

ANALOGUE TO DIGITAL TRANSFORMATION

1. INTRODUCTION

Most industrial processes are continuous in nature and have traditionally been controlled by elements based on continuous time devices. Consequently a great deal of experience is available for the design of continuous controllers.

There are a number of disadvantages, however, associated with continuous controllers. Analogue components or elements are often fixed and cannot cope with unusual plant parameter variations, environmental changes or modifications in overall system requirements. The introduction of digital computer controllers has resulted from the recent availability of the microcomputer with its associated computational speed and power, its reduced cost and improved reliability over early digital computers. The advantages in programming flexibility to suit changes in plant or system requirements are apparent, as is the ability for several separate feedback loops or systems to share a common processor. Other advantages can include weight reduction and compactness for large systems and reduced price, particularly when reprogramming.

The objective of this Item is to examine the means by which an analogue controller element within a system can be replaced by an equivalent digital controller element such that the performance characteristics of the new hybrid system are acceptably similar to those of the original continuous system.

This Item is the second that deals with the modelling of linear continuous systems in discrete form. An earlier Data Item, No. 86037 (Reference 5), introduced the Z-transform and related its use to sampled-data systems in which a system signal is known only at discrete, equally spaced intervals of time. The relationship between the complex s -plane, used in continuous system analysis, and the z -plane, for discrete systems, was explored. The Item also considered the pulse transfer function, system behaviour between sampling instants, reconstruction of approximations to the original continuous signal after sampling and the inverse Z-transform.

This Item describes the main features of a system containing both analogue and digital components within the same loop and the analogue-to-digital and digital-to-analogue components that are necessarily part of such a system. The frequency response characteristics and the frequency content of signals are discussed prior to examining two transformation methods by which digital filter equivalents to analogue filters are designed. The methods are compared by examining the digital equivalents of 'classical' analogue filters which include simple and compound first-order lag and lead filters, second-order lag filters and standard notch filters. The methods are compared for two sampling rates with the aid of Bode plots illustrating the gain and phase variation with frequency. An example illustrates the transformation of a lag-lead controller used in an electromechanical instrument servo.

2. NOTATION

Any coherent set of units may be used in conjunction with this Item.

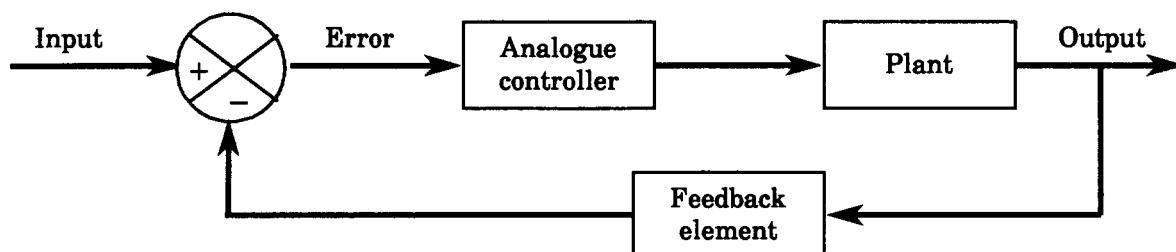
$A(\omega)$	amplitude ratio of a system element as a function of frequency
$\text{Arg}\{\dots\}$	argument of $\{\dots\}$
$a(\omega), b(\omega)$	general constants for real and imaginary parts, respectively, of a system element as a function of frequency
$a_0, a_1, a_2, \dots$	difference equation coefficients of output of an element
$b_0, b_1, b_2, \dots$	difference equation coefficients of input of an element
$C_d(z), E(z), H(z)$	Z-transforms of $c_d^*(t), e^*(t)$ and $h^*(t)$, respectively
$c(t), c_d(t), e(t), h(t), h_d(t)$	continuous functions of time
$c_d^*(t), e^*(t), h^*(t)$	sampled functions of time
$e^*(nT), h^*(nT)$	$e^*(t), h^*(t)$, respectively, for $t = nT, n = 0, 1, 2, \dots$
$\bar{c}(s), \bar{c}_d(s), \bar{e}(s), \bar{h}(s), \bar{h}_d(s)$	Laplace transforms of $c(t), c_d(t), e(t), h(t)$ and $d_d(t)$, respectively
$F(z), G(z)$	pulse transfer functions of digital system elements
$\bar{f}(s), \bar{g}(s)$	Laplace transfer functions
$\text{Im}\{\dots\}, \text{Re}\{\dots\}$	imaginary and real parts of $\{\dots\}$, respectively
i	$\sqrt{-1}$
$L\{\dots\}$	Laplace transform of $\{\dots\}$
n	general integer variable
s	Laplace transform variable
T	sampling interval
t	time
w	w-transformation variable
w'	w' -transformation variable
ZOH	zero-order hold
$Z\{\dots\}$	Z-transform of $\{\dots\}$

z	Z-transform variable
α	constant multiplier in numerator of a compound lag or lead filter
$\gamma, \delta, \zeta, \omega_n$	constants used in second-order lag element
ε	error signal
θ	repeater servo output signal
v, v'	frequency variables in w - and w' -plane, respectively
τ	time constant of lag or lead filter
$\phi(\omega)$	phase angle as a function of frequency
ω, Ω	frequencies
ω_f	folding, or Nyquist, frequency ($=\omega_s/2$)
ω_s	sampling frequency
ω_N, ζ_N	frequency and damping ratio, respectively, of standard notch filter

3. GENERAL DESCRIPTION

An analogue or continuous system is one in which all of the system variables are continuous functions of time. In contrast in a digital system one or more of the system variables, or the signals that constitute the system network, are specified only as a set of discrete values at sampling instants. Such a digital system is known as a sampled-data system.

