Signals and Systems

Lecture 9

Systems Characterized by Linear Constant-coefficient Difference Equations

Sampling

Systems Characterized by
Linear Constant-coefficient Difference Equations

Systems Characterized by Linear Constant-coefficient Difference Equations

A general linear constant-coefficient difference equation for an LTI system with input $x[n]$ and output $y[n]$ is of the form

$$\sum_{k=0}^{N} a_k y[n-k] = \sum_{k=0}^{M} b_k x[n-k]$$

Let $X(e^{j\omega})$, $Y(e^{j\omega})$, and $H(e^{j\omega})$ denote the Fourier transforms of the input $x[n]$, output $y[n]$, and impulse response $h[n]$, respectively.

Using convolution property

$$H(e^{j\omega}) = \frac{Y(e^{j\omega})}{X(e^{j\omega})}$$

Systems Characterized by Linear Constant-coefficient Difference Equations

$$\sum_{k=0}^{N} a_k y[n-k] = \sum_{k=0}^{M} b_k x[n-k]$$

Applying the Fourier transform to both sides of equation and using the linearity and time-shifting properties, we obtain the expression

$$\sum_{k=0}^{N} a_k e^{-jk\omega} Y(e^{j\omega}) = \sum_{k=0}^{M} b_k e^{-jk\omega} X(e^{j\omega}),$$

$$H(e^{j\omega}) = \frac{Y(e^{j\omega})}{X(e^{j\omega})} = \frac{\sum_{k=0}^{M} b_k e^{-jk\omega}}{\sum_{k=0}^{N} a_k e^{-jk\omega}}$$

$H(e^{j\omega})$ is a ratio of polynomials, but in discrete time the polynomials are in the variable $e^{-j\omega}$.

Example 1

Consider the causal LTI system that is characterized by the difference equation

$$y[n] - ay[n-1] = x[n] \qquad |a| < 1$$

The frequency response of this system is

$$H(e^{j\omega}) = \frac{Y(e^{j\omega})}{X(e^{j\omega})} = \frac{\sum_{k=0}^{M} b_k e^{-jk\omega}}{\sum_{k=0}^{N} a_k e^{-jk\omega}}$$

$$H(e^{j\omega}) = \frac{1}{1 - ae^{-j\omega}}$$

The impulse response of the system is the inverse Fourier transfrom of $H(e^{j\omega})$

$$h[n] = a^n u[n]$$

Example 2

Consider the causal LTI system that is characterized by the difference equation

$$y[n] - \frac{3}{4}y[n-1] + \frac{1}{8}y[n-2] = 2x[n]$$

The frequency response of this system is

$$H(e^{j\omega}) = \frac{Y(e^{j\omega})}{X(e^{j\omega})} = \frac{\sum_{k=0}^{M} b_k e^{-jk\omega}}{\sum_{k=0}^{N} a_k e^{-jk\omega}}$$

$$H(e^{j\omega}) = \frac{2}{1 - \frac{3}{4}e^{-j\omega} + \frac{1}{8}e^{-j2\omega}}$$

To obtain the impulse response, we factor the denominator

$$H(e^{j\omega}) = \frac{2}{(1 - \frac{1}{2}e^{-j\omega})(1 - \frac{1}{4}e^{-j\omega})}$$

Example 2 - continued

$$H(e^{j\omega}) = \frac{2}{(1 - \frac{1}{2}e^{-j\omega})(1 - \frac{1}{4}e^{-j\omega})}$$

$H(e^{j\omega})$ can be expanded by the method of partial fractions as

$$H(e^{j\omega}) = \frac{4}{1 - \frac{1}{2}e^{-j\omega}} - \frac{2}{1 - \frac{1}{4}e^{-j\omega}}$$

The inverse transform of each term can be recognized by inspection. The impulse response of the system is

$$h[n] = 4\left(\frac{1}{2}\right)^n u[n] - 2\left(\frac{1}{4}\right)^n u[n].$$

SAMPLING

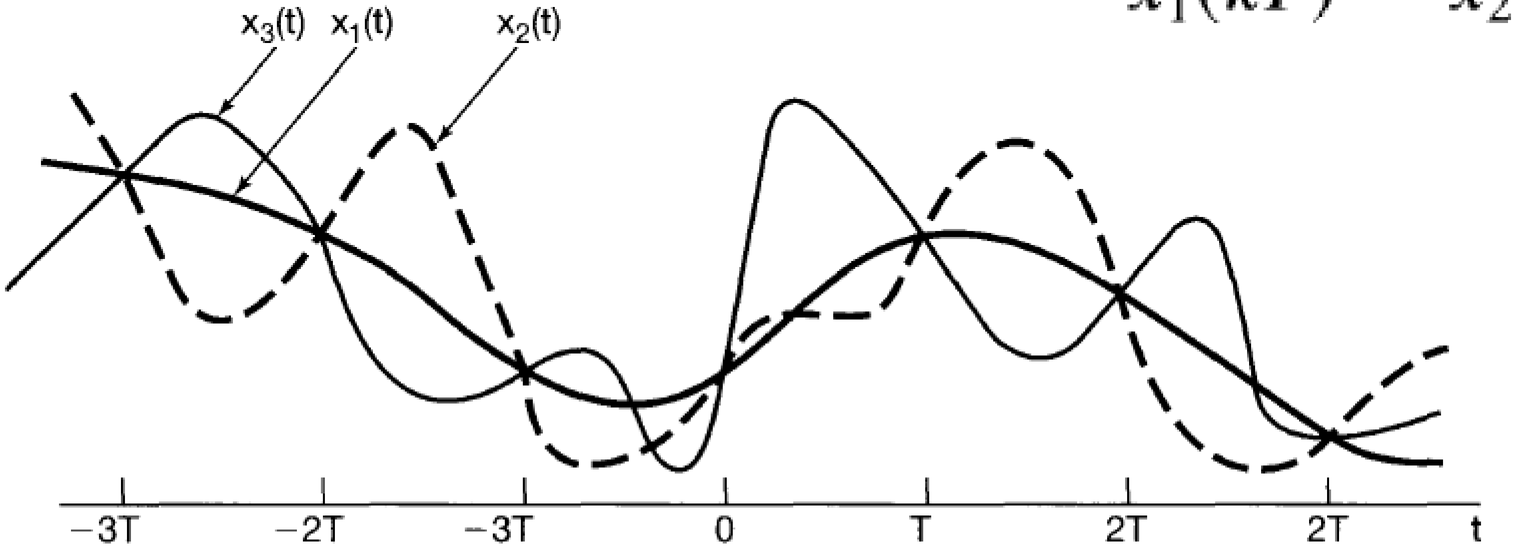
Introduction

- Under certain conditions, a continuous-time signal can be completely represented by and recoverable from knowledge of its values, or samples, at points equally spaced in time.
- This property follows from a basic result that is referred to as the sampling theorem.
- The sampling theorem is a bridge between continuous-time signals and discrete-time signals.
- The sampling theorem provides a mechanism for representing a continuous-time signal by a discrete-time signal.
- In many contexts, processing discrete-time signals is more flexible and is often preferable to processing continuous-time signals.
- This is due to the dramatic development of digital technology, resulting in the availability of inexpensive, lightweight, programmable, and easily reproducible discrete-time systems.

Representation of a continuous-time signal by its samples: The sampling theorem

In general, in the absence of any additional conditions or information, we would not expect that a signal could be uniquely specified by a sequence of equally spaced samples.
For example, below is a figure showing three continuous-time signals with identical values at integer multiples of T.

$$x_1(kT) = x_2(kT) = x_3(kT)$$



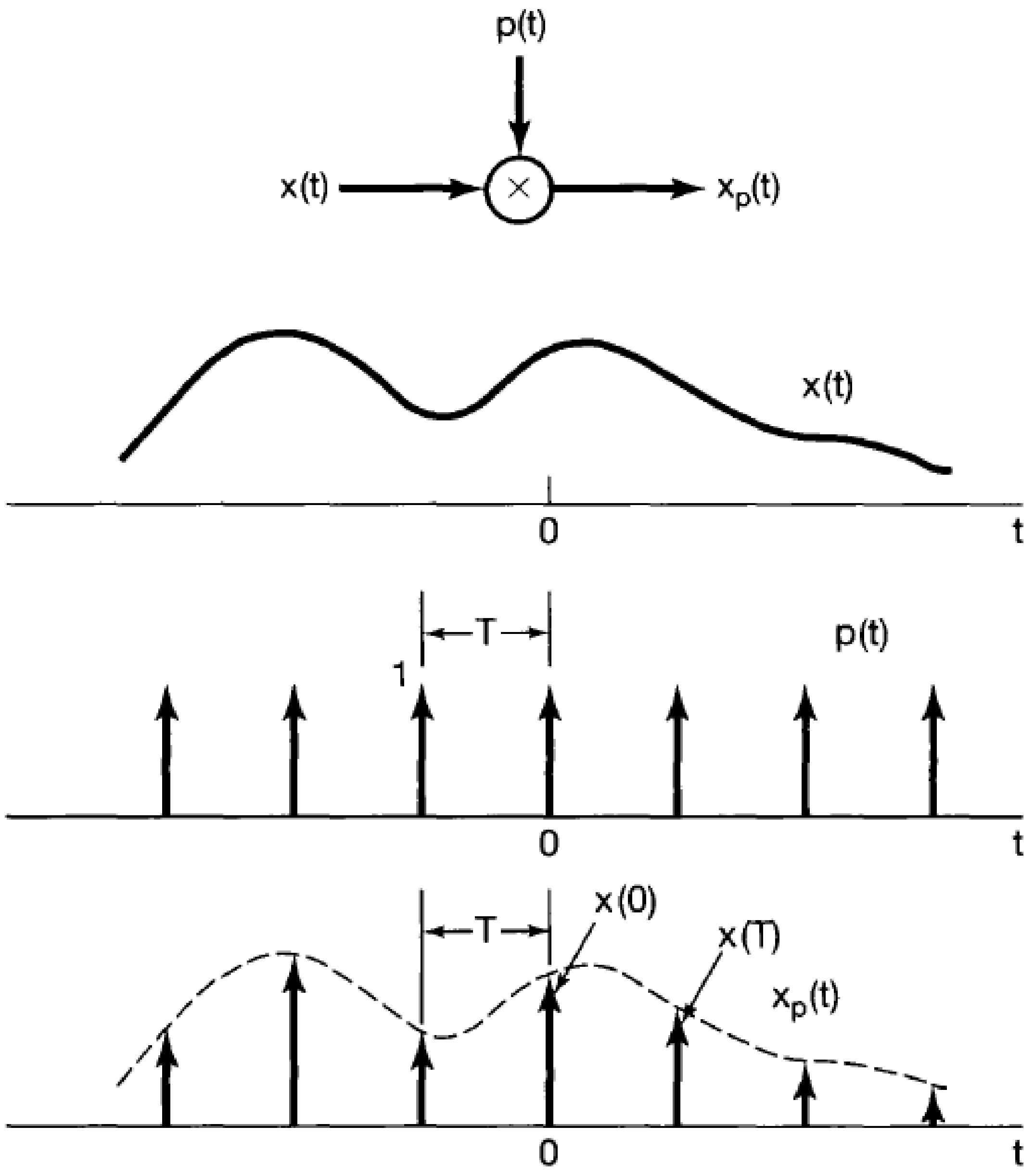
Representation of a continuous-time signal by its samples: The sampling theorem

If a signal is band limited (if its Fourier transform is zero outside a finite band of frequencies) and if the samples are taken sufficiently close together in relation to the highest frequency present in the signal, then the samples uniquely specify the signal, and we can reconstruct it perfectly.

This result is known as the sampling theorem.

Impulse Train Sampling

We need a convenient way to represent the sampling of a continuous-time signal at regular intervals.
A way to do this is through the use of a periodic impulse train multiplied by the continuous-time signal $x(t)$ that we wish to sample.
This mechanism is known as impulse-train sampling.



