Imperial College London

Signals and Systems

DR TANIA STATHAKI

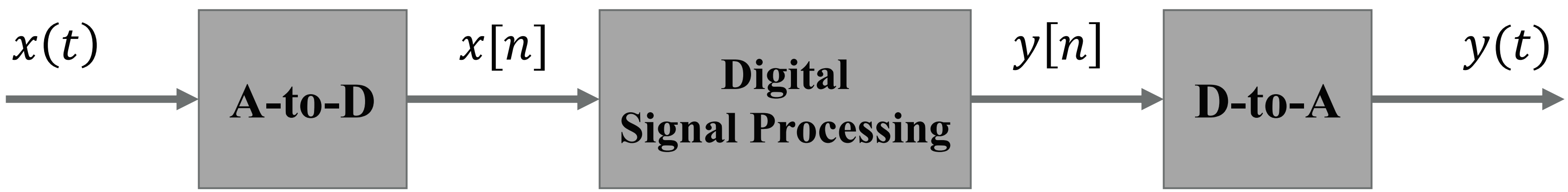
READER (ASSOCIATE PROFESSOR) IN SIGNAL PROCESSING

IMPERIAL COLLEGE LONDON

Continuous time versus discrete time

- **Continuous time systems.**
  - Good for analogue signals and general understanding of signals and systems.
  - Appropriate mostly to analogue electronic systems.

- **Electronic devices are increasingly digital.**
  - E.g. mobile phones are all digital, TV broadcast will be 100% digital in UK.
  - We use digital ASIC chips, FPGAs and microprocessors to implement systems and to process signals.
  - Continuous signals are converted to numbers (discrete signals), they are processed and then converted back to continuous signals.

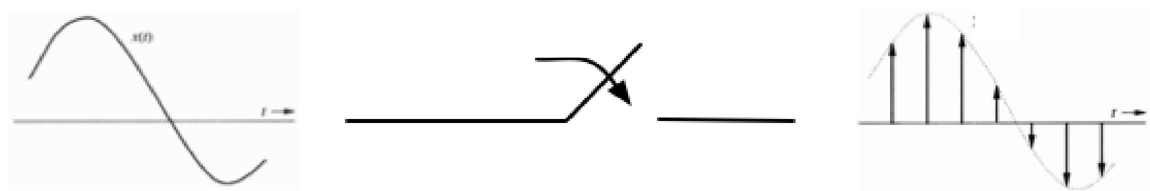


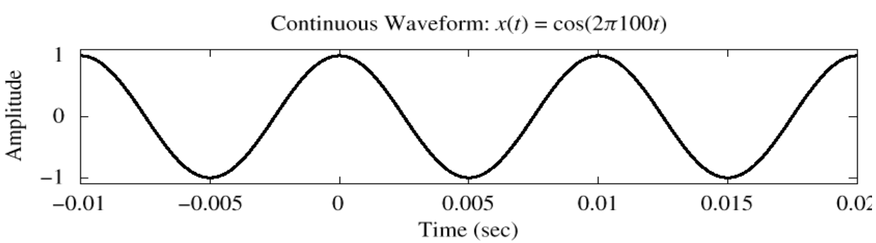
Imperial College London

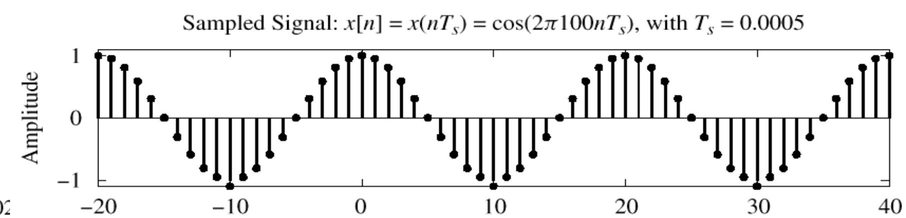
Sampling

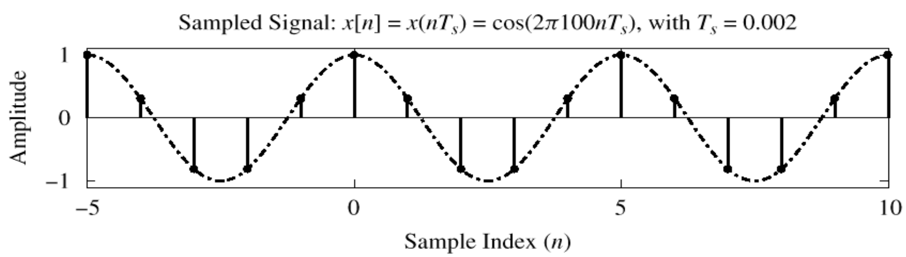
- The sampling process converts continuous signals $x(t)$ into a sequence of numbers $x[n]$.
- A sample is kept every T_s units of time. This process is called **uniform sampling** and $x[n] = x(nT_s)$.

$x(t)$ $\quad\quad$ $T_s = 1/f_s$ $\quad\quad$ $x[n]$







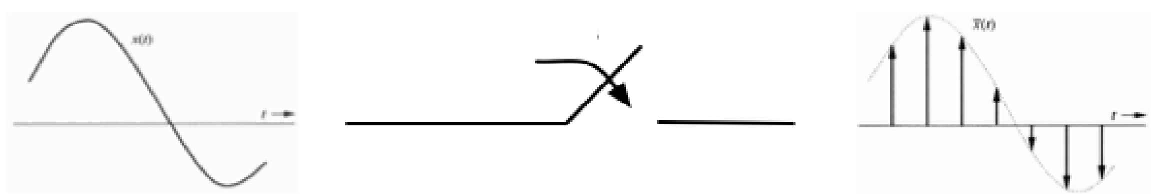


Imperial College London

Sampling theorem

- **Sampling theorem** is the bridge between continuous-time and discrete-time signals.
- It states how often we must sample in order not to loose any information.