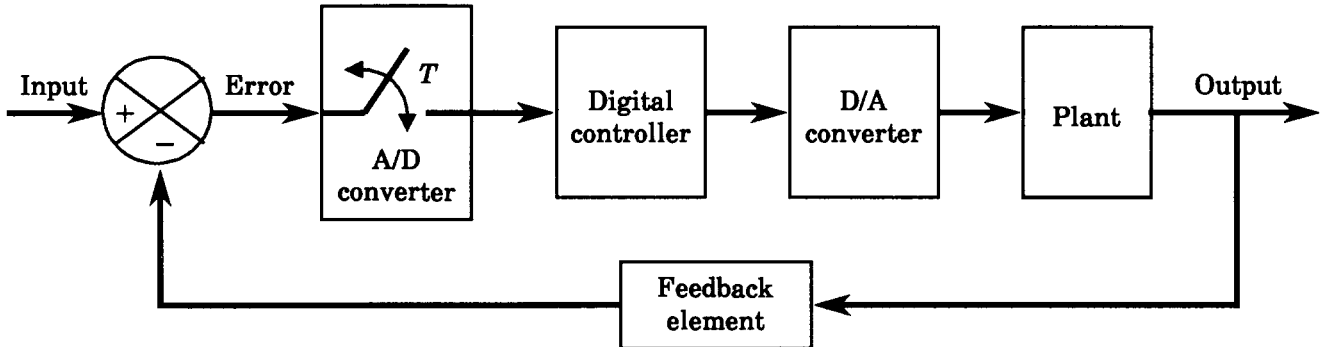
Consider a system, illustrated in Sketch 3.1, in which both the process or plant and the controlling elements are continuous in nature. The controlling elements are assumed to have been previously designed to satisfy a set of performance requirements. The fundamental problem in the application of digital techniques to the control of the assumed process is to replace some of the analogue parts of the system by their digital counterparts, such that the output characteristics of the hybrid system are acceptably similar to those of the continuous system, or at least satisfy the performance specifications required.



Sketch 3.1 Example of a continuous system

Consider now the feedback control system of Sketch 3.2 in which a digital controller is present in the forward path of the system. The controller in this sketch is simply the implementation, say, of a control

algorithm that is shaping the discrete form of the error signal. This element together with the analogue-to-digital (A/D) converter and the digital-to-analogue (D/A) converter have replaced a controlling element, the output of which forms the input to the plant.

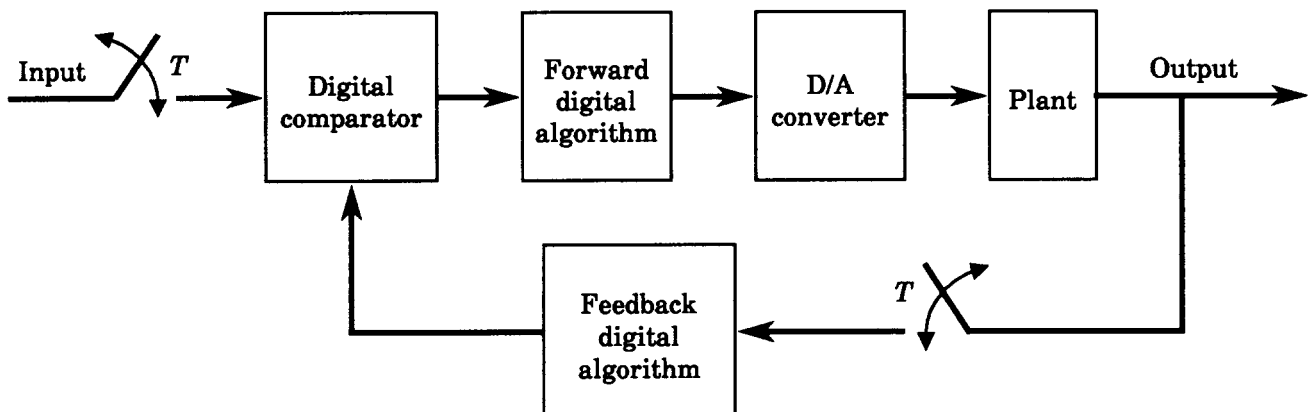


Sketch 3.2 Example of a system containing a digital controller

In the digital form of the system, Sketch 3.2, the error signal is sampled at a sampling interval, T , thus providing a discrete form of the error signal. The sampler, or A/D converter, is shown diagrammatically as a switch, where the time interval between opening and closing the switch is T .

In order that the analogue and digital signals in Sketch 3.2 can be interfaced, the digital processor is preceded by an A/D converter and followed by a D/A converter. The output from the switch is a sequence of discrete signal values or signal pulses available only when the switch is closed. The processor accesses the pulsed signal from the A/D converter and evaluates an output from its computational algorithm for each cycle of the sampler. This output is then converted by the D/A converter to a continuous signal.

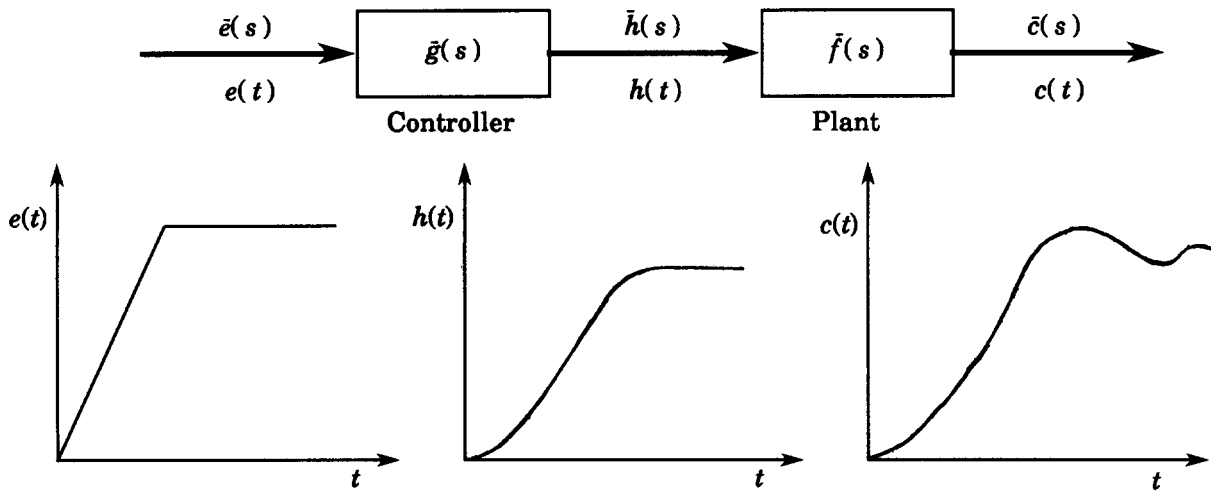
Several parts of a system can be digitised. For example, Sketch 3.3 shows a feedback system with digital input and digital control elements in both the forward and feedback paths of the circuit. In practice a single processor would most likely be used to implement the control algorithms of both the forward and feedback paths, the D/A conversions and the error signal generation. For simplicity the same sampling rate is used throughout the systems described in this Item; however, multirate systems are common.



Sketch 3.3 Example of a system with digital control of input and control elements

The notation to be used in this Item will be defined by reference to the open-loop systems of Sketches 3.4 and 3.5.