Impulse Train Sampling

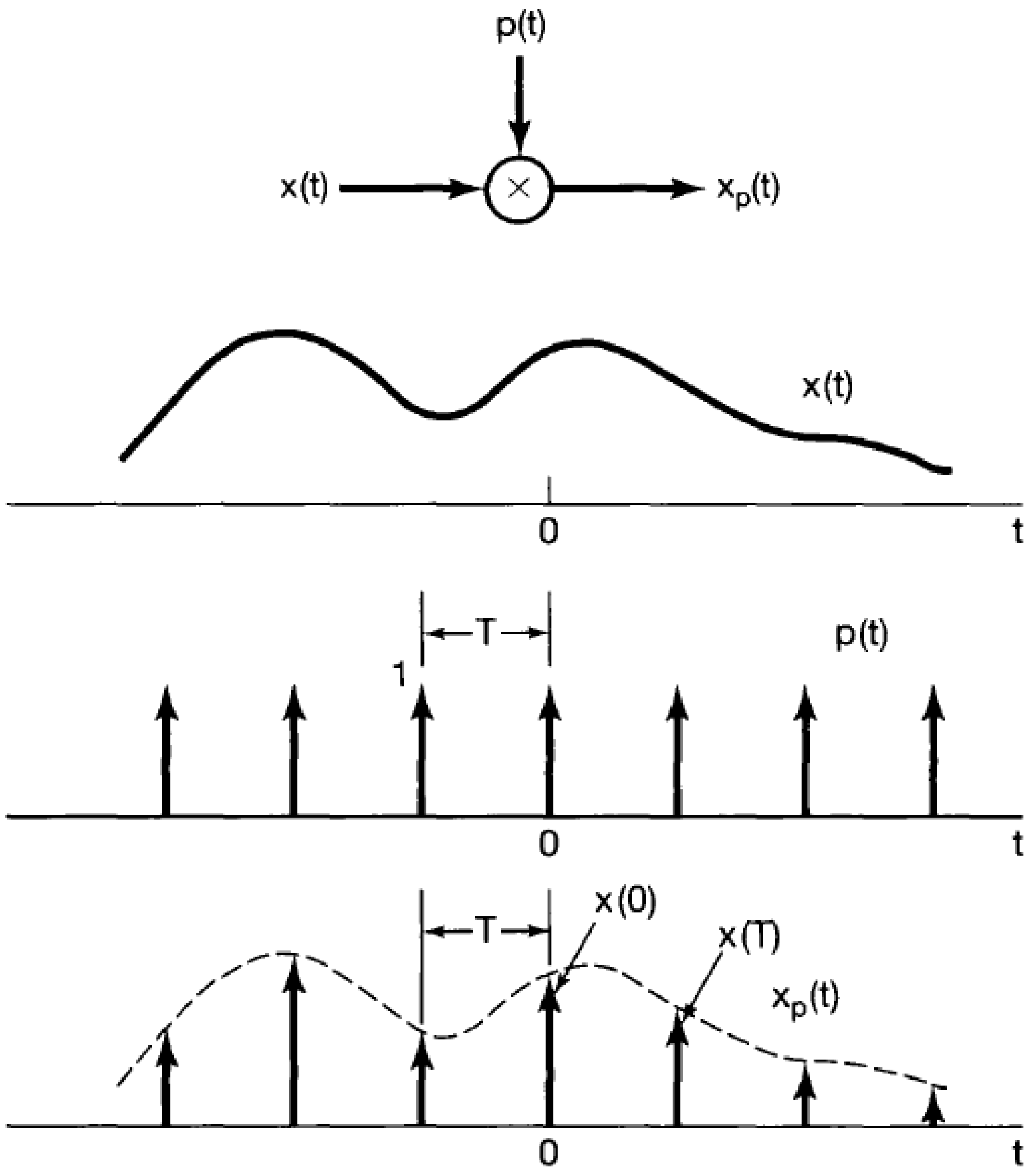
$$p(t) = \sum_{n=-\infty}^{+\infty} \delta(t - nT)$$

The periodic impulse train $p(t)$ is referred to as the sampling function.

The period T is the sampling period.

The fundamental frequency of $p(t)$, $\omega_s = \frac{2\pi}{T}$ is the sampling frequency.

$$x_p(t) = x(t)p(t)$$



Impulse Train Sampling

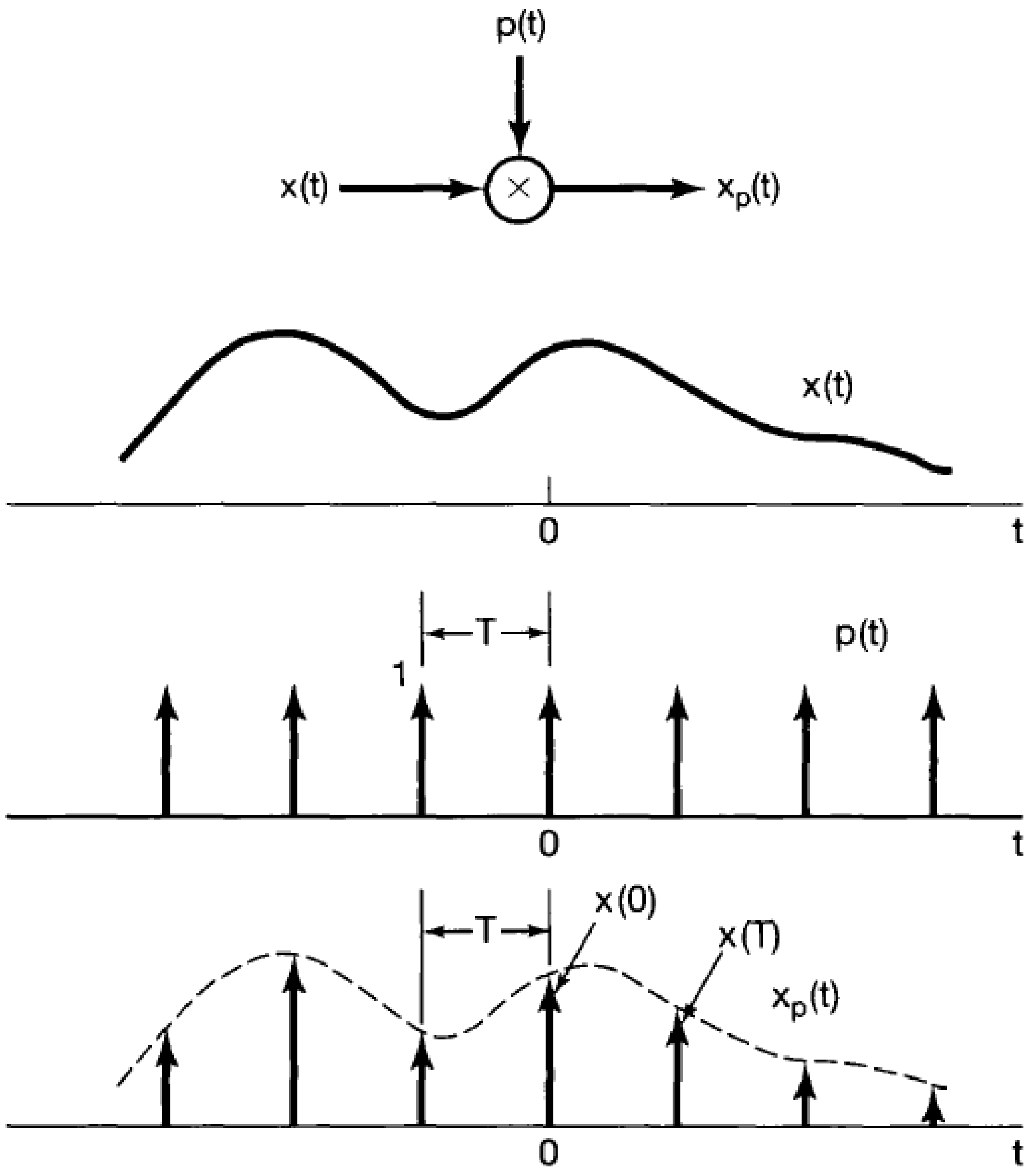
$$p(t) = \sum_{n=-\infty}^{+\infty} \delta(t - nT)$$

$$x_p(t) = x(t)p(t)$$

Multiplying $x(t)$ by a unit impulse samples the value of the signal at the point at which the impulse is located:
$x(t)\delta(t - t_0) = x(t_0)\delta(t - t_0)$.

Then $x_p(t)$ is an impulse train with the amplitudes of the impulses equal to the samples of $x(t)$ at intervals spaced by T:

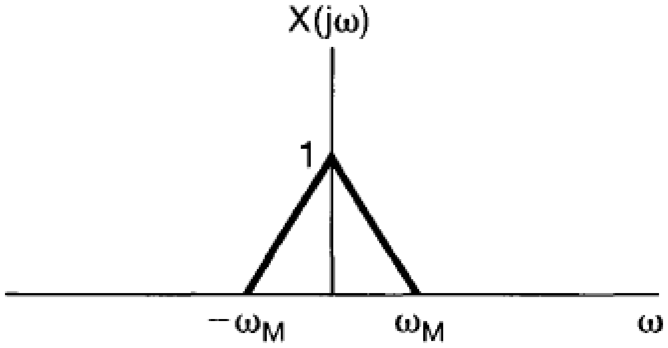
$$x_p(t) = \sum_{n=-\infty}^{+\infty} x(nT)\delta(t - nT)$$



Impulse Train Sampling

$$p(t) = \sum_{n=-\infty}^{+\infty} \delta(t - nT)$$

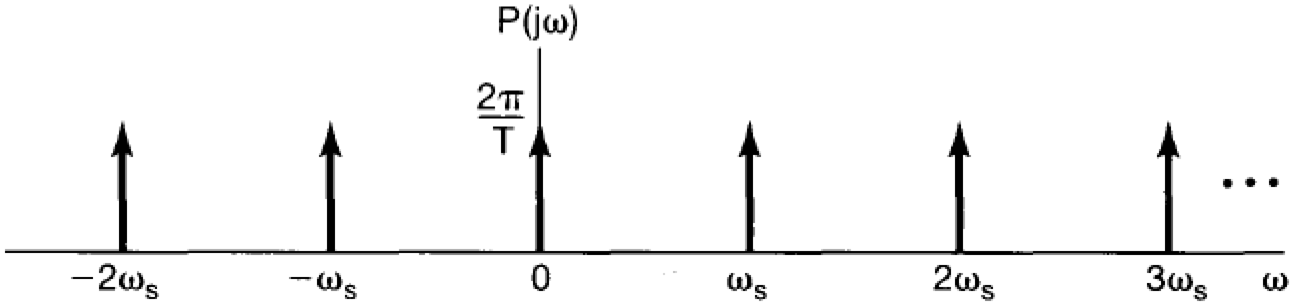
$$x_p(t) = x(t)p(t)$$



$$x_p(t) = \sum_{n=-\infty}^{+\infty} x(nT)\delta(t - nT)$$

The Fourier Transform of $p(t)$ is

$$P(j\omega) = \frac{2\pi}{T} \sum_{k=-\infty}^{+\infty} \delta(\omega - k\omega_s)$$

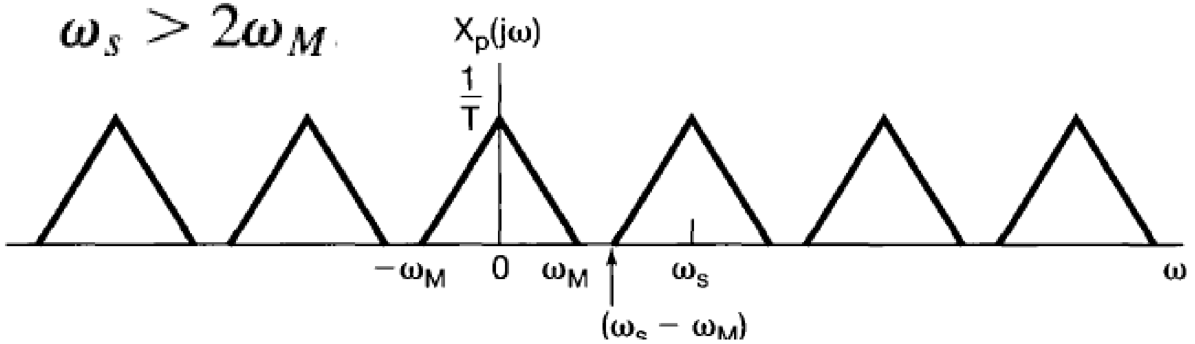


From the multiplication property:

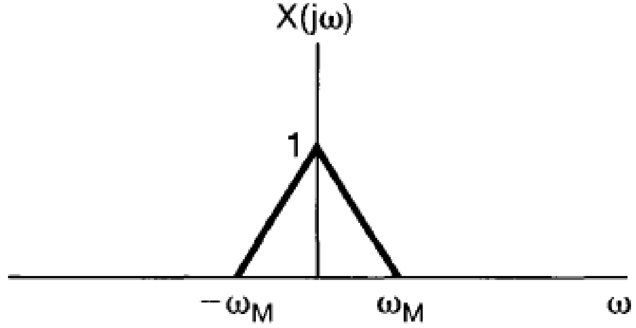
$$X_p(j\omega) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} X(j\theta)P(j(\omega - \theta))d\theta$$

$$X_p(j\omega) = \frac{1}{T} \sum_{k=-\infty}^{+\infty} X(j(\omega - k\omega_s))$$

$\omega_s > 2\omega_M$

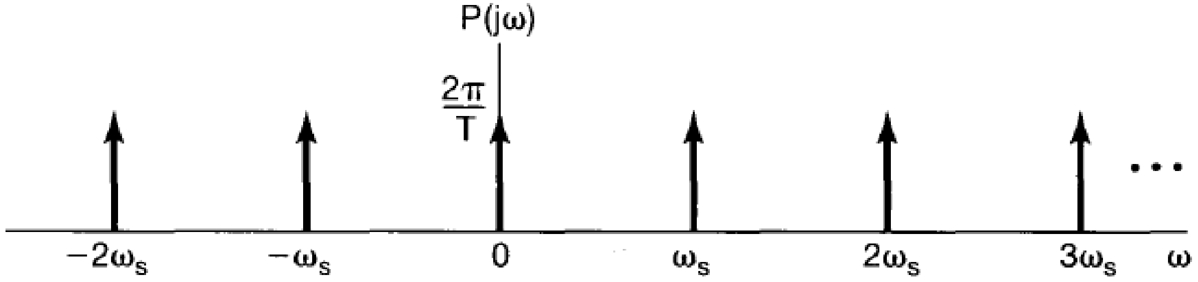


Impulse Train Sampling



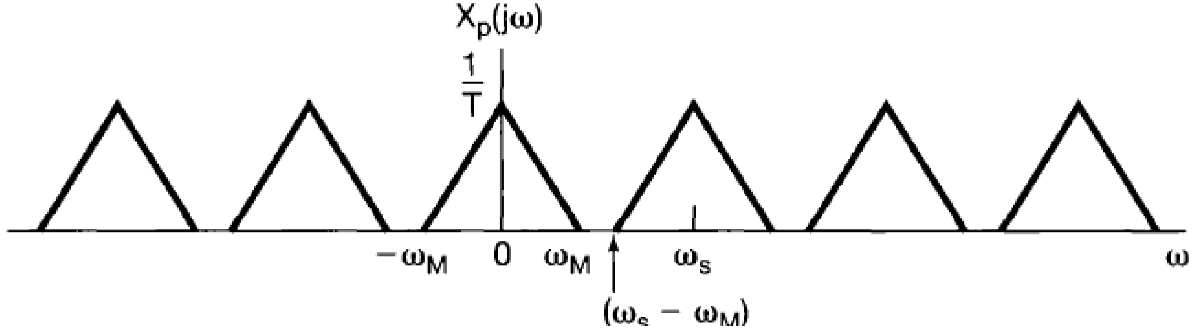
$$X_p(j\omega) = \frac{1}{T}\sum_{k=-\infty}^{+\infty} X(j(\omega - k\omega_s))$$

$X_p(j\omega)$ is a periodic function of ω consisting of a superposition of shifted replicas of $X(j\omega)$, scaled by $1/T$.