$x(t)$ $\qquad$ $T_s = 1/f_s$ $\qquad$ $x[n]$



Sampling theorem

A continuous-time lowpass signal $x(t)$ with frequencies no higher than $f_{max} Hz$ can be perfectly reconstructed from samples taken every T_s units of time, $x[n] = x(nT_s)$, if the samples are taken at a rate $f_s = 1/T_s$ that is greater than $2f_{max} Hz$.

Sampling theorem cont.

- If a lowpass signal has a spectrum bandlimited to BHz, i.e., $X(\omega) = 0$ for $|\omega| > 2\pi B$, it can be reconstructed from its samples without error if these samples are taken uniformly at a rate $f_s > 2B$ samples per second.
- The minimum sampling rate $f_s = 2B$ required to reconstruct $x(t)$ from its samples is called the **Nyquist rate** for $x(t)$ and the corresponding sampling interval $T_s = \frac{1}{2B}$ is called the **Nyquist interval**. Samples of a continuous signal taken at its Nyquist rate are the **Nyquist samples** of that signal.
- In other words the minimum sampling frequency is $f_s = 2BHz$.
- A bandpass signals whose spectrum exists over a frequency band $f_c - \frac{B}{2} < |f| < f_c + \frac{B}{2}$ also has a bandwidth of BHz. Such a signal is still uniquely determined by $2B$ samples per second but the sampling scheme is a bit more complex compared to the case of a lowpass signal.

Sampling theorem: mathematical proof

- Consider a signal bandlimited to BHz with Fourier transform $X(\omega)$.
- The sampled version of the signal $x(t)$ at a rate $f_s Hz$ can be expressed as the multiplication of the original signal with an impulse train as follows:

$$\bar{x}(t) = x(t)\delta_{T_s}(t) = \sum_n x(nT_s)\delta(t - nT_s),\ T_s = 1/f_s$$

- We can express the impulse train using Fourier series as follows:

$$\delta_{T_s}(t) = \frac{1}{T_s}[1 + 2\cos\omega_s t + 2\cos 2\omega_s t + 2\cos 3\omega_s t + \cdots],\ \omega_s = \frac{2\pi}{T_s} = 2\pi f_s$$

- Therefore,

$$\bar{x}(t) = x(t)\delta_{T_s}(t) = \frac{1}{T_s}[x(t) + 2x(t)\cos\boldsymbol{\omega_s} t + 2x(t)\cos\boldsymbol{2\omega_s} t + 2\cos\boldsymbol{3\omega_s} t + \cdots]$$

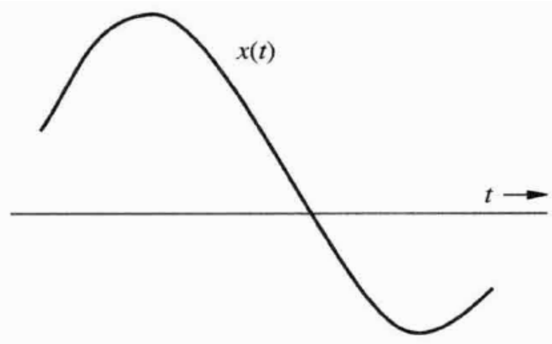
- Since the following holds: $x(t)\cos\omega_s t \Leftrightarrow \frac{1}{2}[X(\omega + \omega_s) + X(\omega - \omega_s)]$

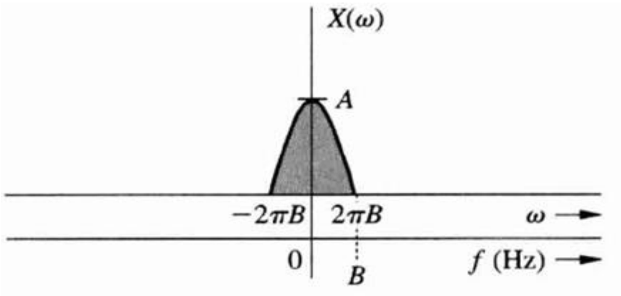
  we have

$$\bar{X}(\omega) = \frac{1}{T_s}\sum_{n=-\infty}^{\infty} X(\omega - n\omega_s)$$

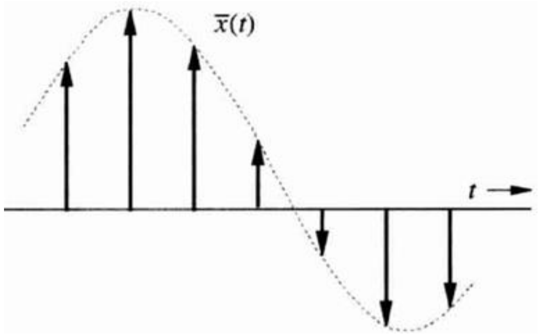
Depiction of previous analysis

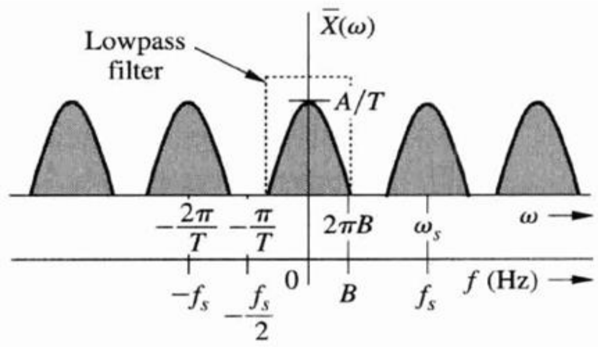
- The previous analysis is depicted below. Consider a signal, bandlimited to BHz, with Fourier transform $X(\omega)$ (depicted real for convenience).