In the continuous linear system of Sketch 3.4 a continuous signal $e(t)$, whose Laplace transform is $\bar{e}(s)$, is an input to a control element whose Laplace transfer function is $\bar{g}(s)$. The continuous output of the control element, $h(t)$, has an associated Laplace transform denoted by $\bar{h}(s)$ and forms the input to the system plant. The plant in turn is denoted by the element with Laplace transform $\bar{f}(s)$ and has a continuous output signal denoted by $c(t)$, with a Laplace transform of $\bar{c}(s)$. Example input and output signals of each of the elements are illustrated in the sketch.



Sketch 3.4 Continuous open-loop linear system

In equation form the transfer functions of the system elements are expressed as

$$\bar{h}(s) / \bar{e}(s) = \bar{g}(s), \quad (3.1)$$

$$\bar{c}(s) / \bar{h}(s) = \bar{f}(s) \quad (3.2)$$

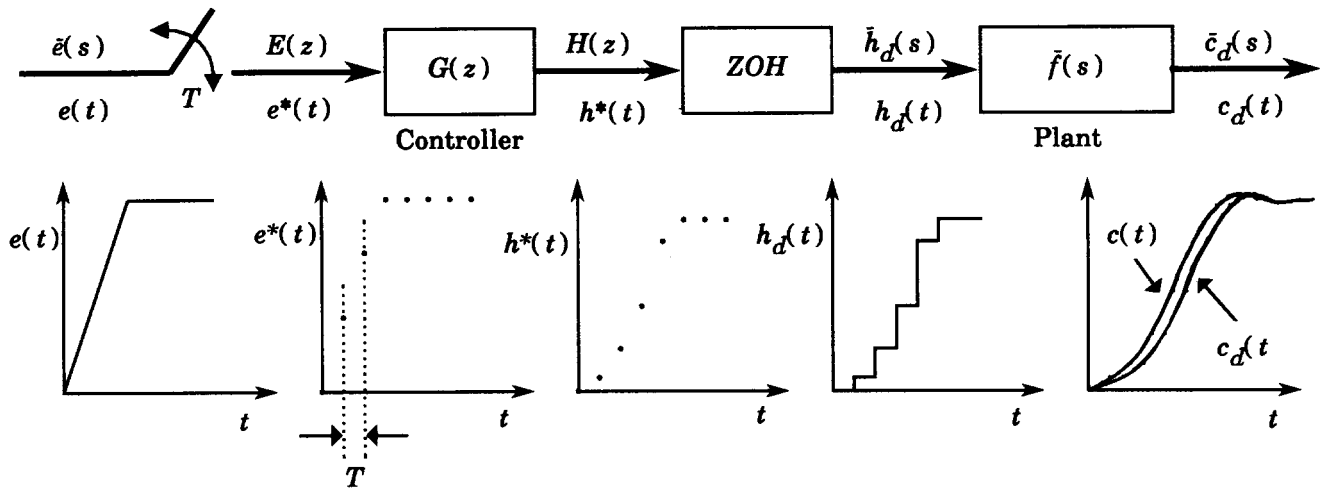
and
$$\bar{c}(s) / \bar{e}(s) = \bar{f}(s) \bar{g}(s), \quad (3.3)$$

where the Laplace transform of a function, say $c(t)$, is defined as

$$\bar{c}(s) = \int_0^{\infty} c(t) e^{-st} dt \quad (3.4)$$

The digital open-loop system of Sketch 3.5 is meant to duplicate the continuous system of Sketch 3.4 where the continuous control element $\bar{g}(s)$ has been replaced by a digital element with a Z-transform denoted by $G(z)$.

The input to the digital element is a pulsed sequence, $e^*(t)$, created by a sampler with sampling interval T from a continuous input signal, $e(t)$. The associated Laplace transform and Z-transform of the input and output of the sampler are $\bar{e}(s)$ and $E(z)$, respectively. The D/A converter is shown as a zero-order hold (ZOH) which will be described later in Section 3.4. The continuous output from the plant, $c_d(t)$, has smoothed the “staircase” or “piecewise constant” input due to the attenuation properties of the plant and is an approximation to the analogue output, $c(t)$. A delay between the analogue plant output, $c(t)$, and the plant output from the discrete system is also shown. This delay is the sum of the delays arising from the zero-order hold and the digital processor. A useful rule of thumb for estimating the delay between $c(t)$ and $c_d(t)$, ignoring any delay due to the controller, is a half of the sampling interval (see Equation (3.29)).



Sketch 3.5 Digital open-loop linear system

A pulse or z transfer function, analogous to the s transfer function for continuous systems, exists for linear sampled-data systems and relates the input and output of a pulsed sequence (see Reference 4). The pulse transfer function of the element, $G(z)$, in the open-loop system of Sketch 3.5 is given by the following expression:

$$H(z) / E(z) = G(z), \quad (3.5)$$

where the Z -transform of a typical signal, say $h^*(t)$, is defined by

$$Z\{h^*(t)\} = H(z) = \sum_{n=0}^{\infty} h^*(nT) z^{-n}. \quad (3.6)$$

The Laplace transform of a pulsed sequence (see Reference 4) is given by

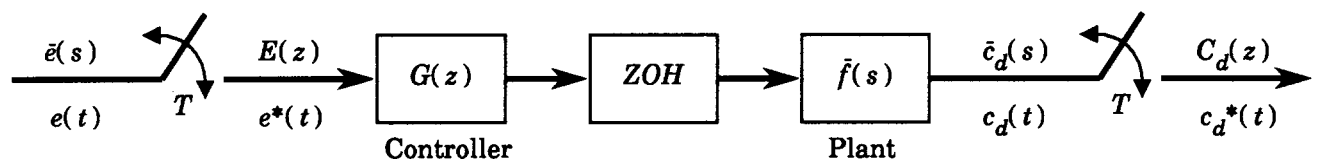
$$L\{h^*(t)\} = h^*(s) = \sum_{n=0}^{\infty} h^*(nT) e^{-nsT}. \quad (3.7)$$

Comparing Equations (3.6) and (3.7)

$$z = e^{sT}. \quad (3.8)$$

Consider the modified circuit shown in Sketch 3.6 where an additional sampler has been added to the output. The continuous signal $c_d(t)$ has been sampled to produce the pulsed sequence $c_d^*(t)$ with a Z -transform of $C_d(z)$. The output to input pulse transfer function is then

$$C_d(z) / MPE(z) = Z\{\tilde{f}(s) \text{ ZOH} \} G(z). \quad (3.9)$$

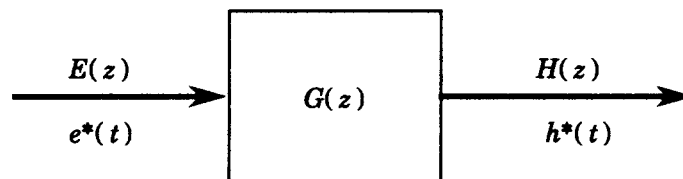


Sketch 3.6 Digital open-loop linear system with input and output samplers

3.1 Frequency Response Characteristics

Digital controllers or filters create new phenomena not encountered in the continuous domain. A main requirement of a digital filter is to match as closely as possible the performance of the original filter, *i.e.* to match the time and frequency responses of digital and analogue filters in order to satisfy the performance specification required. The accuracy with which both requirements can be met will depend on the sampling rate and the method of conversion. The frequency characteristics of digital filters are obtained in the following way.