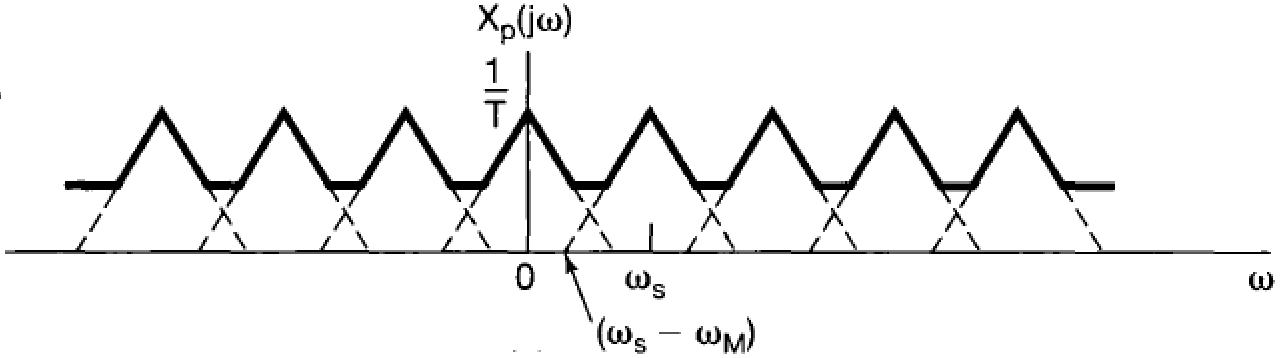
$$\omega_s > 2\omega_M$$

There is no overlap between the shifted replicas of $X(j\omega)$.



$$\omega_s < 2\omega_M$$

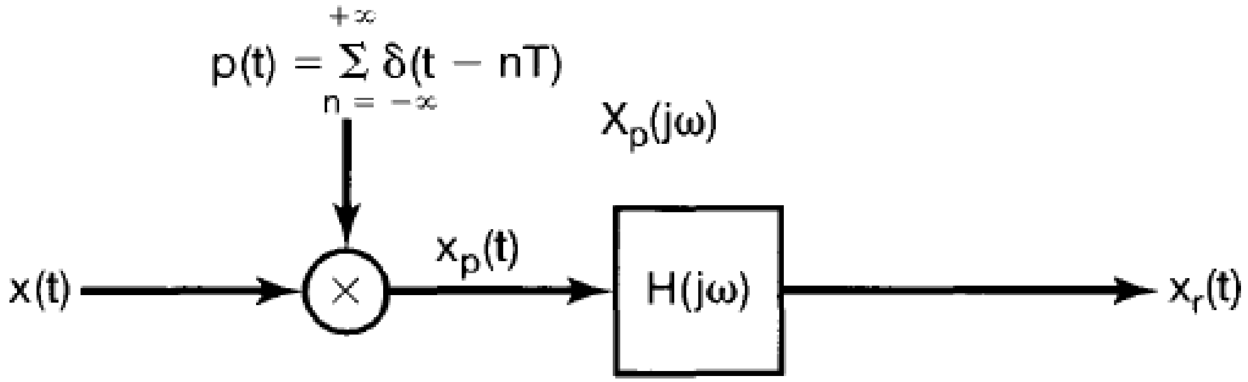
There is overlap between the shifted replicas of $X(j\omega)$.

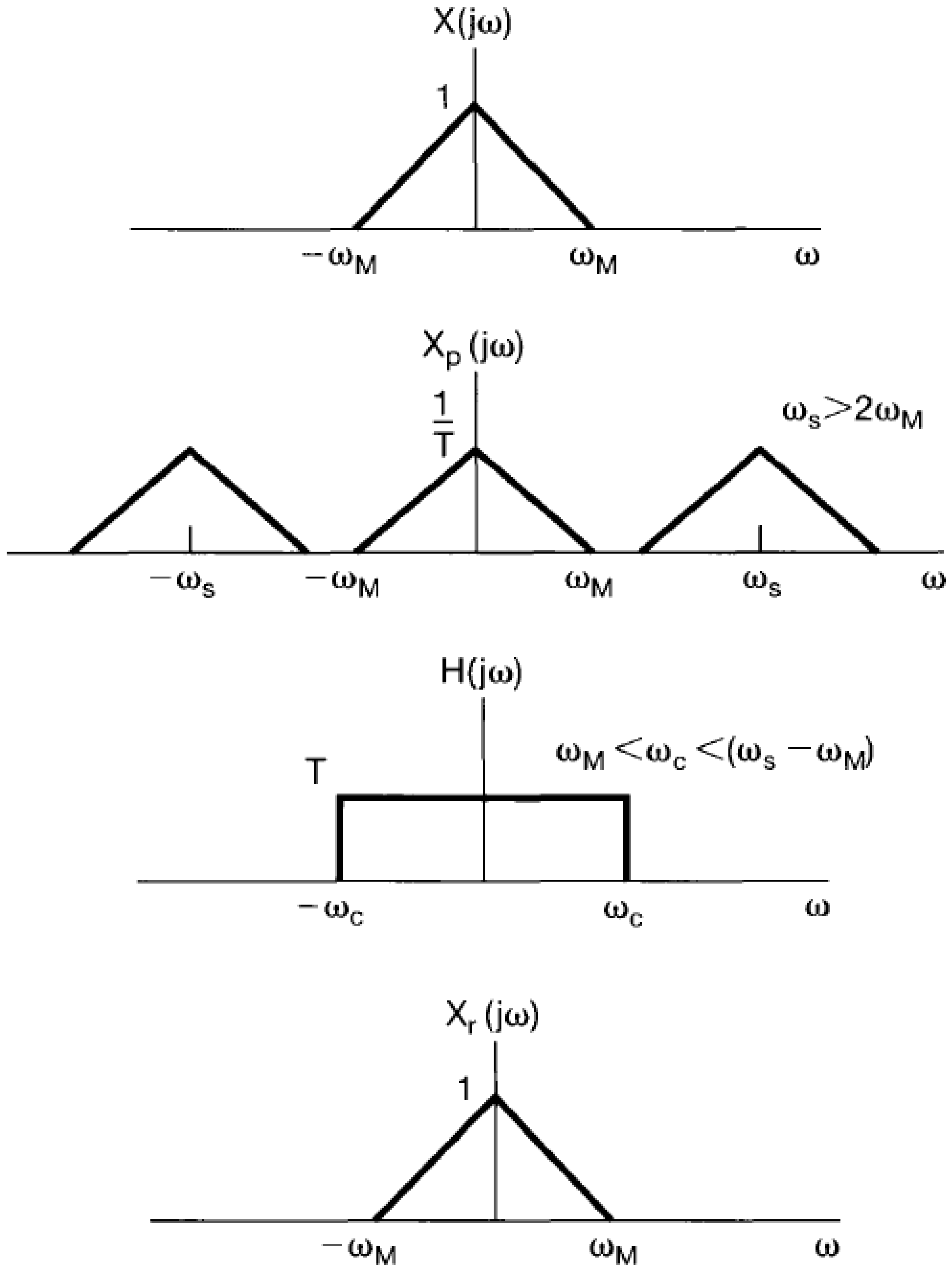


Impulse Train Sampling

$$X_p(j\omega) = \frac{1}{T}\sum_{k=-\infty}^{+\infty} X(j(\omega - k\omega_s))$$

If $\omega_s > 2\omega_M$, $x(t)$ can be recovered exactly from $x_p(t)$ by means of a lowpass filter with gain T and a cutoff frequency greater than ω_M and less than $\omega_s - \omega_M$.





SAMPLING THEOREM

Sampling Theorem:

Let $x(t)$ be a band-limited signal with $X(j\omega) = 0$ for $|\omega| > \omega_M$. Then $x(t)$ is uniquely determined by its samples $x(nT)$, $n = 0, \pm 1, \pm 2, \ldots$, if

$$\omega_s > 2\omega_M,$$

where

$$\omega_s = \frac{2\pi}{T}.$$

Given these samples, we can reconstruct $x(t)$ by generating a periodic impulse train in which successive impulses have amplitudes that are successive sample values. This impulse train is then processed through an ideal lowpass filter with gain T and cutoff frequency greater than ω_M and less than $\omega_s - \omega_M$. The resulting output signal will exactly equal $x(t)$.

SAMPLING THEOREM

Sampling Theorem:
Let $x(t)$ be a band-limited signal with $X(j\omega) = 0$ for $|\omega| > \omega_M$. Then $x(t)$ is uniquely determined by its samples $x(nT)$, $n = 0, \pm 1, \pm 2, \ldots$, if

$$\omega_s > 2\omega_M,$$

where

$$\omega_s = \frac{2\pi}{T}.$$

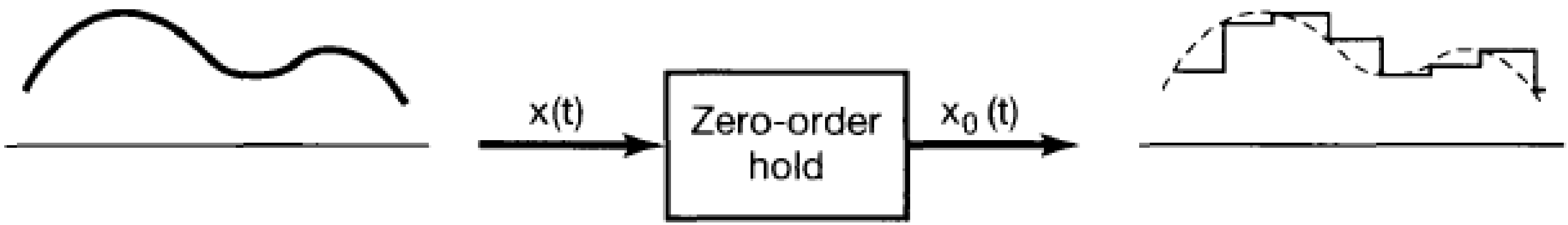
Given these samples, we can reconstruct $x(t)$ by generating a periodic impulse train in which successive impulses have amplitudes that are successive sample values. This impulse train is then processed through an ideal lowpass filter with gain T and cutoff frequency greater than ω_M and less than $\omega_s - \omega_M$. The resulting output signal will exactly equal $x(t)$.

The frequency $2\omega_M$, which, under the sampling theorem, must be exceeded by the sampling frequency, is commonly referred to as the *Nyquist rate.*

Sampling with a Zero-Order Hold

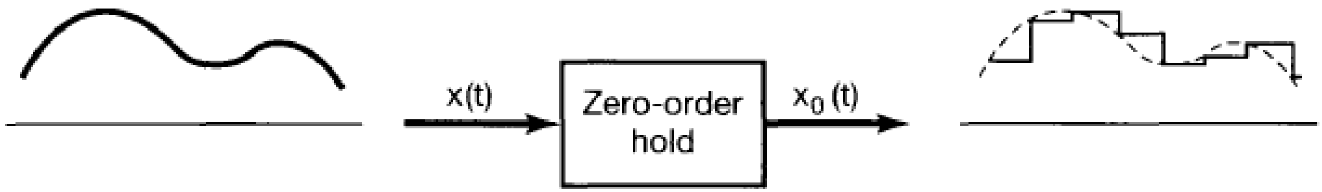
The sampling theorem, which is most easily explained in terms of impulse-train sampling, establishes the fact that a band-limited signal is uniquely represented by its samples.
In practice, narrow, large-amplitude pulses, which approximate impulses, are difficult to generate.

It is often more convenient to generate the sampled signal in a form referred to as a zero-order hold.
Such a system samples $x(t)$ at a given instant and holds that value until the next instant at which a sample is taken.