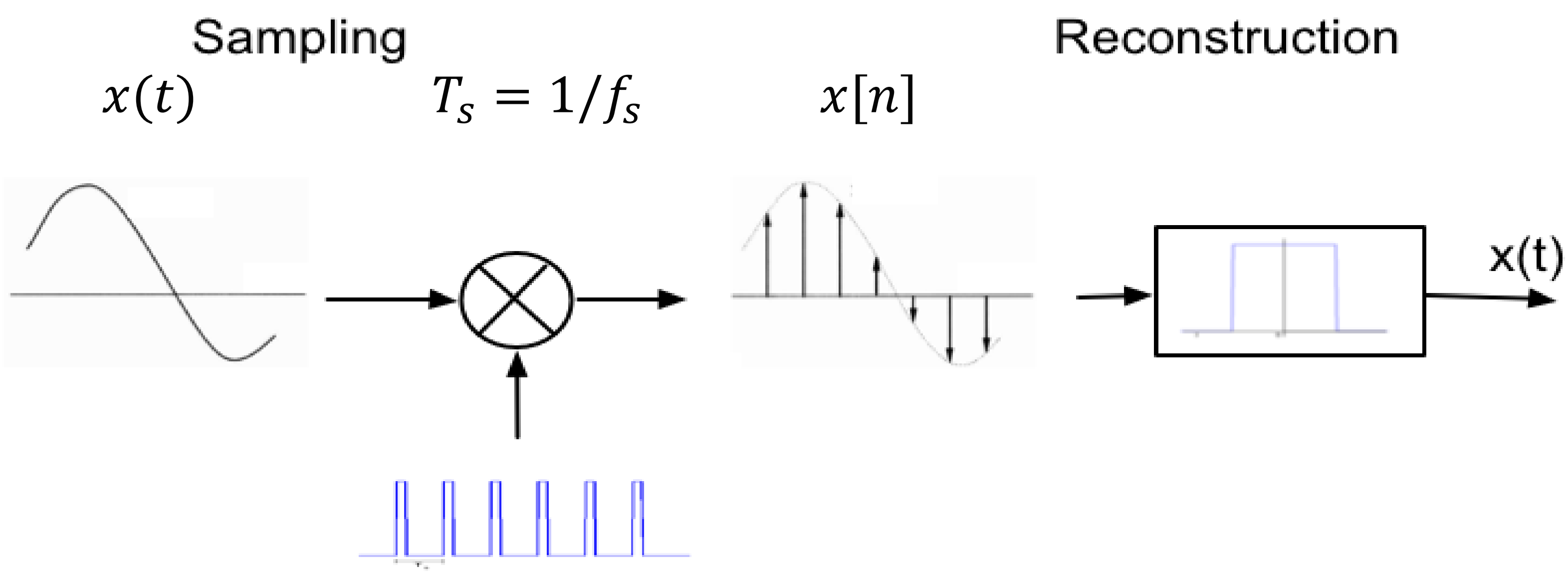
- The sampled signal has the following spectrum.





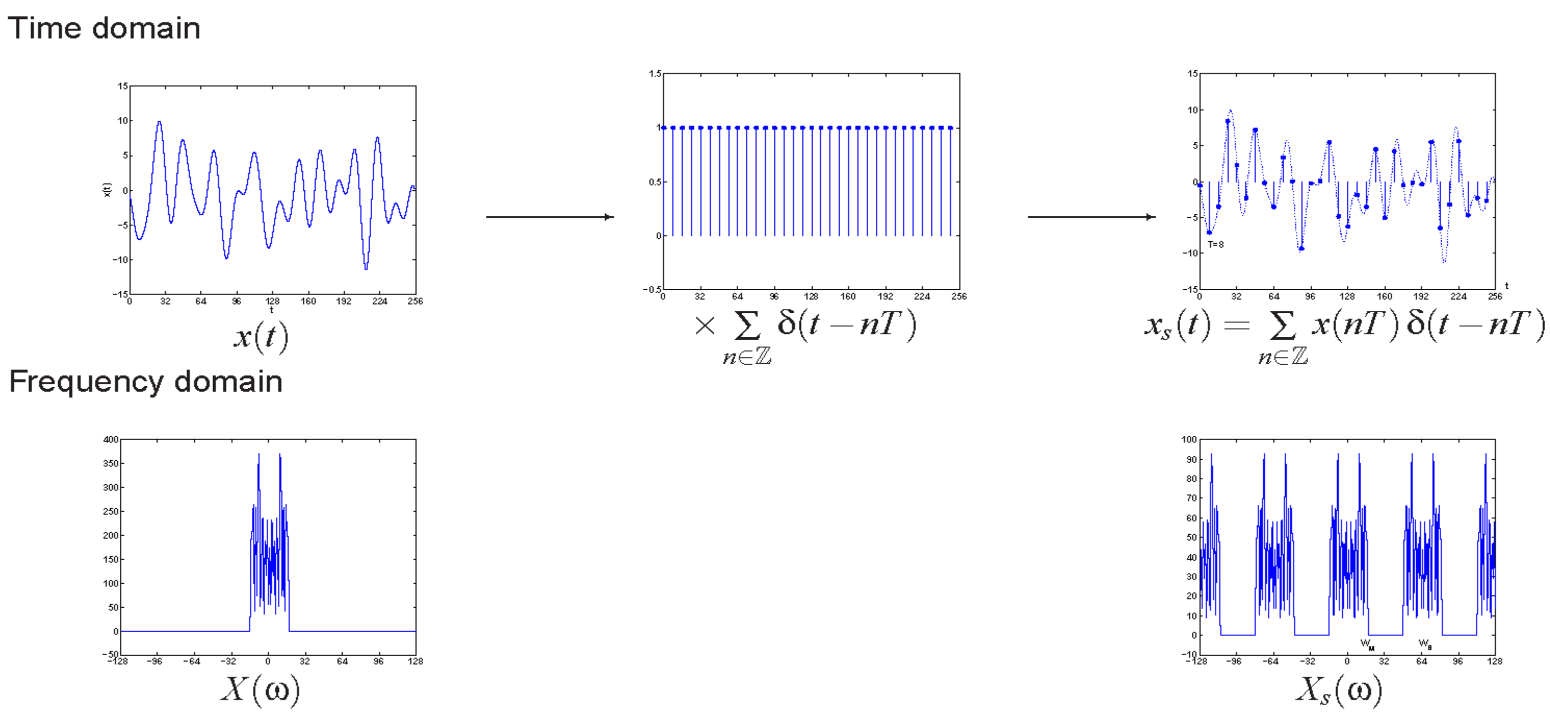
Depiction of previous analysis

- We graphically illustrate below the collection of the above mentioned processes.
  - The signal is multiplied by a train of impulses (in reality these are very narrow pulses).
  - The sampled signal is generated.
  - A lowpass filter is required in order to isolate the main period of the Fourier spectrum.