Consider a digital element, illustrated in Sketch 3.7, with pulse transfer function $G(z)$ and input $E(z)$.



Sketch 3.7 A digital element

The output from the element may be written as

$$H(z) = G(z) E(z). \quad (3.10)$$

For a periodic sampled input of the form $e^*(nT) = \sin(n\omega T)$, for $n = 0, 1, 2, \dots$, the steady-state output, $h^*(nT)$, from a linear element is

$$h^*(nT) = G(e^{i\omega T}) e^*(nT) = A(\omega) \sin(n\omega T + \phi(\omega)), \quad (3.11)$$

where $A(\omega)$ is the amplitude ratio of the response and $\phi(\omega)$ is the phase angle. Assuming the element $G(e^{i\omega T})$ is expressed as $G(e^{i\omega T}) = a(\omega T) + ib(\omega T)$, the frequency response can be characterised by the following two quantities.

- (i) The amplitude response

$$A\{\omega\} = |G(e^{i\omega T})| = \{a^2 + b^2\}^{1/2}. \quad (3.12)$$

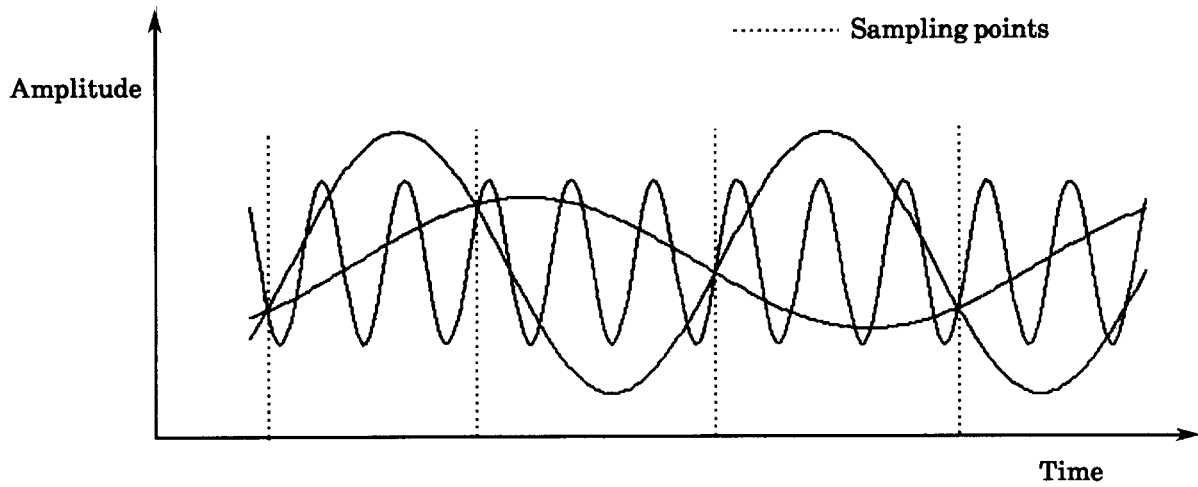
- (ii) The phase response

$$\phi\{\omega\} = \text{Arg}\{G(e^{i\omega T})\} = \tan^{-1}\{b/a\}. \quad (3.13)$$

(Note: $\phi + ve$ is a lead and $\phi - ve$ is a lag.)

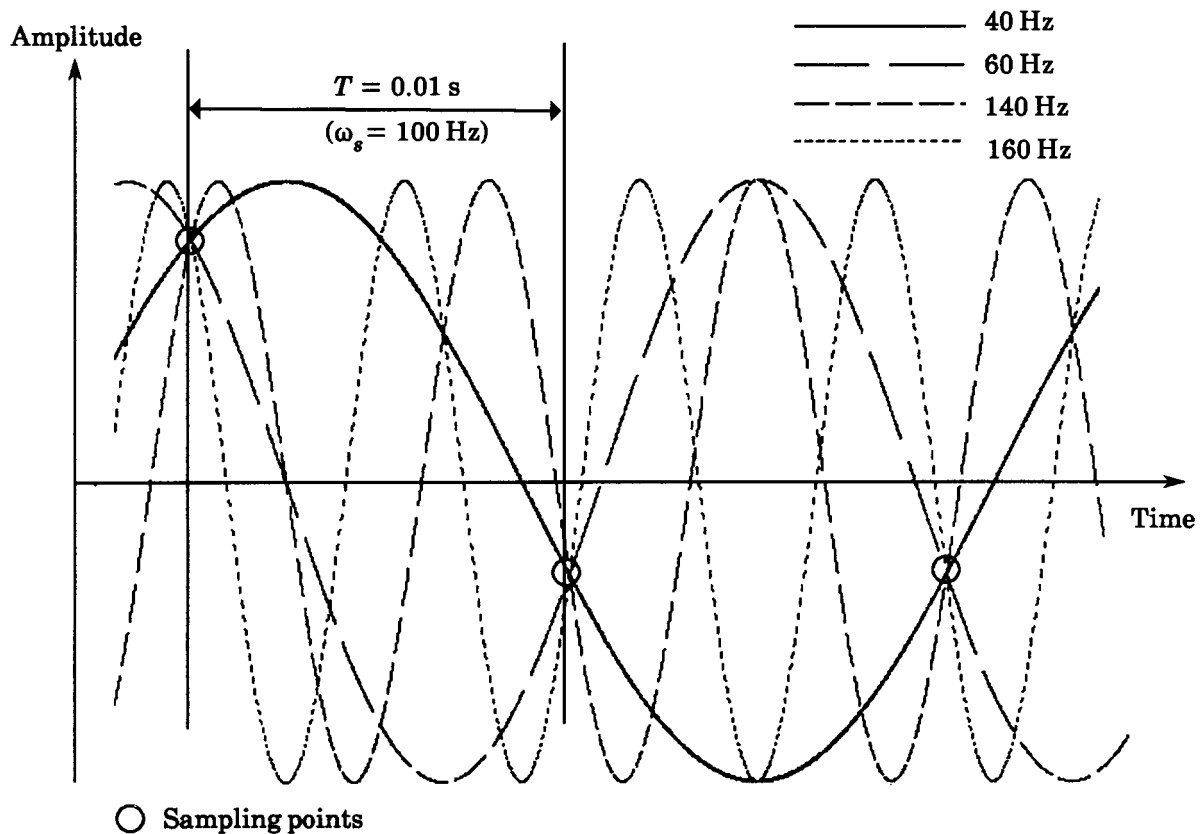
3.2 Frequency Aliasing in the Sampling Process