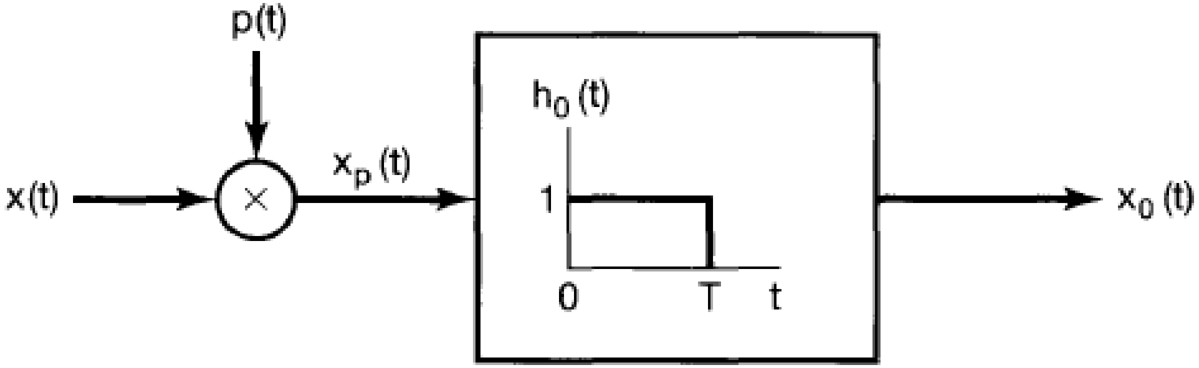


Sampling with a Zero-Order Hold

A zero-order hold system samples $x(t)$ at a given instant and holds that value until the next instant at which a sample is taken.

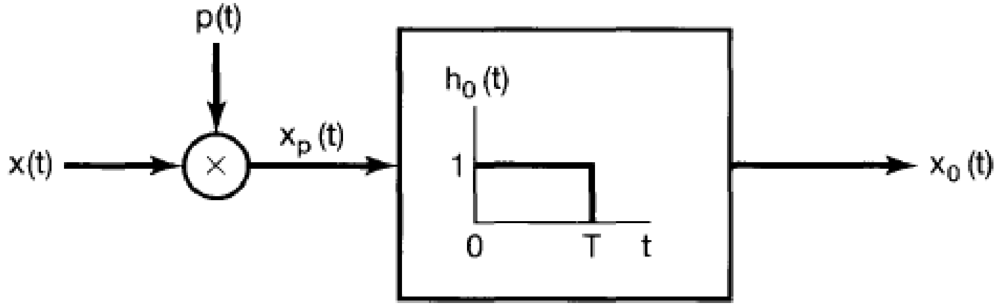


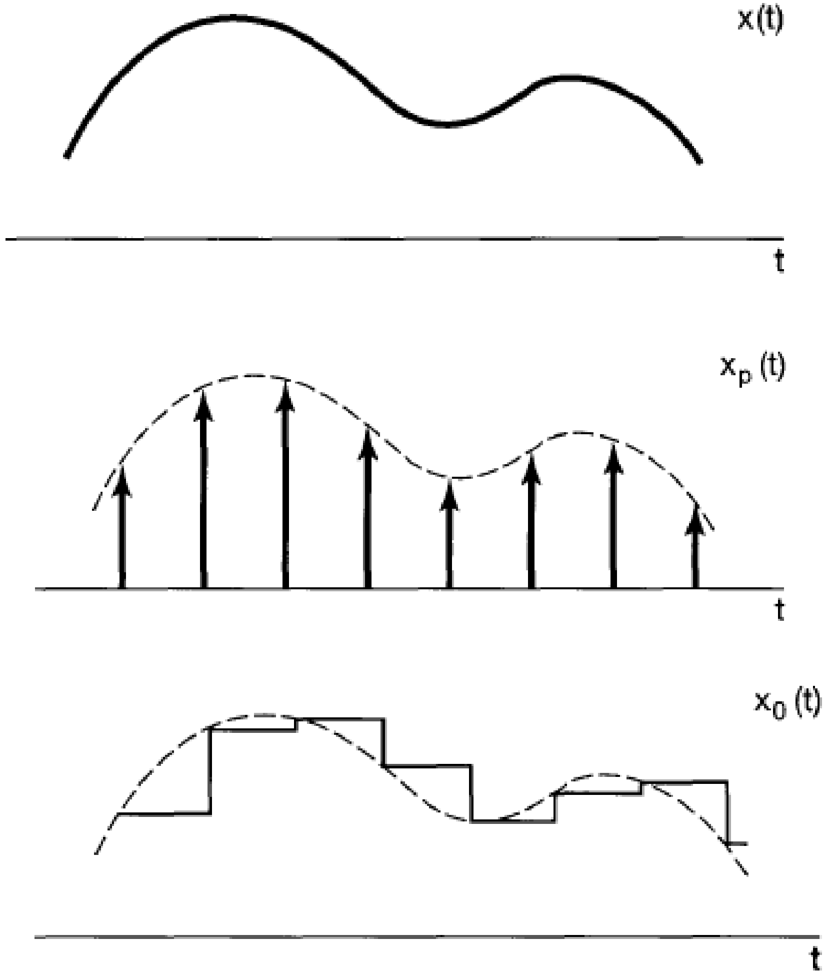
The output $x_0(t)$ of the zero-order hold can in principle be generated by impulse-train sampling followed by an LTI system with a rectangular impulse response.



Sampling with a Zero-Order Hold

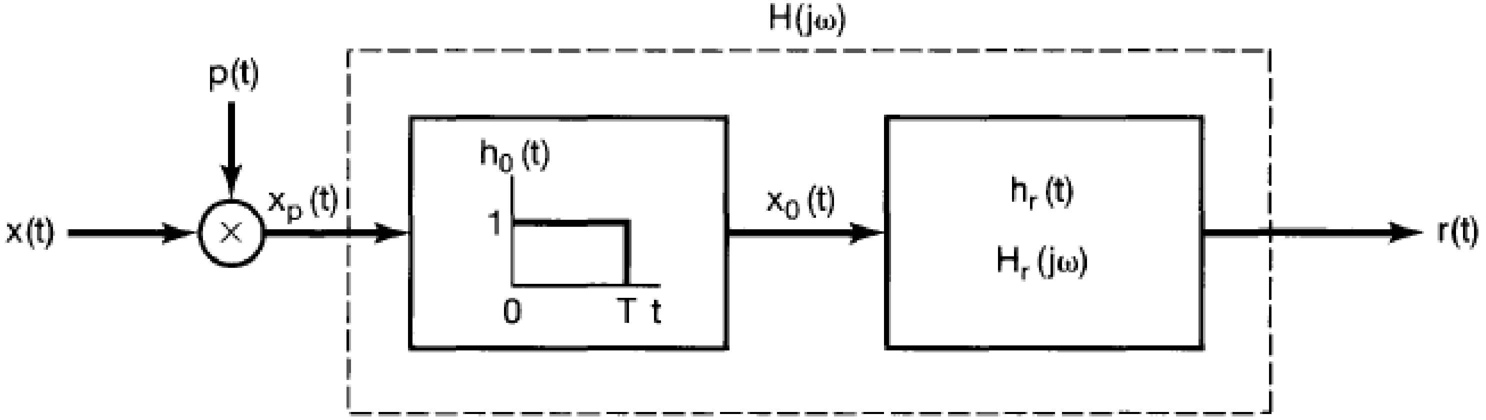
The output $x_0(t)$ of the zero-order hold can in principle be generated by impulse-train sampling followed by an LTI system with a rectangular impulse response.



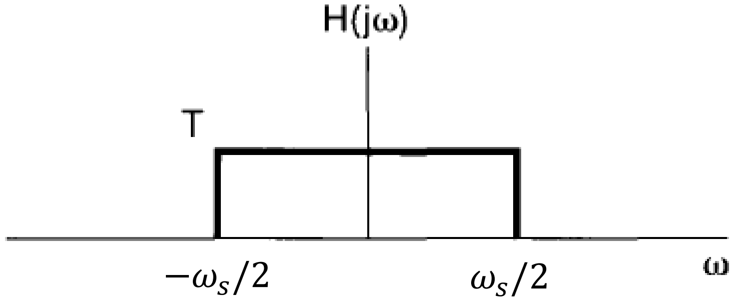


Sampling with a Zero-Order Hold

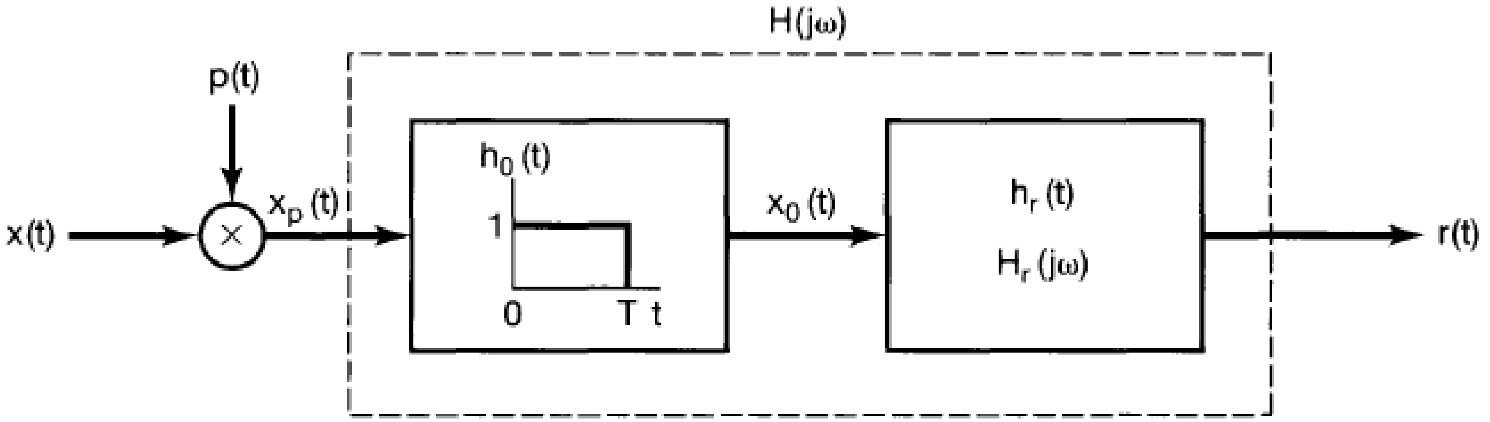
To reconstruct $x(t)$ from $x_0(t)$, we consider processing $x_0(t)$ with an LTI system with impulse response $h_r(t)$ and frequency response $H_r(j\omega)$.



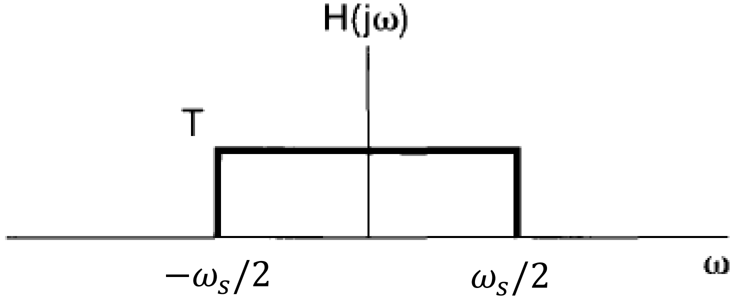
We wish to specify $H_r(j\omega)$ so that $r(t) = x(t)$.
For that $H(j\omega)$ should be the ideal low-pass filter.



Sampling with a Zero-Order Hold



$\mathrm{H}(j\omega) = H_0(j\omega)H_r(j\omega)$ should be the ideal low-pass filter.



The Fourier transform of $h_0(t)$ is

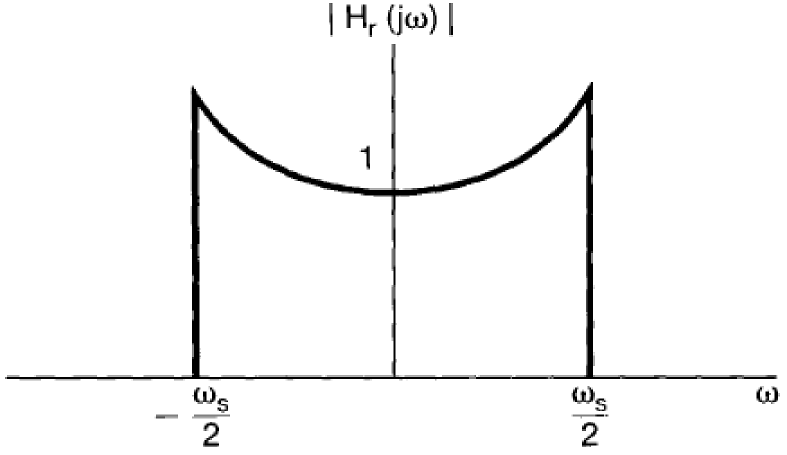
$$H_0(j\omega) = e^{-j\omega T/2}\left[\frac{2\sin(\omega T/2)}{\omega}\right]$$

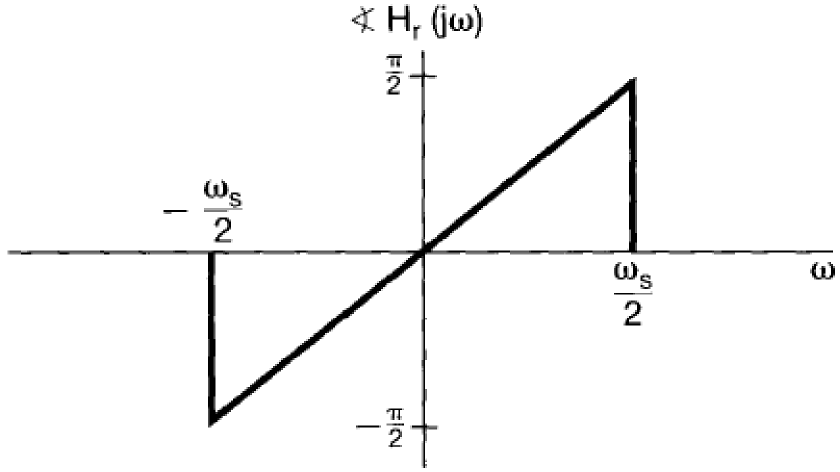
$$H_r(j\omega) = \frac{e^{j\omega T/2}H(j\omega)}{\dfrac{2\sin(\omega T/2)}{\omega}}$$

Sampling with a Zero-Order Hold

The reconstruction filter for a zero-order hold:

$$H_r(j\omega) = \frac{e^{j\omega T/2}H(j\omega)}{\frac{2\sin(\omega T/2)}{\omega}}$$





In practice this frequency response $H_r(j\omega)$ cannot be exactly realized.
We need to design an adequate approximation to the original signal.