Depiction of previous analysis cont.

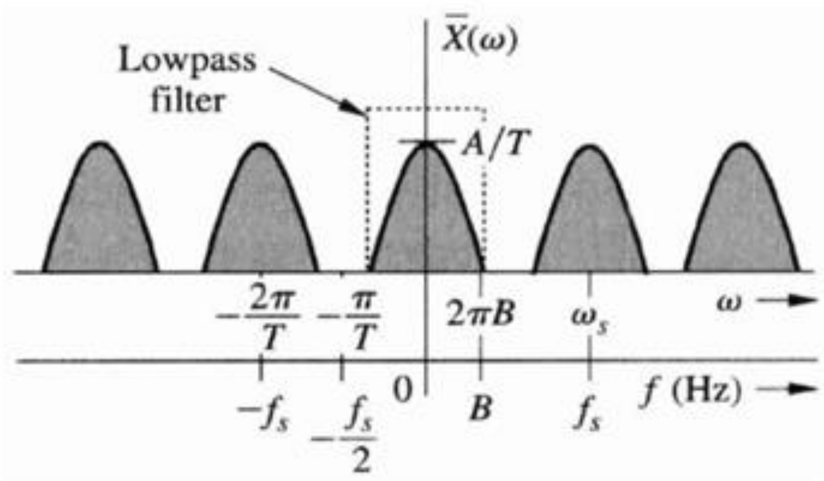
- A continuous signal with its Fourier transform is depicted on the left.
- The sampled signal and its Fourier transform is depicted on the right.
  - **Note that, for simplicity, the Fourier transform is considered to be real in this case.**



Imperial College London

Reconstruction of the original signal

- The gap between two adjacent spectral repetitions is $(f_s - 2B)Hz$.
- In order to reconstruct the original signal $x(t)$ we can use an ideal lowpass filter on the sampled spectrum which has a bandwidth of any value between B and $(f_s - BHz)$.

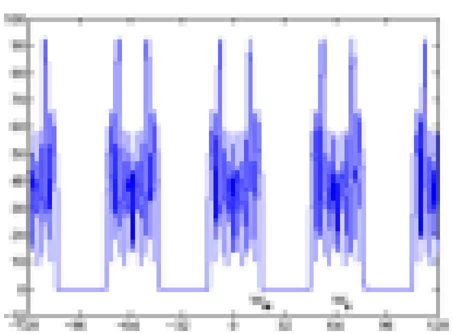


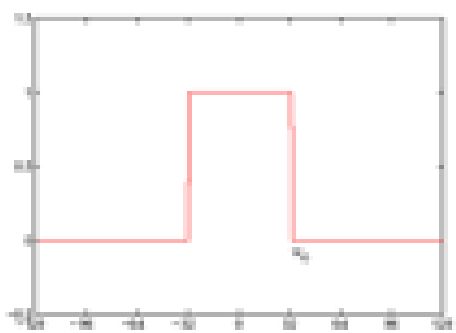
- Reconstruction process is possible only if the shaded parts do not overlap. This means that f_s must be greater that twice B.
- We can also visually verify the sampling theorem in the above figure.

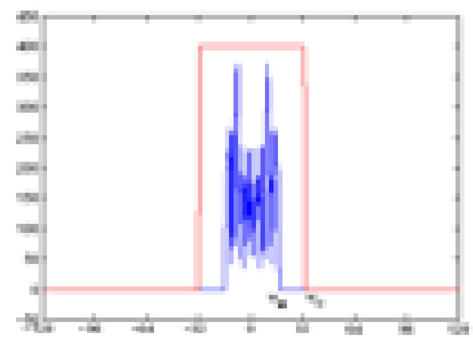
Reconstruction generic example

- The signal $x(nT_s)$ has a spectrum $X_s(\omega)$ which is multiplied with a rectangular pulse in frequency domain in order to isolate the main period which is the spectrum of the original continuous time signal.
    - **Note that, for simplicity, the Fourier transform is considered to be real in this case.**

$X_s(\omega)$





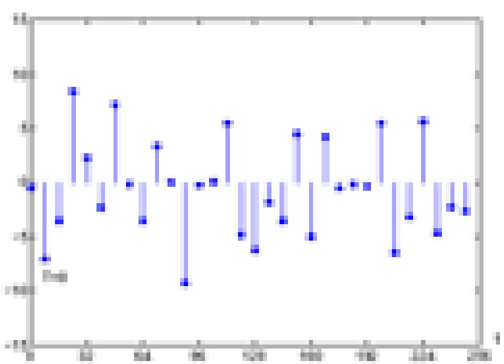


Reconstruction generic example cont.

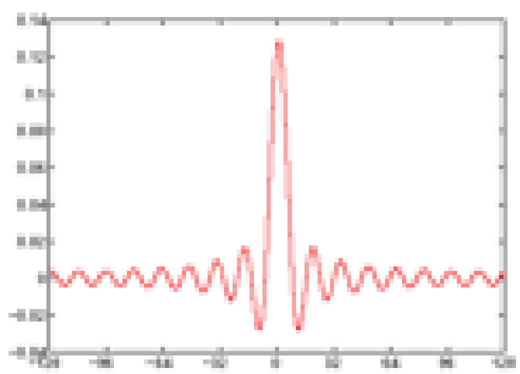
- The signal $x(nT_s)$ is convolved with a sinc function, which is the time domain version of a rectangular pulse in frequency domain centred at the origin.