A particular difficulty that arises when sampling a continuous signal is *aliasing*. The nature of this problem is illustrated in Sketch 3.8 in which the amplitude of a signal is plotted against time and shows that the same set of sampled data points can describe a number of different time histories that are indistinguishable to the digital processor.



Sketch 3.8 Different time series having a single set of data points

Aliasing can be quantitatively described by considering a series of related sampled time histories, an example of which is illustrated by Sketch 3.9.



Sketch 3.9 Sinusoids illustrating aliasing

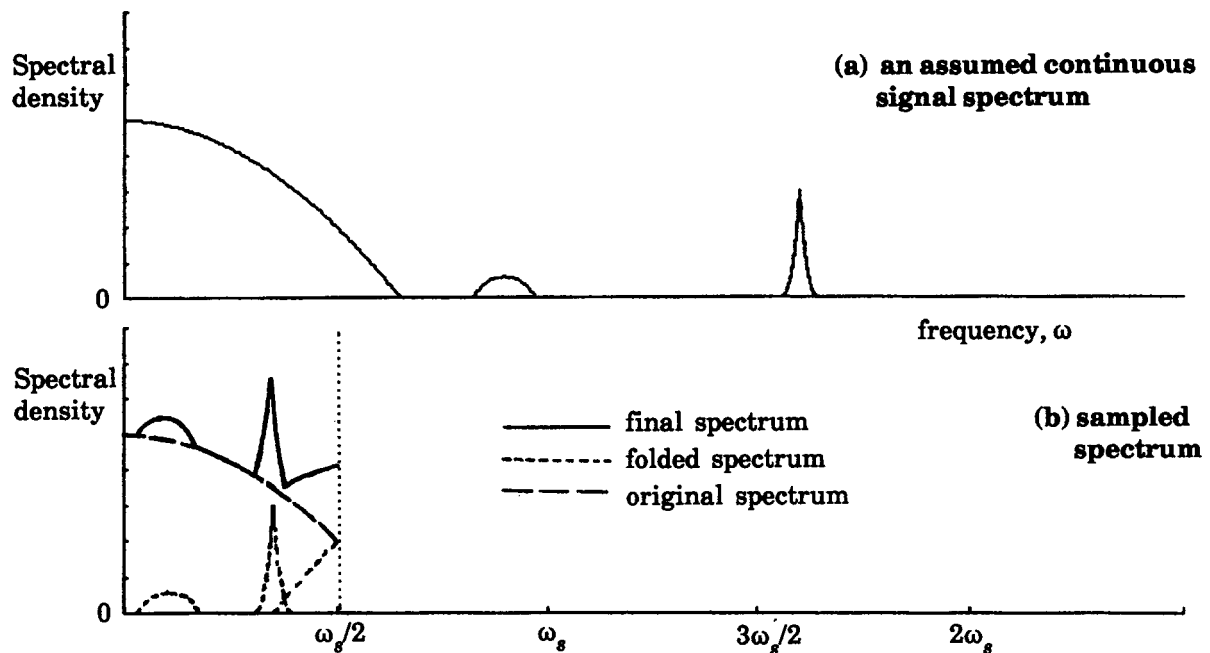
In this sketch a sinusoidal signal with a frequency equivalent to 40 Hz is shown plotted and assumed sampled with a sampling period and associated sampling frequency of 0.01 s and 100 Hz, respectively. The sampling points are shown as small circles. Also shown in this sketch are sinusoidal signals with frequencies equivalent to 60 Hz, 140 Hz and 160 Hz that are seen to pass through the same sampling points as the 40 Hz signal. These higher frequency signals, sampled with the same sampling rate as the 40 Hz signal, are indistinguishable from each other by the sampler, and from the lowest frequency signal. Consequently the

higher frequency signals are interpreted by the sampler as lower frequency signals where signals with frequencies between $\omega_s/2$ and ω_s are accepted as though they had frequencies between $\omega_s/2$ and 0, i.e. they are folded back to the lower range of frequencies. Also signals with frequencies between ω_s and $2\omega_s$, $2\omega_s$ and $3\omega_s$, etc, will all be interpreted as signals with frequencies in the range 0 to ω_s . The frequency $\omega_s/2$ ($= \pi/T$) is called the Nyquist or folding frequency and represents a limit to the sampled data, above which an unambiguous reconstruction of the original signal is not possible.

In the example of Sketch 3.9 the signals with associated frequencies of 60 Hz, 140 Hz and 160 Hz are known as the aliases of the signal with the frequency of 40 Hz. All frequencies, ω_0 , in the range 0 to $\omega_s/2$ will have a set of aliases given by the following sequence

$$\left(\frac{1}{T} - \omega_0\right), \left(\frac{1}{T} + \omega_0\right), \left(\frac{2}{T} - \omega_0\right), \left(\frac{2}{T} + \omega_0\right), \dots, \left(\frac{n}{T} \pm \omega_0\right), \dots$$

It is essential when sampling that aliasing is avoided. Aliasing can be further illustrated by considering a continuous signal with a gain/frequency relationship given by Sketch 3.10(a). This signal is then sampled at a sampling frequency of ω_s . Sketch 3.10(b) shows the sampled spectrum where frequencies exceeding the Nyquist frequency that were present in the original continuous signal have been irretrievably folded into the low frequency signals.