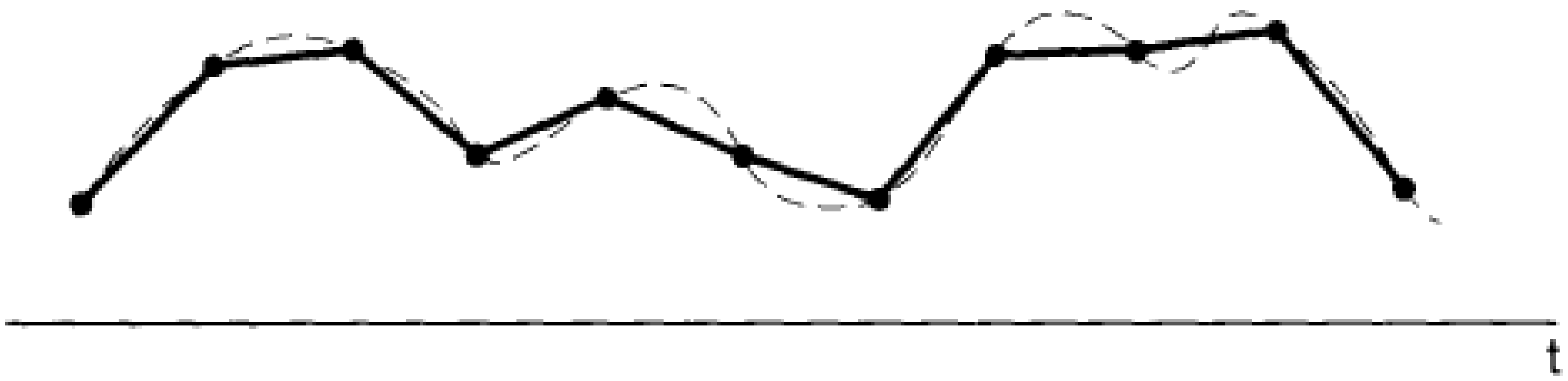
Reconstruction of a signal from its samples using interpolation

Interpolation, that is, the fitting of a continuous signal to a set of sample values, is a commonly used procedure for reconstructing a function, either approximately or exactly, from samples.

One simple interpolation procedure is the zero-order hold.

Another useful form of interpolation is linear interpolation, whereby adjacent sample points are connected by a straight line.

In more complicated interpolation formulas, sample points may be connected by higher order polynomials or other mathematical functions.



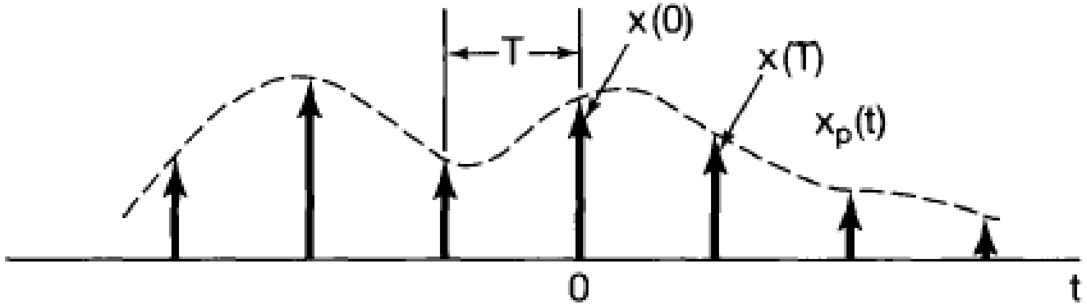
Linear interpolation between sample points. The dashed curve represents the original signal and the solid curve the linear interpolation.

Reconstruction of a signal from its samples using interpolation

For a band-limited signal, if the sampling instants are sufficiently close, then the signal can be reconstructed exactly; i.e., through the use of a lowpass filter, exact interpolation can be carried out between the sample points.

Consider the signal $x_p(t)$ sampled by an impulse-train:

$$x_p(t) = \sum_{n=-\infty}^{+\infty} x(nT)\delta(t - nT)$$



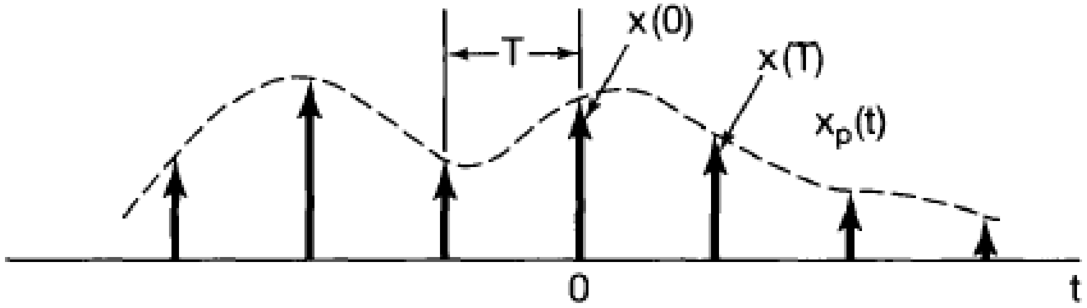
To reconstruct the original signal from $x_p(t)$ we apply lowpass filtering.
In time-domain, the output of the low-pass filter is:

$$x_r(t) = x_p(t) * h(t)$$

$$x_r(t) = \sum_{n=-\infty}^{+\infty} x(nT)h(t - nT)$$

Reconstruction of a signal from its samples using interpolation

$$x_p(t) = \sum_{n=-\infty}^{+\infty} x(nT)\delta(t - nT)$$



To reconstruct the original signal from $x_p(t)$ we apply lowpass filtering.
In time-domain, the output of the low-pass filter is

$$x_r(t) = \sum_{n=-\infty}^{+\infty} x(nT)h(t - nT)$$

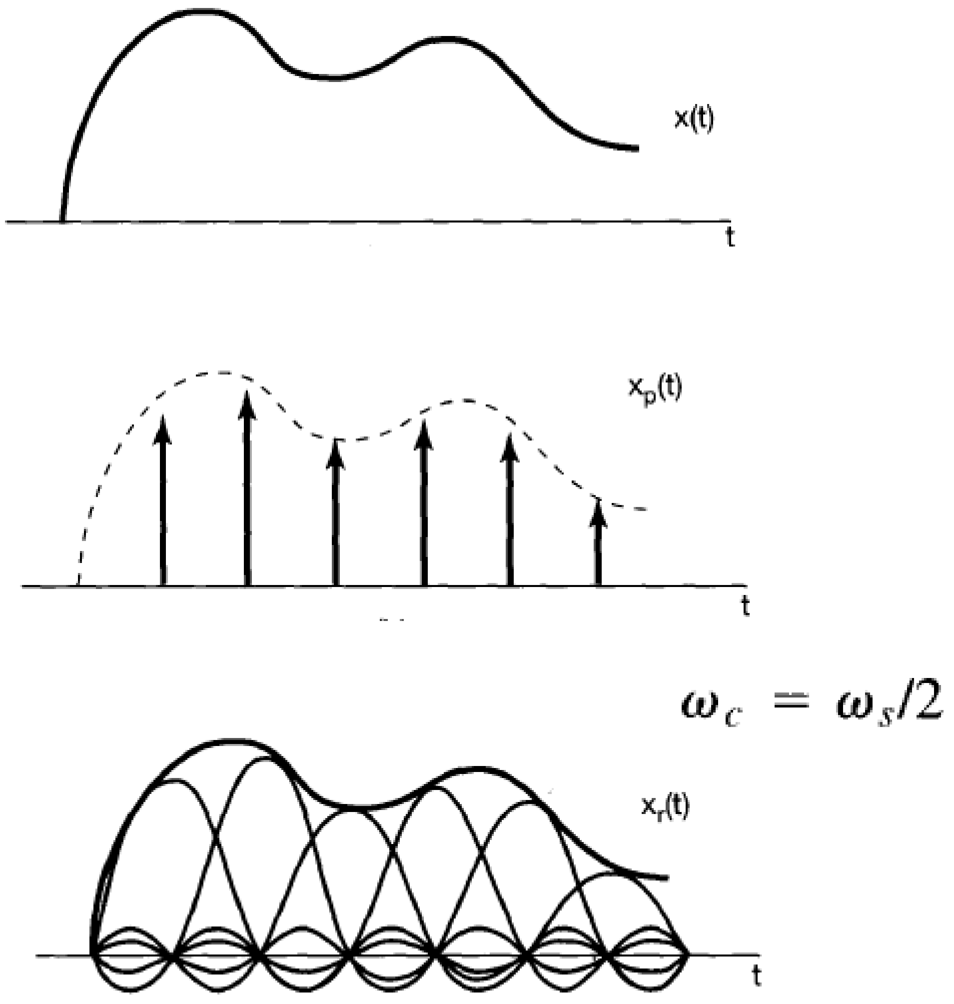
For the ideal lowpass filter the impulse response is

$$h(t) = \frac{\omega_c T \sin(\omega_c t)}{\pi \omega_c t}$$

$$x_r(t) = \sum_{n=-\infty}^{+\infty} x(nT)\frac{\omega_c T}{\pi}\frac{\sin(\omega_c(t - nT))}{\omega_c(t - nT)}$$

Reconstruction of a signal from its samples using interpolation

$$x_r(t) = \sum_{n=-\infty}^{+\infty} x(nT)\frac{\omega_c T}{\pi}\frac{\sin(\omega_c(t-nT))}{\omega_c(t-nT)}$$



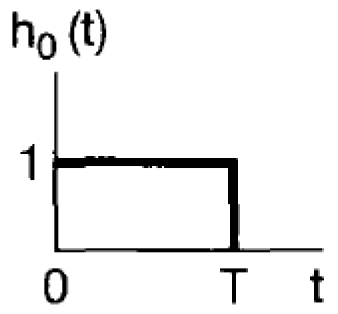
This operation corresponds to ideal band-limited interpolation using the sinc function.
With the superposition of the individual terms in the sum the original signal is reconstructed for the case $\omega_c = \omega_s/2$.

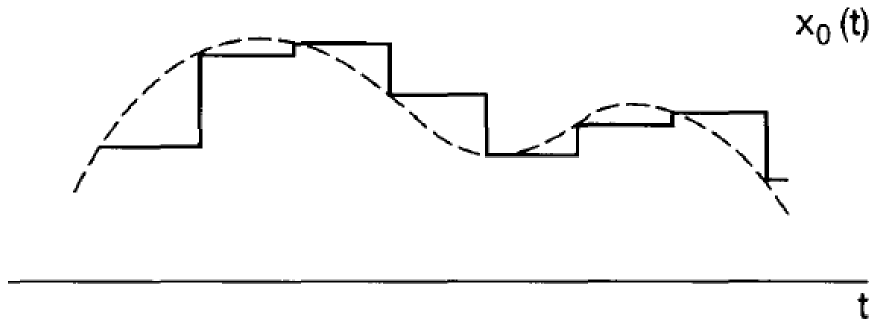
Interpolation using the impulse response of an ideal lowpass filter is commonly referred to as band-limited interpolation, since it implements exact reconstruction if $x(t)$ is band limited and the sampling frequency ω_s satisfies the conditions of the sampling theorem.

Reconstruction of a signal from its samples using interpolation

In many cases it is preferable to use a less accurate, but simpler, filter or, equivalently, a simpler interpolating function than the sinc function.

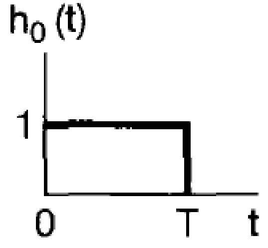
For example, the zero-order hold can be viewed as a form of interpolation between sample values in which the interpolating function $h(t)$ is the impulse response $h_0(t)$.

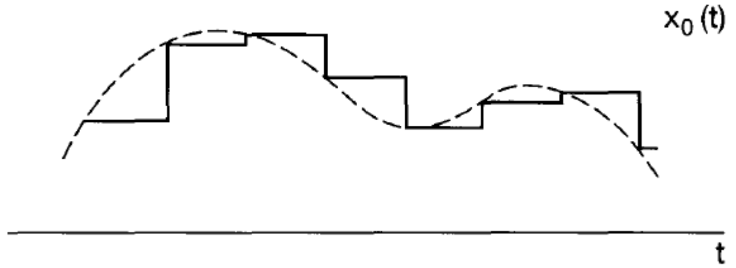




With $x_0(t)$ in the figure corresponding to the approximation to $x(t)$, the systems $h_0(t)$ represents an approximation to the ideal lowpass filter required for the exact interpolation.