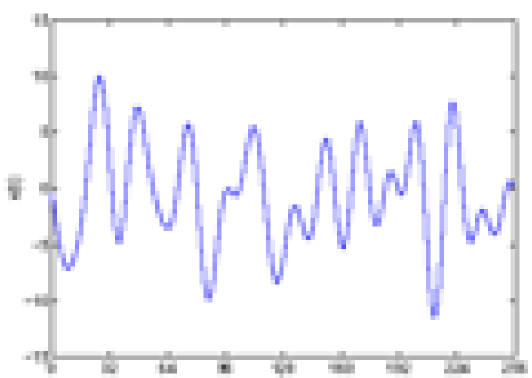
$$x(nT_s) * \mathrm{sinc}(at) = \sum_n x[n]\,\mathrm{sinc}[a(t - nT_s)]$$

$x(nT_s)$



$\mathrm{sinc}(at)$

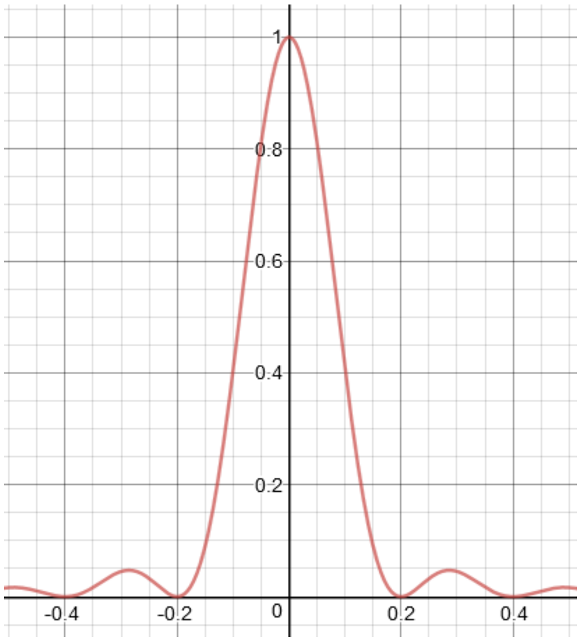


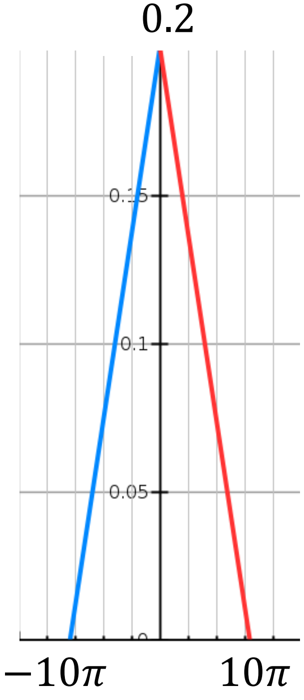


Imperial College London

Example

- Consider the signal $x(t) = \mathrm{sinc}^2(5\pi t)$.
- The property $(\frac{W}{2\pi})\mathrm{sinc}^2(Wt/2) \Leftrightarrow \Delta(\omega/2W)$ holds with:
- $\Delta(x) = \begin{cases} 0 & |x| \geq \frac{1}{2} \\ 1 - 2|x| & |x| < \frac{1}{2} \end{cases}$
- Using the above property with $W = 10\pi$ we obtain $X(\omega) = 0.2\Delta(\omega/20\pi)$
- The bandwidth is $B = 5Hz(10\pi rad/s)$.
- Consequently, the Nyquist sampling rate is $f_s = 10Hz$; we require at least 10 samples per second.
- The Nyquist interval is $T_s = \frac{1}{10} = 0.1s$.

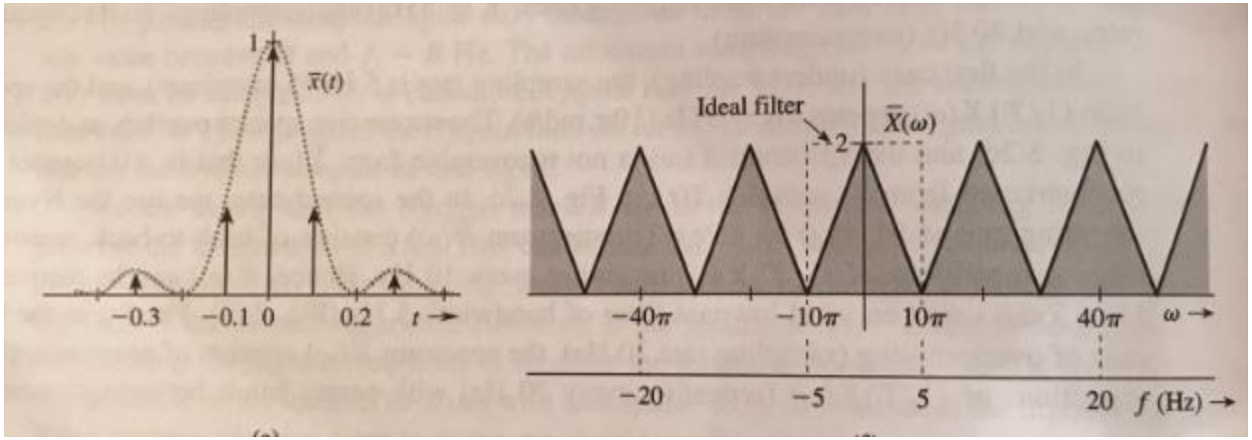




Imperial College London

Example cont.
Nyquist sampling: Just about the correct sampling rate

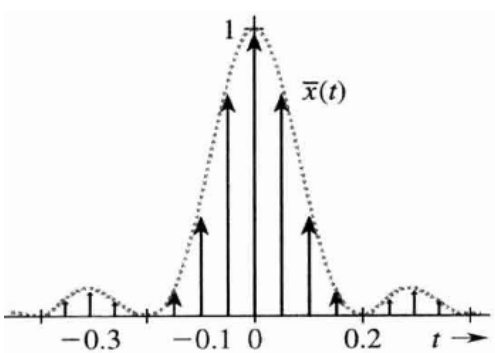
- In that case we use the Nyquist sampling rate of $10Hz$.
- The spectrum $\bar{X}(\omega)$ consists of back-to-back, non-overlapping repetitions of $\frac{1}{T_s}X(\omega)$ repeating every $f_s = 2B = 10Hz$.
- In order to recover $X(\omega)$ from $\bar{X}(\omega)$ we must use an ideal lowpass filter of bandwidth $5Hz$. This is shown in the right figure below with the dotted line.
- **Obviously, it is not possible to design and implement such a filter.**