Sketch 3.10 Illustration of folding of frequencies

This same phenomenon can be looked at in a slightly different way by considering the mapping from the s-plane to the z-plane using Equation (3.8), or

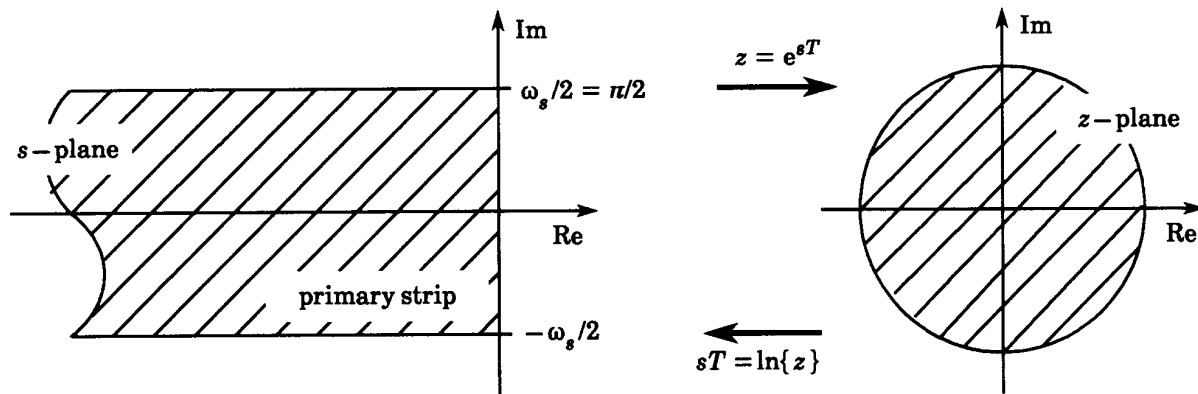
$$z = e^{sT}.$$

Replacing s by $i\omega$ to obtain the frequency response the following relationship is obtained

$$z = e^{i\omega T} = \cos(\omega T) + i \sin(\omega T) \quad (3.14)$$

and so $G(z) = G(e^{i\omega T}) = G(e^{i(\omega + 2\pi n)T}), n = 0, 1, 2, 3, \dots$ (3.15)

which shows that the frequency range is completely specified by the range of ωT of 2π , the usual band being the range $-\pi \leq \omega T \leq \pi$. Graphically this is shown in Sketch 3.11 where poles on the s -plane, whose frequencies differ by integer multiples of the sampling frequency, are transformed to the same position on the z -plane. Other mapping relationships between poles in the s -plane and their equivalents in the z -plane are given in Reference 4. The fundamental frequency poles are located in what is called the primary strip and relate to the frequencies for $n = 0$.



Sketch 3.11 Mapping of the s -plane to the z -plane

3.3 The w -Plane and w' -Plane Variable Transformations

Construction of the approximate Bode diagram to obtain the frequency response characteristics of a factored transfer function can be achieved using the sum of the contributions made by each factor. A quick method of obtaining an approximation to the gain of a transfer function, particularly useful for sketching purposes, can be obtained by using a straight line approximation to each factor. These approximations are particularly useful to describe the asymptotic behaviour of the frequency response curve as $\omega \rightarrow \infty$. (These construction methods are fully described in Appendix B of Data Item No. 74021, Reference 3. Plots of amplitude ratio and phase angle versus frequency of the steady-state response of a first- and a second-order system to a sinusoidal response are given in Data Item No. 69005, Reference 1.)

Analysis of basic digital elements in terms of the conventional frequency response variable, ω , does not yield frequency responses that are as amenable to simple asymptotic construction as the responses of analogue networks. For this reason it is useful to specify the frequency response characteristics as a function of a modified frequency variable so that digital responses will obey the same set of asymptotic construction rules as analogue networks obey in terms of ω .

The w -plane transformation is a bilinear mapping of the z -plane onto the w -plane in which the interior of the unit circle in the z -plane is mapped into the complete left half of the w -plane and points on the unit circle itself map onto the imaginary axis. The transformation is given by

$$w = \frac{z-1}{z+1}, \quad (3.16a)$$

$$\text{or} \quad z = \frac{1+w}{1-w}. \quad (3.16b)$$

Alternatively, comparing the s -plane and the w -plane,

$$w = \frac{e^{sT}-1}{e^{sT}+1} = \tanh\left(\frac{sT}{2}\right), \quad (3.17)$$

in which the primary strip of the s -plane is transformed into the left half of the w -plane. Setting $s = i\omega$

Equation (3.17) becomes

$$w = \tanh\left(\frac{i\omega T}{2}\right) = i \tan\left(\frac{\omega T}{2}\right), \quad (3.18a)$$

or $w = iv$, (3.18b)

where v is the new frequency variable in the w -plane, related to ω , the frequency in the s -plane, by

$$v = \tan\left(\frac{\omega T}{2}\right), \quad (3.19)$$

where as $\omega \rightarrow \pi$, $v \rightarrow \infty$. Thus the frequency response of $G\{z(\omega)\}$ is $G(e^{i\omega T})$ and the frequency response of $G\{z(v)\}$ is $G(iv)$ which provides the means of calculating the frequency response of a discrete network by transforming into the w -plane and using existing s -plane techniques to obtain the frequency. However, the frequency scale has been modified to the fictitious frequency v and must be interpreted in terms of ω with the use of Equation (3.19).

Another bilinear mapping using a scaling factor $T/2$ is known as the w' -transformation and is defined as

$$w' = \frac{2}{T} \left(\frac{z-1}{z+1} \right) \quad (3.20a)$$

and so $z = \frac{1 + \frac{T}{2}w'}{1 - \frac{T}{2}w'}$. (3.20b)

The associated frequency variable is given by

$$v' = \frac{2}{T} \tan\left(\omega \frac{T}{2}\right) \quad (3.21)$$

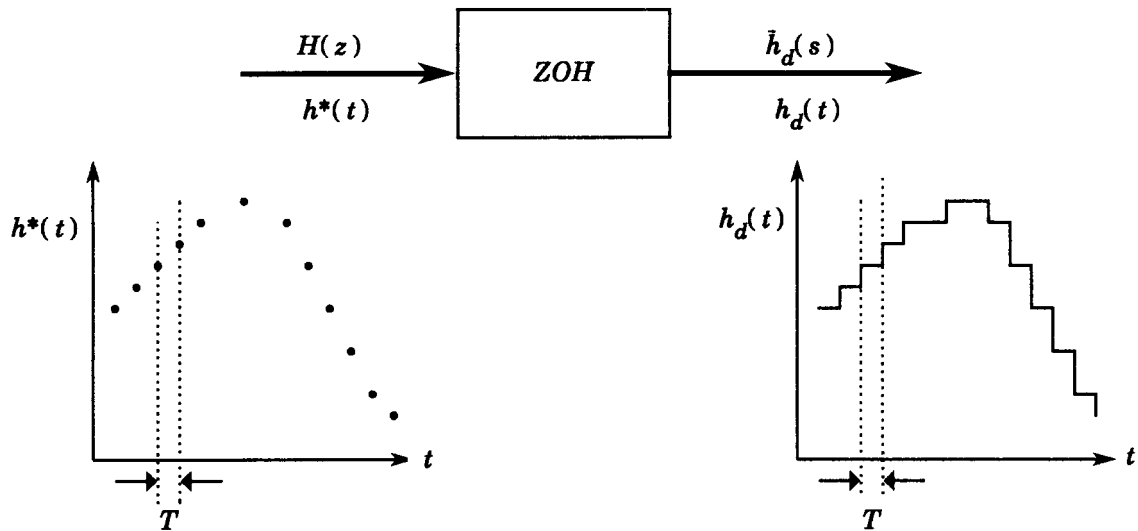
and is seen to have an advantage over the w -transformation since, for high sampling rates and at low frequency, $v' \approx \omega$, *i.e.* when $\omega T/2$ is small v' is approximately equal to ω .