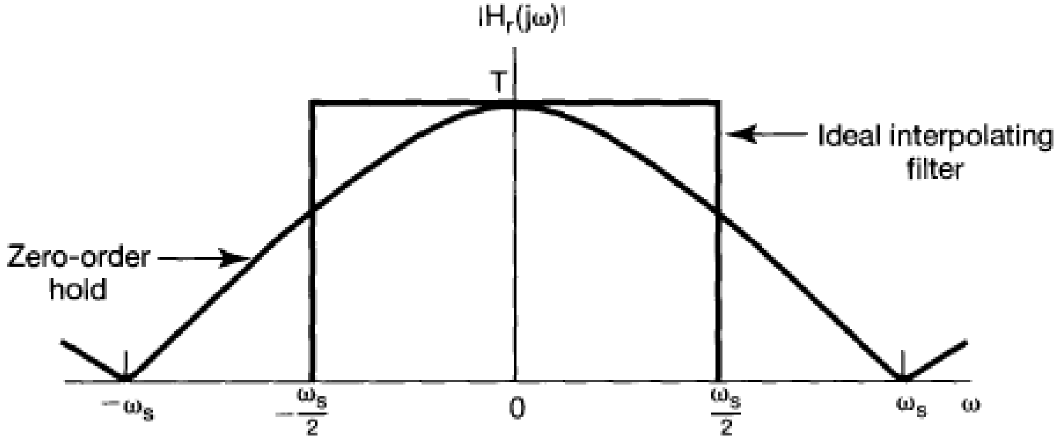
Reconstruction of a signal from its samples using interpolation





With $x_0(t)$ in the figure corresponding to the approximation to $x(t)$, the systems $h_0(t)$ represents an approximation to the ideal lowpass filter required for the exact interpolation.

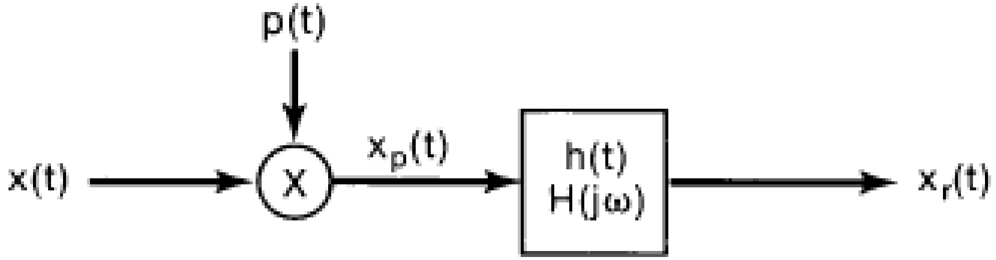
The magnitude of the frequency response of the zero-order-hold interpolating filter, superimposed on the desired frequency response of the exact interpolating filter is given below:

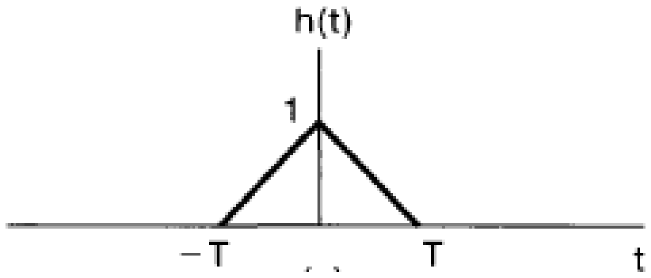


Reconstruction of a signal from its samples using interpolation

The interpolation provided by the zero-order hold is crude.
We can use a variety of smoother interpolation strategies, some of which are known collectively as higher order holds.

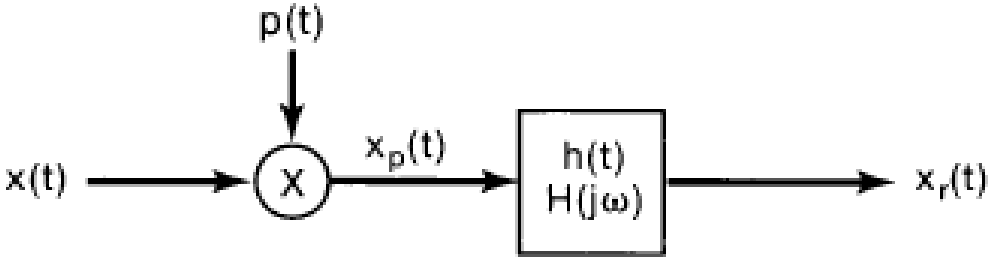
Linear interpolation, which is sometimes referred to as a first-order hold, can also be viewed as interpolation with $h(t)$ triangular.

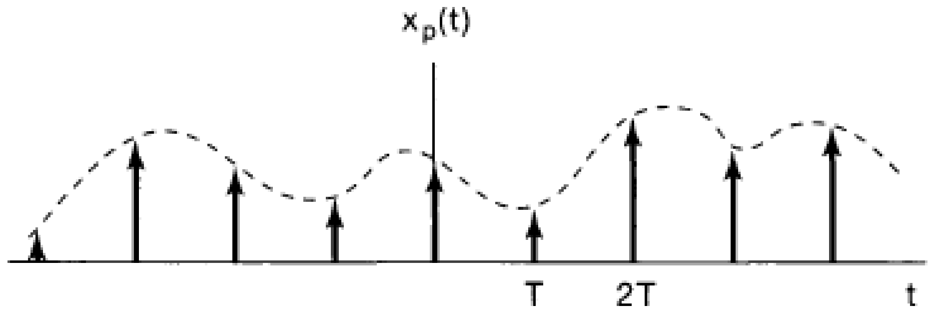


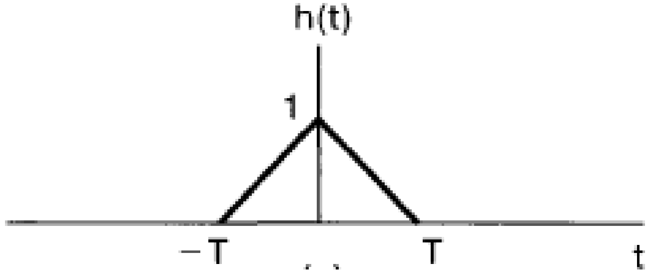


Reconstruction of a signal from its samples using interpolation

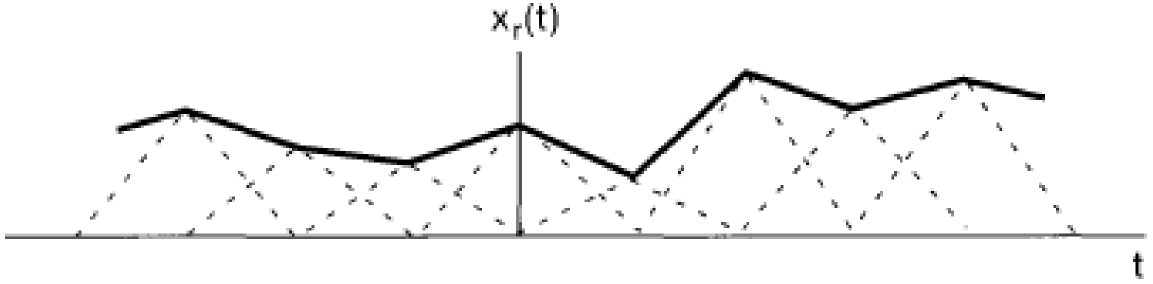
Linear interpolation, which is sometimes referred to as a first-order hold, can also be viewed as interpolation with $h(t)$ triangular.





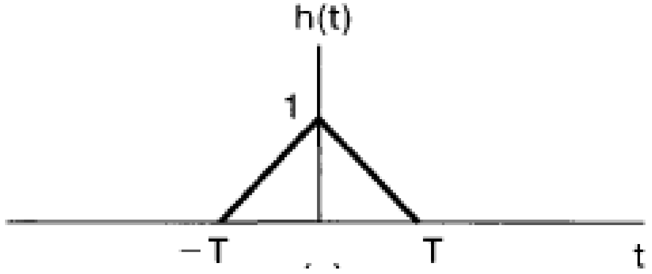


Linear interpolation yields reconstructions that are continuous, although with discontinuous derivatives due to the changes in slope at the sample points



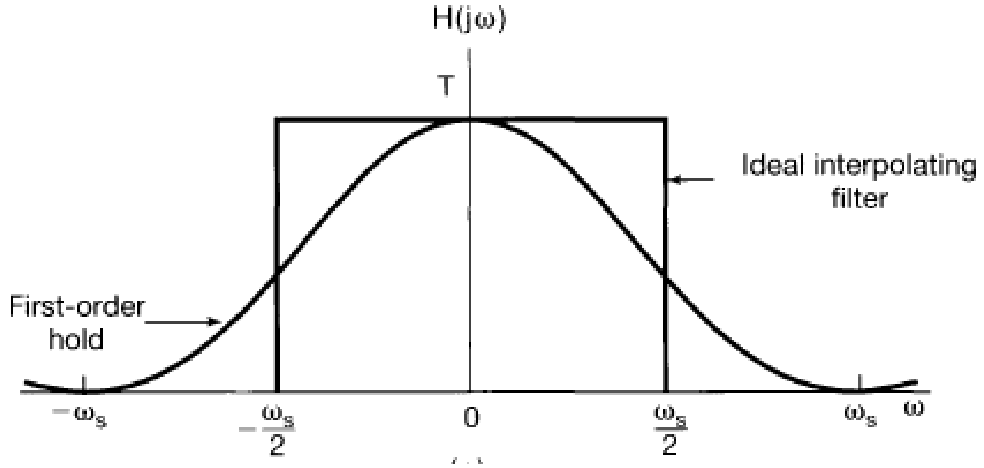
Reconstruction of a signal from its samples using interpolation

The impulse function for linear interpolation $h(t)$:



The associated frequency response is

$$H(j\omega) = \frac{1}{T}\left[\frac{\sin(\omega T/2)}{\omega/2}\right]^2$$

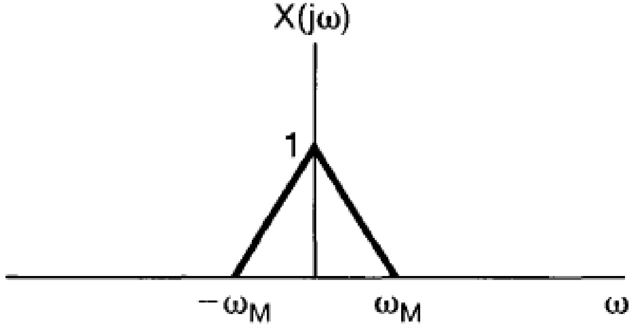


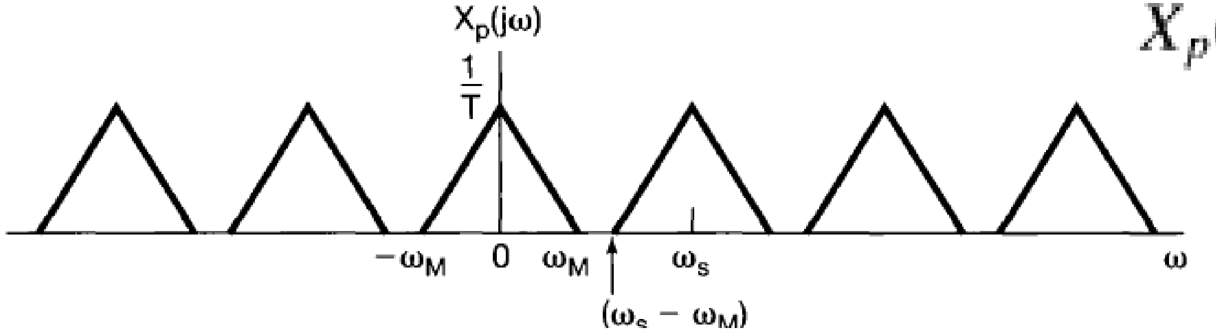
The magnitude of the frequency response of the first-order-hold interpolating filter, superimposed on the frequency response of the exact interpolating filter.

The effect of undersampling: Aliasing

We assumed that the sampling frequency (ω_s) was sufficiently high that the conditions of the sampling theorem were met ($\omega_s > 2\omega_M$).

When $\omega_s > 2\omega_M$, the spectrum of the sampled signal $x_p(t)$ consists of scaled replications of the spectrum of $x(t)$, and this forms the basis for the sampling theorem.