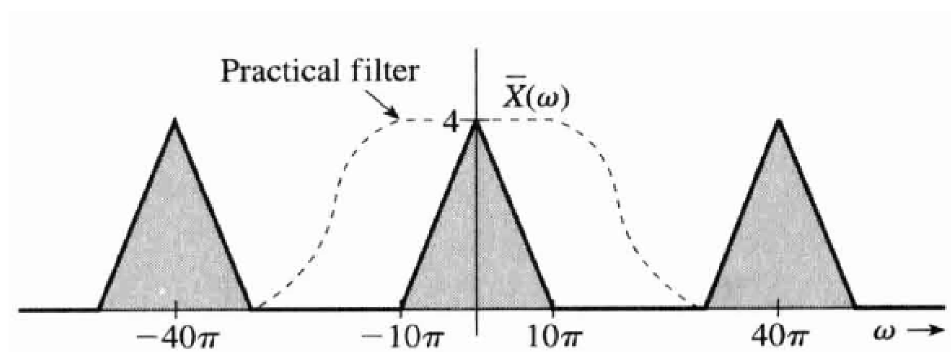


Example cont.

Oversampling: What happens if we sample too quickly?

- Sampling at higher than the Nyquist rate (in this case $20Hz$) makes reconstruction easier.
- The spectrum $\bar{X}(\omega)$ consists of non-overlapping repetitions of $\frac{1}{T_s}X(\omega)$, repeating every $20Hz$ with empty bands between successive cycles.
- In order to recover $X(\omega)$ from $\bar{X}(\omega)$ we can use a **practical** lowpass filter and not necessarily an ideal one. This is shown in the right figure below with the dotted line.





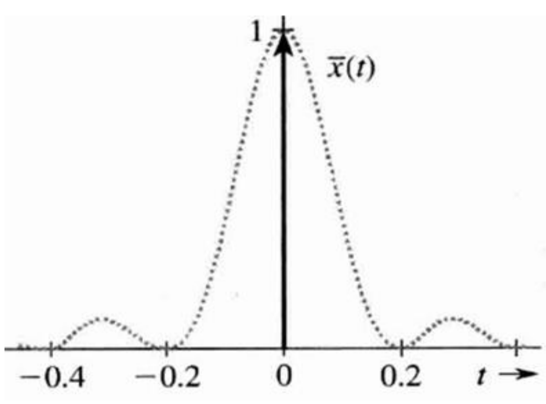
- The filter we use for reconstruction must have gain T_s and bandwidth of any value between B and $(f_s - B)Hz$.

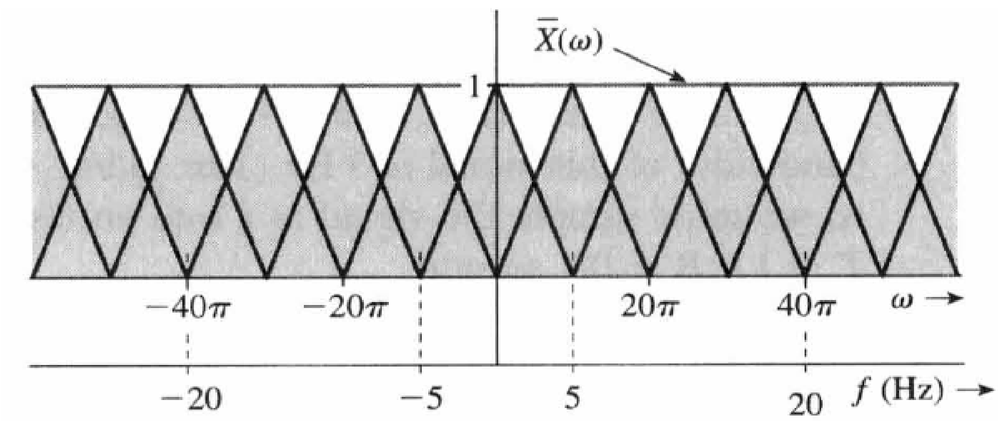
Imperial College London

Example cont.

Undersampling: What happens if we sample too slowly?

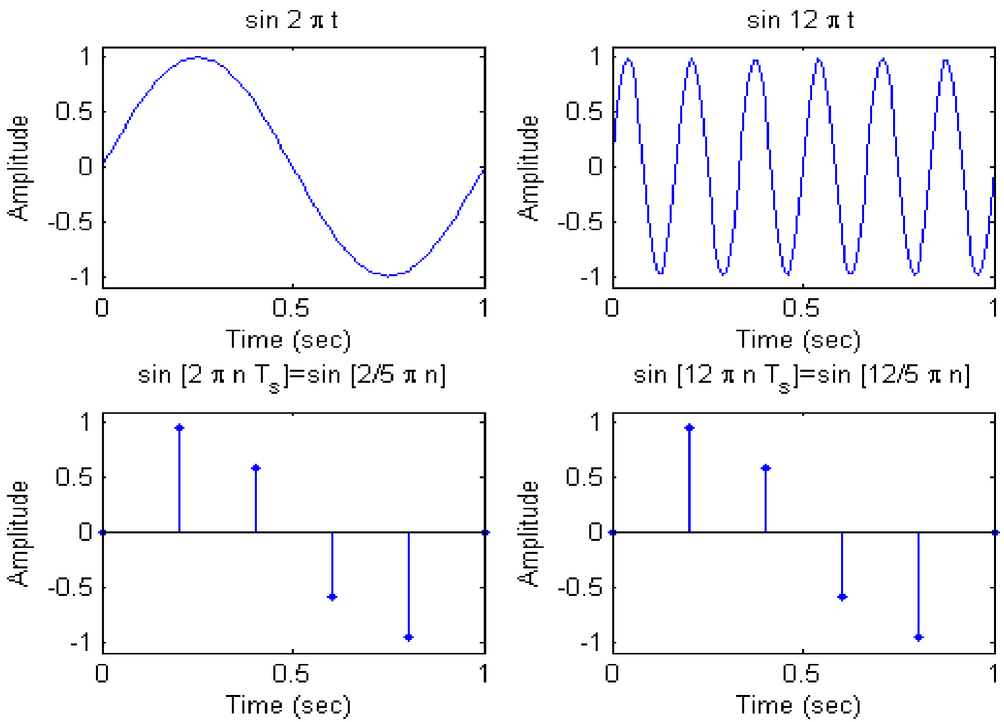
- Sampling at lower than the Nyquist rate (in this case $5Hz$) makes reconstruction impossible.
- The spectrum $\bar{X}(\omega)$ consists of overlapping repetitions of $\frac{1}{T_S}X(\omega)$ repeating every $5Hz$.
- $X(\omega)$ is not recoverable from $\bar{X}(\omega)$.
- Sampling below the Nyquist rate corrupts the signal. This type of distortion is called **aliasing**.