3.4 Digital-to-Analogue Conversion

The error when reconstructing a continuous signal from a sampled signal is in the interpolation between samples. The sampled version of an analogue signal is a sequence of pulses which have frequency components not only between 0 and $\omega_s/2$ corresponding to the true signal, but also components between $\omega_s/2$ and ∞ due to the pulse-like nature of the sampled signal. In order to reconstruct the original signal from the sampled version, these higher frequency components must be suppressed by means of an ideal filter. Such a filter would have rectangular characteristics, *i.e.* a gain characteristic that is unity up to the folding frequency and zero thereafter and a phase characteristic that is zero throughout. Such a filter, however, is physically unrealisable and approximate reconstruction of the original signal is carried out by a *hold* filter.

A hold filter is a smoothing filter which reconstructs a continuous signal from a sampled signal by fitting

an m th-order polynomial through the $m + 1$ previous points and extrapolating forward in time along this polynomial for each period, T . The polynomial is then called an m th-order (polynomial) hold. Two practical hold circuits are the zero-order hold (ZOH) and the first-order hold, although the former is mostly used due to its simplicity. The zero-order hold is illustrated in Sketch 3.12 where the output is a stepped version of the original signal with the output value held constant between samples at the level previously reached. The zero-order hold is a converter element accepting as an input a pulsed sequence of signals and delivering a continuous stepped output.



Sketch 3.12 The zero-order hold

The zero-order hold is usually defined in terms of a Laplace transfer function given by

$$ZOH = \frac{\bar{h}_d(s)}{\bar{h}^*(s)}, \quad (3.22)$$

where $\bar{h}^*(s) = K\{h^*(t)\}$ as given by Equation (3.7).

A sample and hold element over one time step can be thought of as the sum of two step functions of opposite sign occurring at nT and $(n + 1)T$ which will have a transfer function given by

$$\frac{e^{-nTs}}{s} - \frac{e^{-(n+1)Ts}}{s}.$$

The output from the zero-order hold can now be described by the following equation

$$\bar{h}_d(s) = \sum_{n=0}^{\infty} h^*(nT) \left\{ \frac{e^{-nTs}}{s} - \frac{e^{-(n+1)Ts}}{s} \right\}, \quad (3.23a)$$

or

$$\bar{h}_d(s) = \left\{ \sum_{n=0}^{\infty} h^*(nT) e^{-nTs} \right\} \left\{ \frac{1 - e^{-Ts}}{s} \right\}, \quad (3.23b)$$

which may be compared with

$$\bar{h}_d(s) = \bar{h}^*(s) ZOH = \left\{ \sum_{n=0}^{\infty} h^*(nT) e^{-nTs} \right\} ZOH \quad (3.24)$$

and so the zero-order hold transfer function is given by

$$ZOH = \frac{1 - e^{-sT}}{s} . \quad (3.25)$$

Incorporating a sampler gain of $1/T$ in order that the ZOH has a low frequency gain of unity yields

$$ZOH = \frac{1 - e^{-sT}}{sT} . \quad (3.26)$$

The frequency response of this hold is obtained by substituting $s = i\omega$ into this equation, which gives

$$ZOH = \frac{1 - e^{-i\omega T}}{i\omega T} \quad (3.27a)$$

or
$$ZOH = \frac{\sin(\omega T/2)}{\omega T/2} e^{-i\omega T/2} , \quad (3.27b)$$

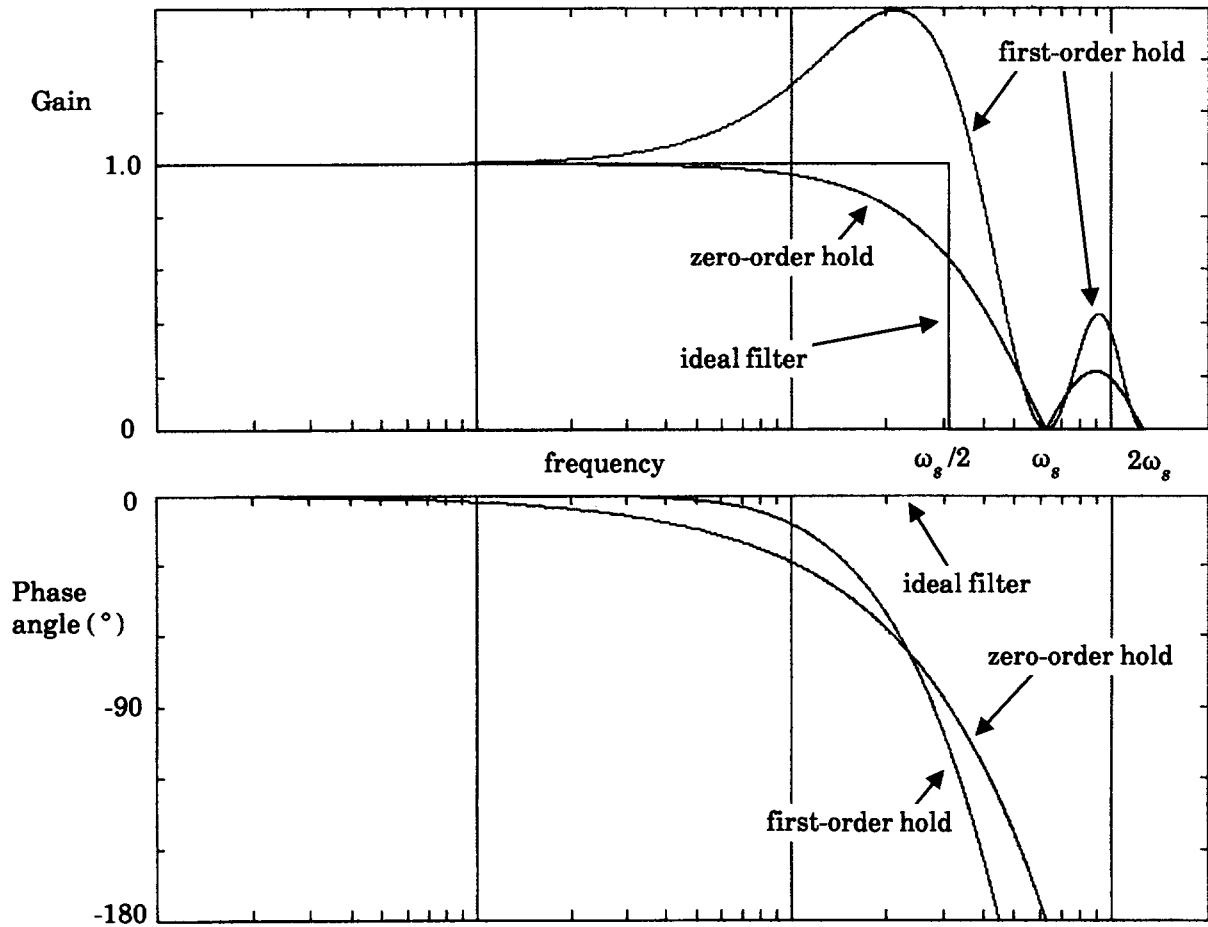
in which the gain and phase, respectively, are