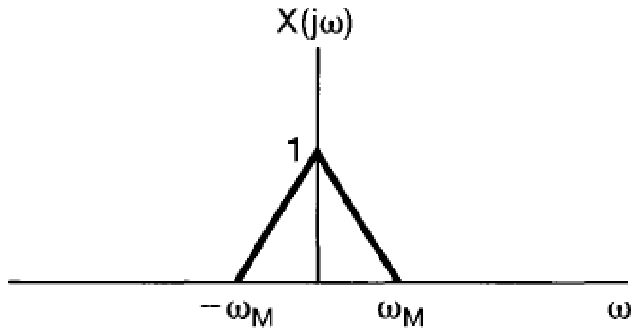


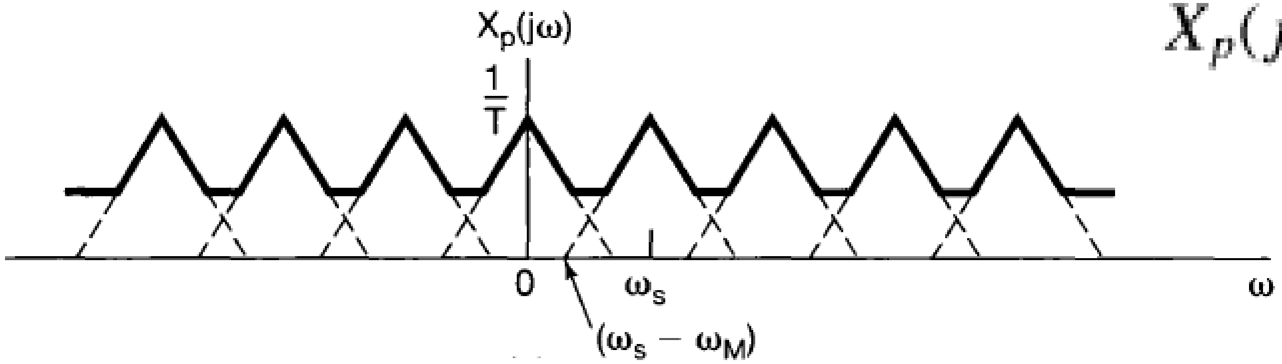


$$X_p(j\omega) = \frac{1}{T}\sum_{k=-\infty}^{+\infty} X(j(\omega - k\omega_s)).$$

The effect of undersampling: Aliasing

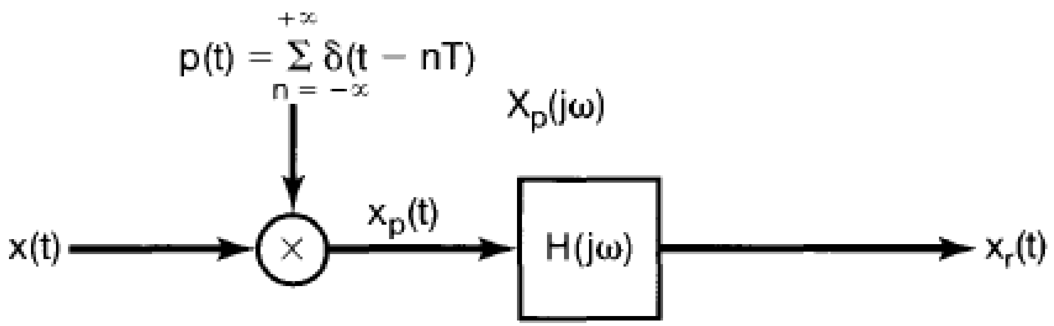
When $\omega_s < 2\omega_M$, $X(j\omega)$, the spectrum of $x(t)$, is no longer replicated in $X_p(j\omega)$ and thus is no longer recoverable by lowpass filtering.
This effect, in which the individual terms in the summation of the shifted replicas overlap, is referred to as aliasing.





$$X_p(j\omega) = \frac{1}{T}\sum_{k=-\infty}^{+\infty} X(j(\omega - k\omega_s)).$$

The effect of undersampling: Aliasing



If the system is applied to a signal with $\omega_s < 2\omega_M$, the reconstructed signal $x_r(t)$ will no longer be equal to $x(t)$.
However, the original signal $x(t)$ and the signal $x_r(t)$ that is reconstructed using bandlimited interpolation will always be equal at the sampling instants; that is, for any choice of ω_s,

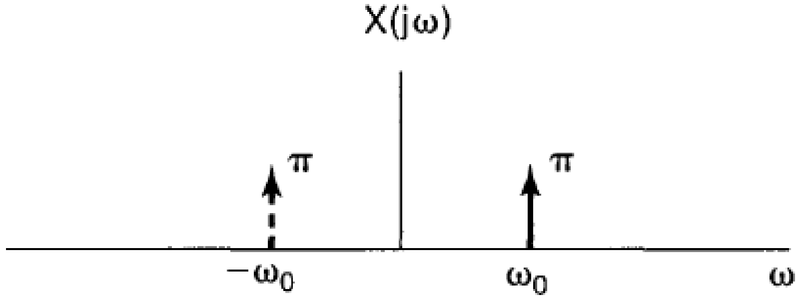
$$x_r(nT) = x(nT), \qquad n = 0, \pm 1, \pm 2, \ldots.$$

The effect of undersampling: Aliasing

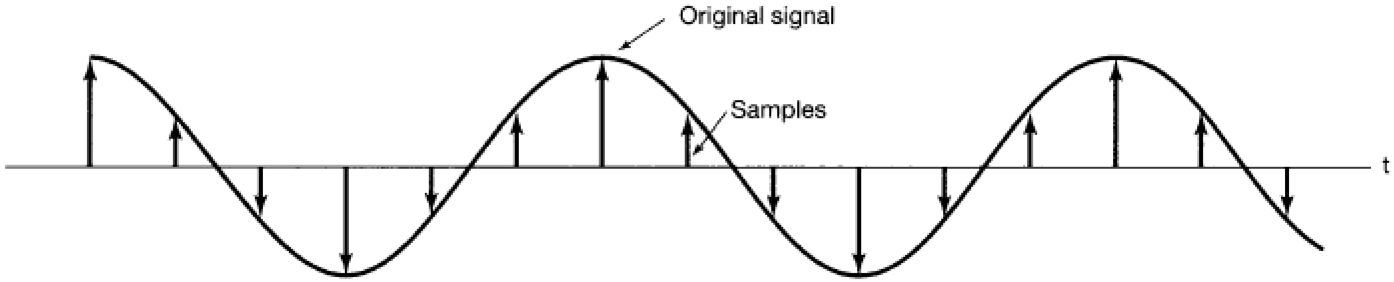
We will continue discussing aliasing with a specific example of the original signal $x(t)$. Let

$$x(t) = \cos \omega_0 t$$

with Fourier transform $X(j\omega)$.



Let sample this signal with the sampling frequency ω_s fixed and observe the effect of changing the frequency ω_0 of the cosine.

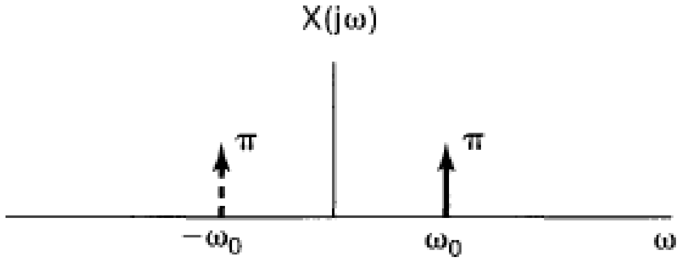


The effect of undersampling: Aliasing

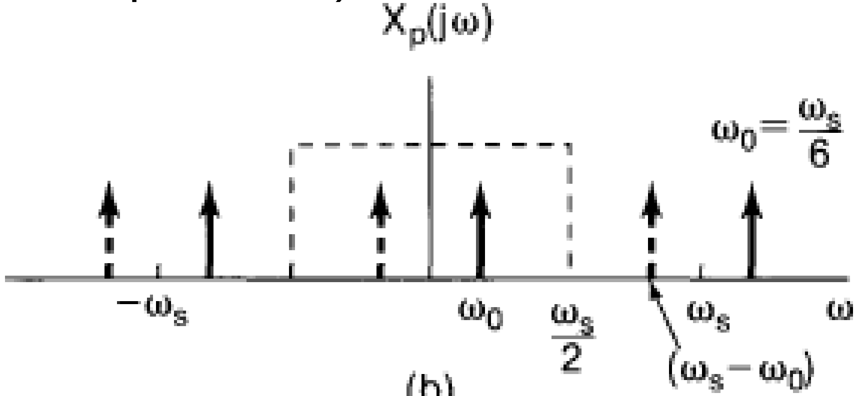
Let sample this signal with the sampling frequency ω_s fixed and observe the effect of changing the frequency ω_0 of the cosine.

$$\omega_0 = \frac{\omega_s}{6} \qquad \omega_s > 2\omega_M$$

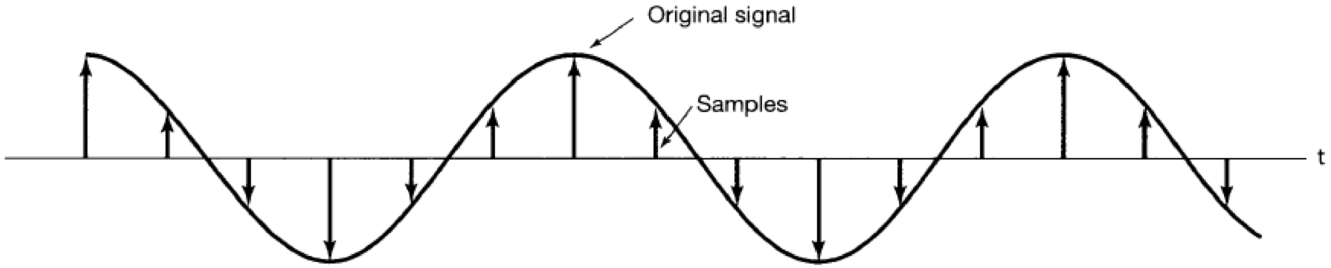
Spectrum of the original signal:



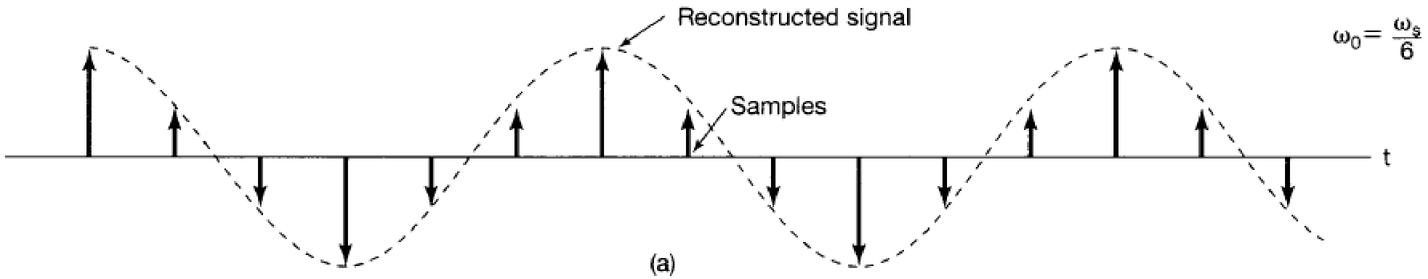
Spectrum of the sampled signal (dashed line is the passband of the lowpass filter):



The original sinusoidal signal (solid curve) and its samples:



The reconstructed signal (dashed line):



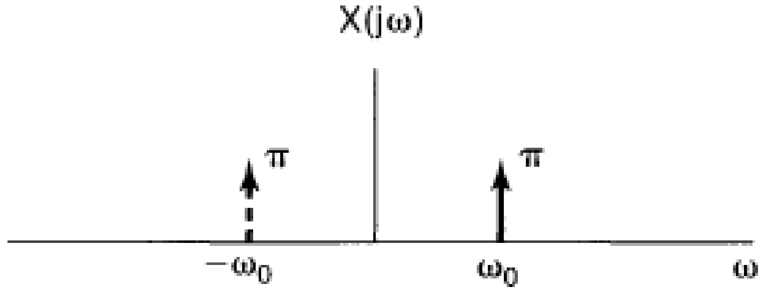
$$x_r(t) = \cos \omega_0 t = x(t)$$

No aliasing occurs in this case.

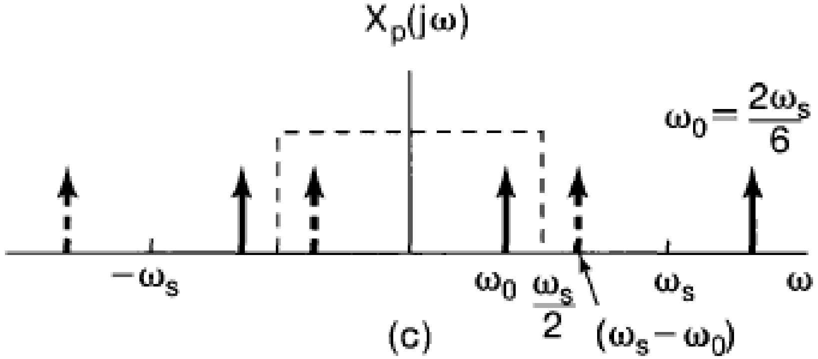
The effect of undersampling: Aliasing

$\omega_0 = \frac{2\omega_s}{6}$ $\qquad \omega_s > 2\omega_M$

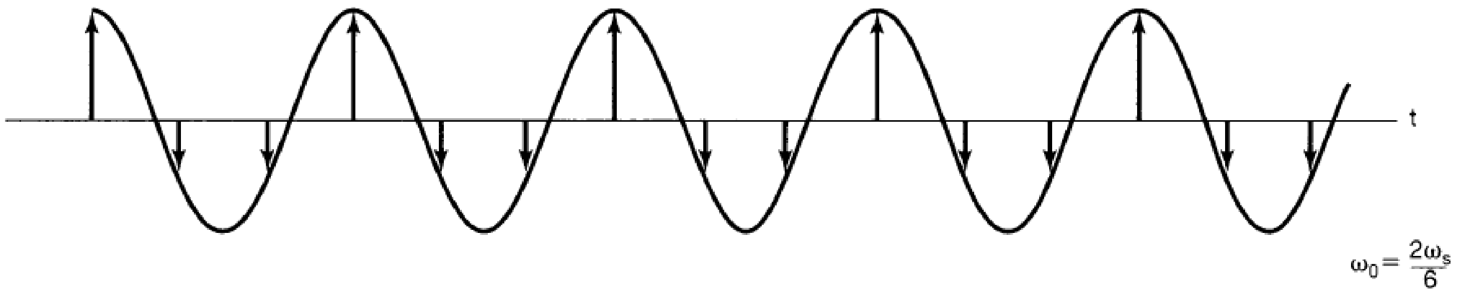
Spectrum of the original signal:



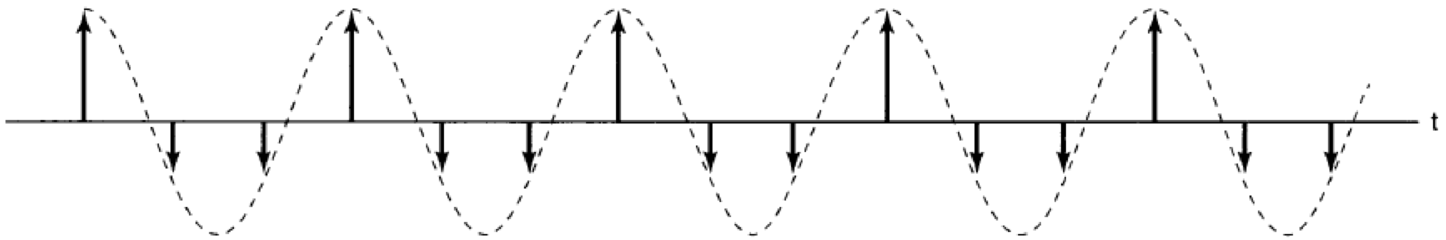
Spectrum of the sampled signal (dashed line is the passband of the lowpass filter):



The original sinusoidal signal (solid curve) and its samples:



The reconstructed signal (dashed line):



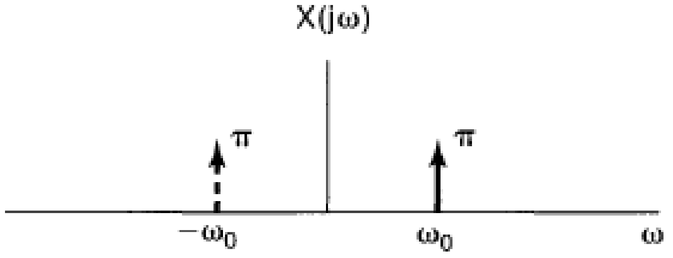
$$x_r(t) = \cos \omega_0 t = x(t)$$

No aliasing occurs in this case.

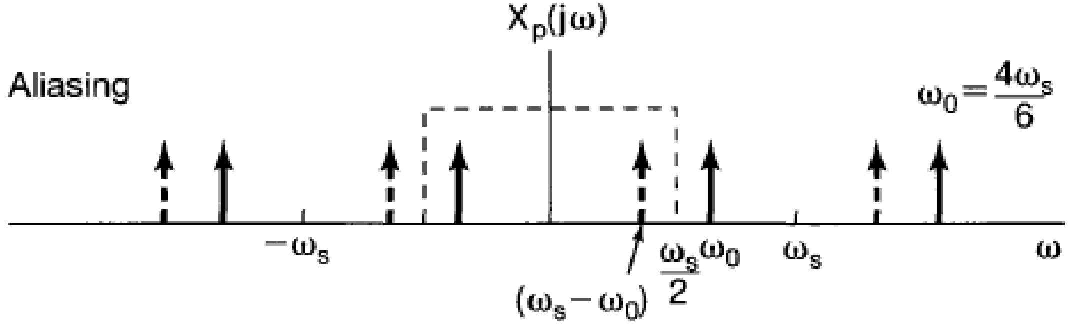
The effect of undersampling: Aliasing

$$\omega_0 = \frac{4\omega_s}{6} \qquad \omega_s < 2\omega_M$$

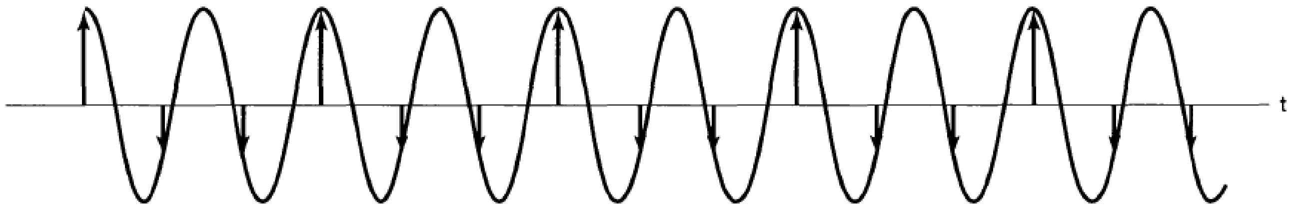
Spectrum of the original signal:



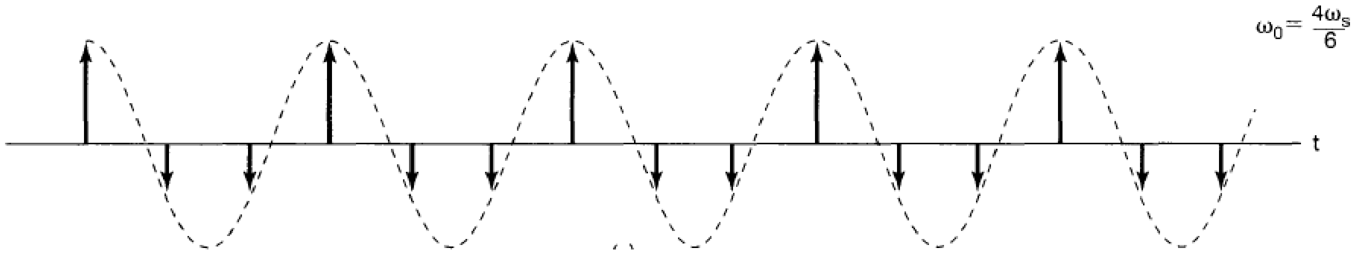
Spectrum of the sampled signal (dashed line is the passband of the lowpass filter). There is a change in the impulses falling within the passband of the ideal lowpass filter.



The original sinusoidal signal (solid curve) and its samples:



The reconstructed signal (dashed line):



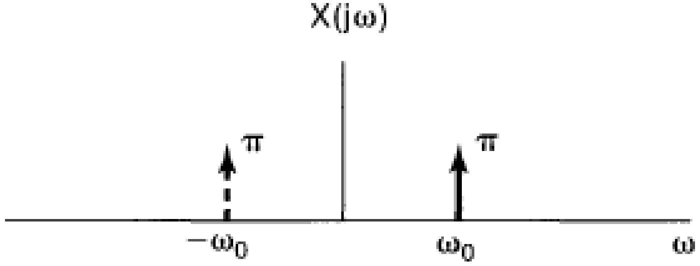
$$x_r(t) = \cos(\omega_s - \omega_0)t \neq x(t)$$

Aliasing!

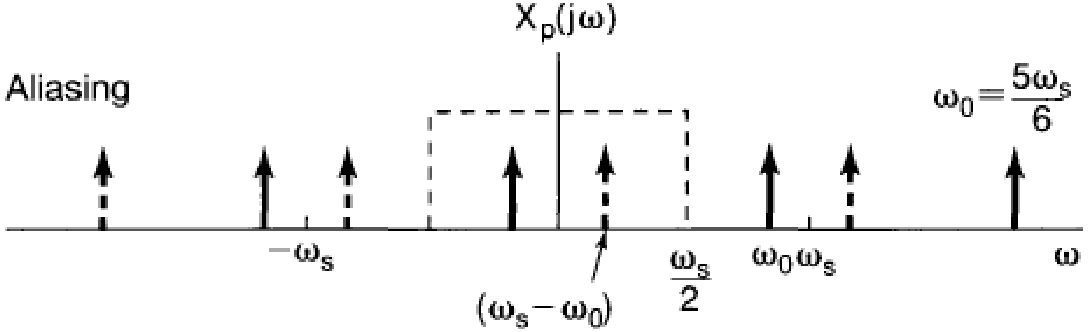
The effect of undersampling: Aliasing

$$\omega_0 = \frac{5\omega_s}{6} \qquad \omega_s < 2\omega_M$$

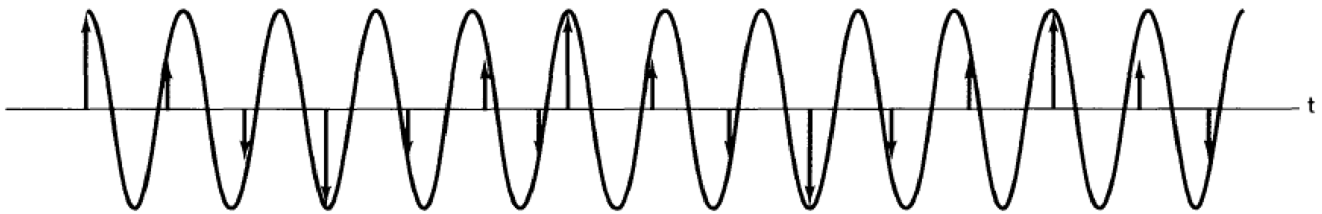
Spectrum of the original signal:



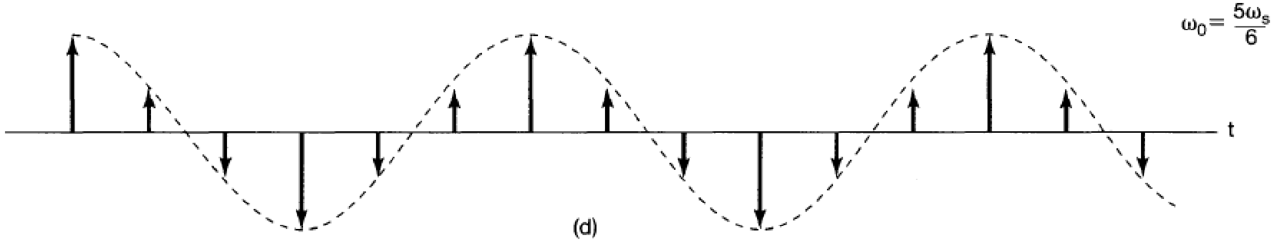
Spectrum of the sampled signal (dashed line is the passband of the lowpass filter). There is a change in the impulses falling within the passband of the ideal lowpass filter.



The original sinusoidal signal (solid curve) and its samples:



The reconstructed signal (dashed line):



$$x_r(t) = \cos(\omega_s - \omega_0)t \neq x(t)$$

Aliasing!

The effect of undersampling: Aliasing

When aliasing occurs, the original frequency ω_0 takes on the identity of a lower frequency, $\omega_s - \omega_0$.

For $\omega_s/2 < \omega_0 < \omega_s$, as ω_0 increases relative to ω_s, the output frequency $\omega_s - \omega_0$ decreases.

When $\omega_s = \omega_0$, for example, the reconstructed signal is a constant.

This is consistent with the fact that, when sampling once per cycle, the samples are all equal and would be identical to those obtained by sampling a constant signal.

The effect of undersampling: Aliasing

The sampling theorem explicitly requires that the sampling frequency be greater than twice the highest frequency in the signal, rather than greater than or equal to twice the highest frequency.

$$\omega_s > 2\omega_M$$

Sampling a signal at exactly twice its highest frequency is NOT sufficient.

Example

Consider the sinusoidal signal

$$x(t) = \cos\left(\frac{\omega_s}{2}t + \phi\right)$$

Suppose that this signal is sampled, using impulse sampling, at exactly twice the frequency of the sinusoid, i.e., at sampling frequency ω_s.
You can show that if this impulse-sampled signal is applied as the input to an ideal lowpass filter with cutoff frequency $\omega_s/2$, the resulting output is

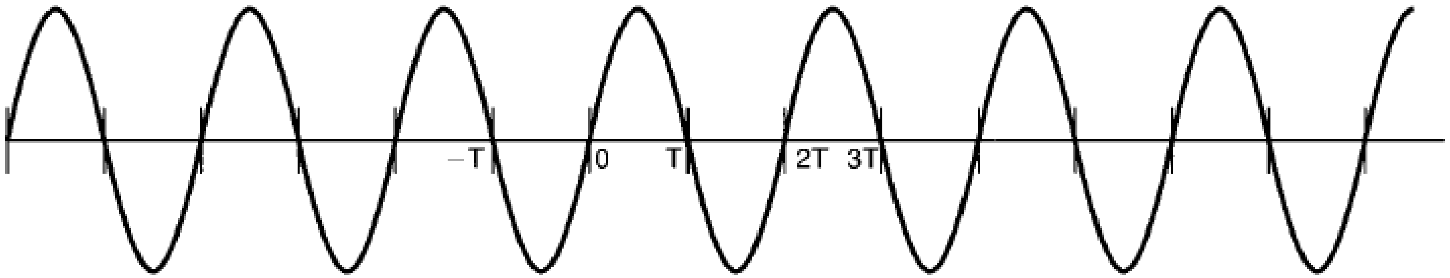
$$x_r(t) = (\cos\phi)\cos\left(\frac{\omega_s}{2}t\right)$$

We see that perfect reconstruction of $x(t)$ occurs only in the case in which the phase ∅ is zero (or an integer multiple of 2π).
Otherwise, the reconstructed signal $x_r(t)$ does not equal $x(t)$.

Example - continued

Consider the case in which $\emptyset = -\pi/2$, so that

$$x(t) = \cos\left(\frac{\omega_s}{2}t + \phi\right) = \sin\left(\frac{\omega_s}{2}t\right)$$



Sampling at the rate ω_s produces a signal that is identically zero, and when this zero input is applied to the ideal lowpass filter, the resulting output $x_r(t)$ is also identically zero.

$$x_r(t) = 0$$