Aliasing

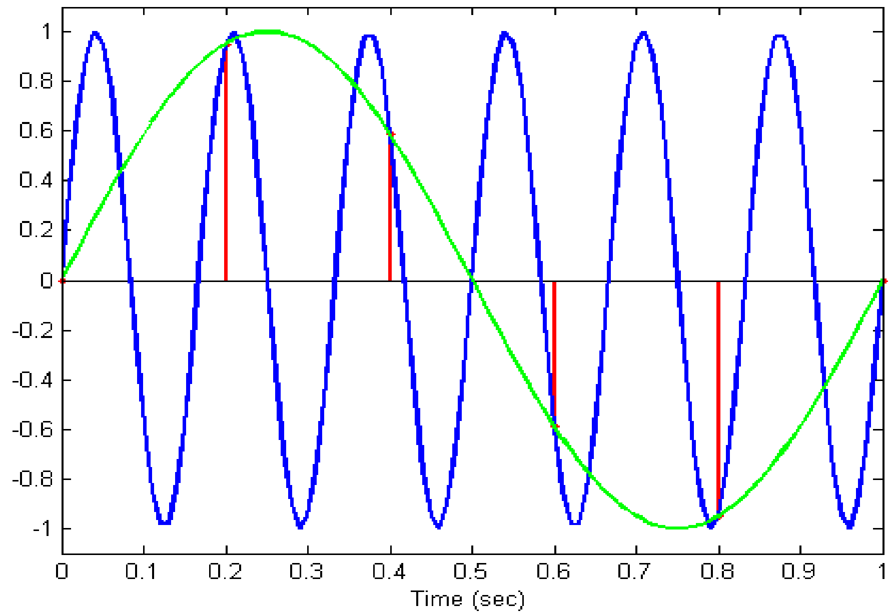
- Consider what happens when a $1Hz$ and a $6Hz$ sinewaves are sampled at a rate of 5Hz.



- The $1Hz$ and $6Hz$ sinewaves are indistinguishable after sampling. The two discrete signals produced are identical.

Aliasing cont.

- The two original signals are shown together with the sampled values. As verified, the sampled values are not efficient to recover the original shape of the high frequency signal (shown in blue).



Aliasing example (recall Signals and Systems Lab)

- Consider a signal bandlimited to BHz with Fourier transform $X(\omega)$. The sampled version of the signal $x(t)$ at a rate $f_s Hz$ can be expressed as the multiplication of the original signal with an impulse train as follows:
  $\bar{x}(t) = x(t)\delta_{T_s}(t) = \sum_n x(nT_s)\delta(t - nT_s),\, T_s = 1/f_s$

- Assume that $x(t) = \sin(2\pi f t)$. Consider two frequencies f_1 and f_2 which satisfy the condition $-f_1 + f_2 = f_s$. **This happens in the previous example.**

- Take the signal $x(t) = \sin(2\pi f_1 t)$. Its sampled version is
  $\bar{x}(t) = x(t)\delta_{T_s}(t) = \sum_n x(nT_s)\delta(t - nT_s),\, T_s = 1/f_s$
  $x(nT_s) = \sin(2\pi f_1 nT_s) = \sin[2\pi(-f_s + f_2)nT_s] = \sin(2\pi f_2 nT_s - 2\pi f_s nT_s)$

- We know that $f_s T_s = 1$. Therefore, the above becomes:
  $$x(nT_s) = \sin(2\pi f_2 nT_s - 2\pi n) = \sin(2\pi f_2 nT_s)$$

- We see that if $f_2 = f_s + f_1$ the two sinusoidals look the same after sampling.

Aliasing and the wagon wheel effect

- Consider making a video of a clock face.
- The second hand makes one revolution per minute ($f_{max} = \frac{1}{60} Hz$).
- Critical Sampling: $f_s > \frac{1}{30} Hz$ ($T_s < 30\text{sec}$), anything below that sampling frequency will create problems.

  For example:
  - When $T_s = 60\text{sec}$ ($f_s \sim 0.167 Hz$), the second hand will not move.
  - When $T_s = 59\text{sec}$ ($f_s \sim 0.169 Hz$), the second hand will move backwards.

- Watch the videos (optional)