$$|ZOH| = \frac{\sin(\omega T/2)}{\omega T/2} \quad (3.28)$$

and
$$\text{Arg}\{ZOH\} = -\omega T/2 , \quad (3.29)$$

which is seen to be a time delay of $T/2$ at a frequency of ω . A first-order hold replaces the rectangular stepform by a ramp whose slope is dependent on a simple backward difference obtained from the previous pair of sample values.

Sketch 3.13 shows the gain and phase characteristics of the zero-order and first-order holds and compares them with the ideal 'rectangular' filter. The curves are shown plotted on a logarithmic scale to emphasise the lower frequency variation. Comparing the hold filters with the ideal filter, significant error in signal reconstruction will occur unless the sampling frequency is high. For low frequencies, below about $\omega_s/4$ the first-order filter is to be preferred to the zero-order filter since it introduces less phase lag. However, the additional complexity is rarely justified in terms of improved performance.



Sketch 3.13 Amplitude and phase characteristics of hold filters

3.5 Physical Realisability

Consider a digital element with a pulsed input sequence of $e^*(nT)$ and a pulsed output sequence of $h^*(nT)$. The pulse transfer function of this element is given by $G(z) = H(z)/E(z)$ as previously defined in Equation (3.5), may be written as

$$G(z) = \frac{b_0 + b_1 z^{-1} + b_2 z^{-2} + \dots + b_n z^{-n}}{a_0 + a_1 z^{-1} + a_2 z^{-2} + \dots + a_n z^{-n}} = \frac{H(z)}{E(z)}, \quad (3.30)$$

where b_n is equivalent to $b^*(nT)$.

It may be shown, however, that z^{-1} is a pure time delay as given by the time shift property of the Z-transform (see Reference 4), and so the inverse Z-transform of the product $z^{-j}E(z)$ is

$$Z^{-1}\{z^{-j}E(z)\} = e^*\{(n-j)T\} = e_{n-j}. \quad (3.31)$$

This inverse Z-transform applied to Equation (3.30) gives the following difference equation

$$a_0 h_n + a_1 h_{n-1} + a_2 h_{n-2} + \dots = b_0 e_n + b_1 e_{n-1} + b_2 e_{n-2} + \dots \quad (3.32a)$$

or
$$h_n = (b_0 e_n + b_1 e_{n-1} + b_2 e_{n-2} + \dots) / a_0 - (a_1 h_{n-1} + a_2 h_{n-2} + \dots) / a_0 \quad (3.32b)$$

This difference equation is the way in which digital elements are implemented on a computer where the output h_n depends on past outputs h_{n-1} , h_{n-2} , .. and inputs e_n , e_{n-1} , e_{n-2} . For this network to be realisable the coefficient a_0 must be non zero. In practice a_0 is often chosen to be unity by suitable scaling.

4. ANALOGUE TO DIGITAL TRANSFORMATION

Several methods are available for obtaining an approximate z-plane transfer function from an s-plane transfer function representation, *i.e.* an approximation of the algebraic relationship $z = e^{sT}$, as given by Equation (3.8). A transformation formed using this algebraic substitution would result in an exact z-plane frequency response representation of s-plane characteristics but would not provide a time implementation. The most important of these approximate methods are the following:

- (i) impulse invariance transformation or Z-transform,
 - (ii) step invariance transformation,
 - (iii) mapping of differentials,
 - (iv) bilinear transformation or Tustin's approximation,
 - (v) bilinear transformation and frequency prewarping
- and (vi) matched Z-transform.

The methods can be sorted into three types; Methods (i) and (ii) are based on the matching of specific transient responses, Methods (iii), (iv) and (v) are approximate integration methods and Method (vi) maps poles and zeros between the s-plane and the z-plane. Another method by which continuous systems can be represented in discrete form is the state space method. This Item, however, will only consider Methods (i) and (iv), which will be described in general terms and then compared by examination of the frequency responses of standard elements.

4.1 Impulse Invariance Transformation

This method, also known simply as the Z-transform method, matches the impulse response of an analogue element to that of a digital element at the sampling instants. An algebraic substitution of the relation $s = \ln(z) / T$ will not implement this transformation.

Consider an element with a Laplace transform of $\bar{g}(s)$. The Z-transform of this element is

$$G(z) = Z\{\bar{g}(s)\}, \quad (4.1)$$

where the Z-transform is defined by Equation (3.6). Often it is necessary after transformation to match the gains of the analogue and digital elements at a required frequency. For example in the case of a low pass filter the gains would be matched at zero frequency. Usually this transformation is accomplished with the use of standard tables. One such source which describes the Z-transform process and includes a table of transforms for some commonly used discrete time functions is presented in Item No. 86037 (Reference 5). Although standard tables of transformations can be used, a complicated transfer function will probably not be tabulated and partial fraction manipulation will be required. Other properties of this transformation are that the digital element will have the same impulse response as the analogue element but will not preserve the frequency response. If the analogue transfer function is stable then the digitally transformed element will also be stable.