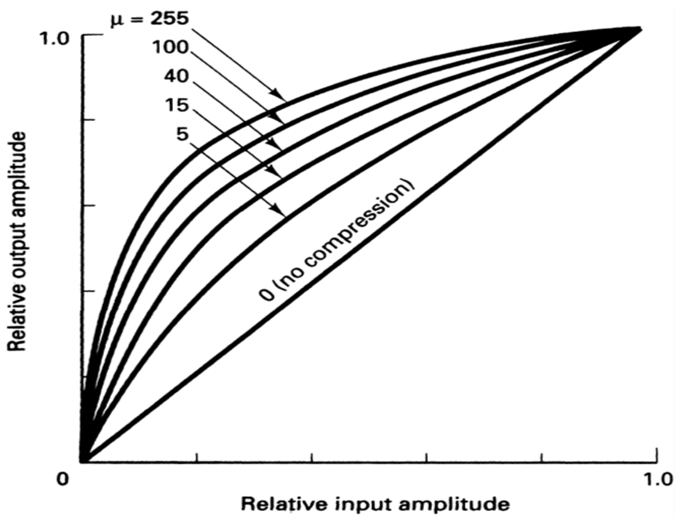
--In the receiver, the PCM code is converted to a PAM signal, filtered, and then expanded back to its original dynamic range.

-- There are two methods of analog companding currently being used that closely approximate a logarithmic function and are often called log-PCM codes.

The two methods are 1) μ-law and

2) A-law

μ-law companding



$$V_{out} = \frac{V_{\max} \ln\left(1 + \mu \, V_{in}/V_{\max}\right)}{\ln\left(1+\mu\right)}$$

where Vmax = maximum uncompressed analog input amplitude(volts)

Vin = amplitude of the input signal at a particular instant of time (volts)

μ = parameter used tio define the amount of compression (unitless)

Vout = compressed output amplitude (volts)

A-law companding

--A-law is superior to μ-law in terms of small-signal quality

--The compression characteristic is given by

$$|y| = \begin{cases} \dfrac{A|x|}{1+\log A}, & 0 \le |x| \le \dfrac{1}{A} \\ \dfrac{1+\log(A|x|)}{1+\log A}, & \dfrac{1}{A} \le |x| \le 1 \end{cases}$$

where y = Vout

x = Vin / Vmax

Digital Companding: Block diagram refer in text book.

--With digital companding, the analog signal is first sampled and converted to a linear PCM code, and then the linear code is digitally compressed.

-- In the receiver, the compressed PCM code is expanded and then decoded back to analog.

-- The most recent digitally compressed PCM systems use a 12- bit linear PCM code and an 8-bit compressed PCM code.

Digital compression error

--To calculate the percentage error introduced by digital compression

$$\%error = \frac{\text{12-bit encoded voltage - 12-bit decoded voltage}}{\text{12-bit decoded voltage}} \times 100$$

PCM Line speed

--Line speed is the data rate at which serial PCM bits are clocked out of the PCM encoder onto the transmission line. Mathematicaly,

$$\text{Line speed} = \frac{\text{samples}}{\text{second}} \times \frac{\text{bits}}{\text{sample}}$$

Delta Modulation

--*Delta modulation* uses a single-bit PCM code to achieve digital transmission of analog signals.

--With conventional PCM, each code is a binary representation of both the sign and the magnitude of a particular sample. Therefore, multiple-bit codes are required to represent the many values that the sample can be.

--With delta modulation, rather than transmit a coded representation of the sample, only a single bit is transmitted, which simply indicates whether that sample is larger or smaller than the previous sample.

--The algorithm for a delta modulation system is quite simple.

--If the current sample is smaller than the previous sample, a logic 0 is transmitted.

--If the current sample is larger than the previous sample, a logic 1 is transmitted.

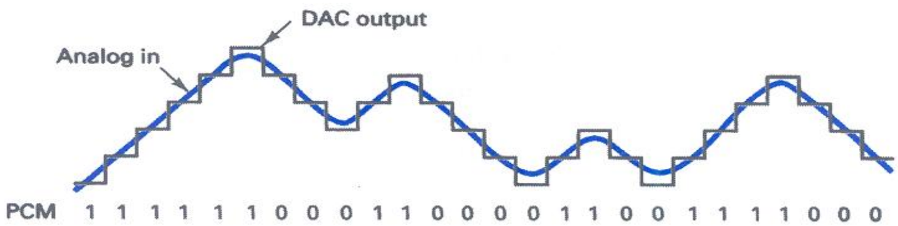


FIGURE 10-21 Ideal operation of a delta modulation encoder

<u>Differential DM</u>

--With Differential Pulse Code Modulation (DPCM), the difference in the amplitude of two successive samples is transmitted rather than the actual sample. Because the range of sample differences is typically less than the range of individual samples, fewer bits are required for DPCM than conventional PCM.