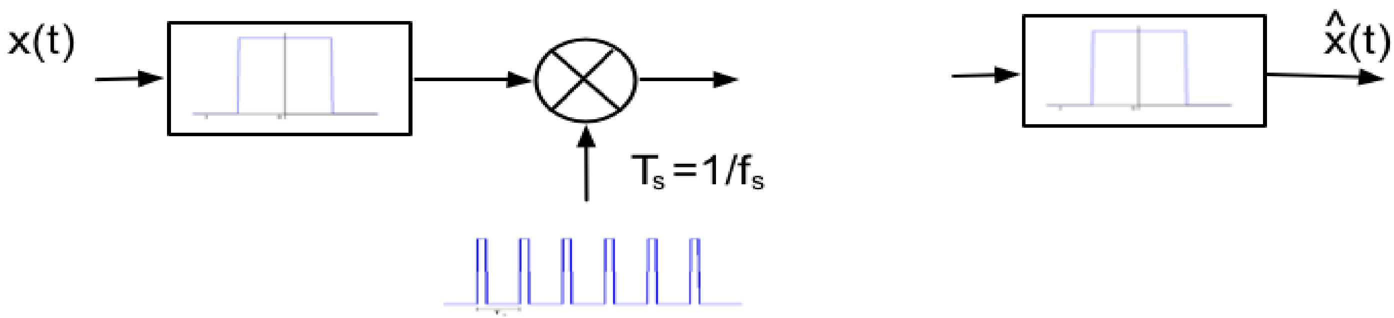
  https://www.youtube.com/watch?v=VNftf5qLpiA

  https://www.youtube.com/watch?v=QOwzkND_ooU

Anti-aliasing filter

- To avoid distortion of a signal after sampling, one must ensure that the signal being sampled at f_s is bandlimited to a frequency B, where $B < \frac{f_s}{2}$.
- If the signal does not obey the above restriction, we may apply a lowpass filter with cut-off frequency $\frac{f_s}{2}$ before sampling.

Sampling **Reconstruction**



- If the original signal $x(t)$ is not bandlimited to $\frac{f_s}{2}$, perfect reconstruction is not possible when sampling at f_s. However, the reconstructed signal $\hat{x}(t)$ is the best bandlimited approximation to $x(t)$ in the least-square sense.

Practical sampling (Lathi page 778)

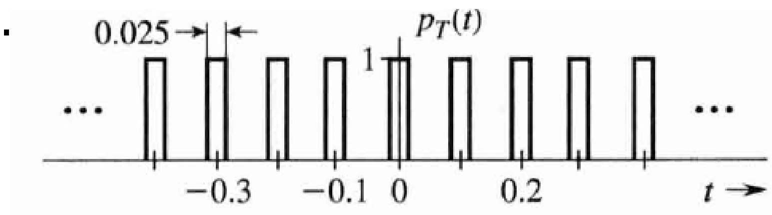
- The impulse train is not a very practical sampling signal.
- In practice we may use a train of pulses $p_T(t)$ as the one shown below.
  - The pulse height is $A = 1$, its width is $\tau = 0.025sec$ and the period is $T = 0.1\text{sec}$.
- We use the same signal as previously, i.e.,

$$x(t) = \text{sinc}^2(5\pi t) \text{ with } X(\omega) = 0.2\Delta(\omega/20\pi)$$

- For the train of pulses we use Fourier Series (Lathi, Chapter 6):

$$p_{T_s}(t) = C_0 + \sum_{n=1}^{\infty} C_n \cos n\omega_s t,\ C_0 = 1/4,\ C_n = \frac{2}{n\pi}\sin(\frac{n\pi}{4})$$

- $\bar{x}(t) = x(t)p_{T_s}(t) = \frac{1}{4}x(t) + C_1 x(t)\cos 20\pi t + C_2 x(t)\cos 40\pi t + C_3 x(t)\cos 60\pi t + \cdots$
- $\bar{X}(\omega) = \frac{1}{4}X(\omega) + \mathbf{2\pi}\{\frac{C_1}{2}[X(\omega - 20\pi) + X(\omega + 20\pi)] + \frac{C_2}{2}[X(\omega - 40\pi) + X(\omega + 40\pi)] + \frac{C_3}{2}[X(\omega - 60\pi) + X(\omega + 60\pi)] + \cdots\}$
- Lowpass filter can be used to recover $X(\omega)$.

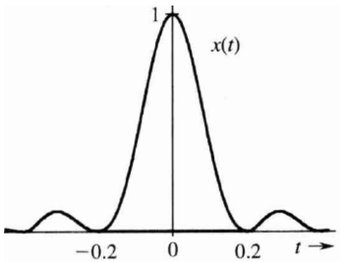


Imperial College London

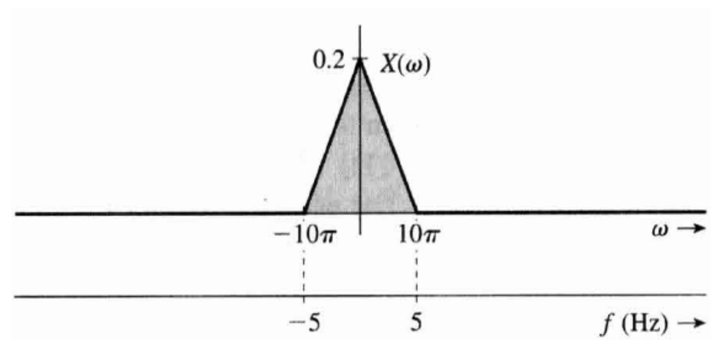
Practical sampling cont.

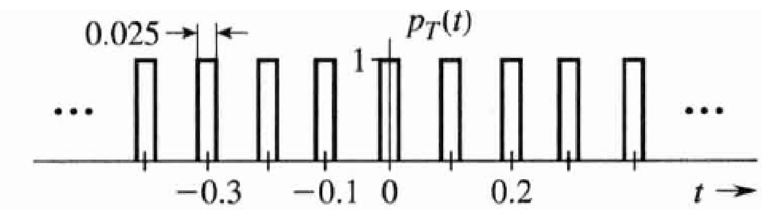
$x(t) = \text{sinc}^2(5\pi t)$

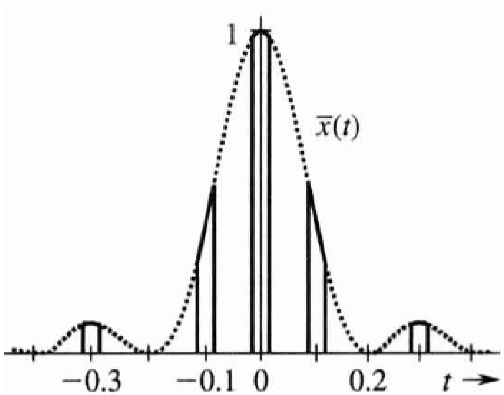
$X(\omega) = 0.2\Delta(\omega/20\pi)$



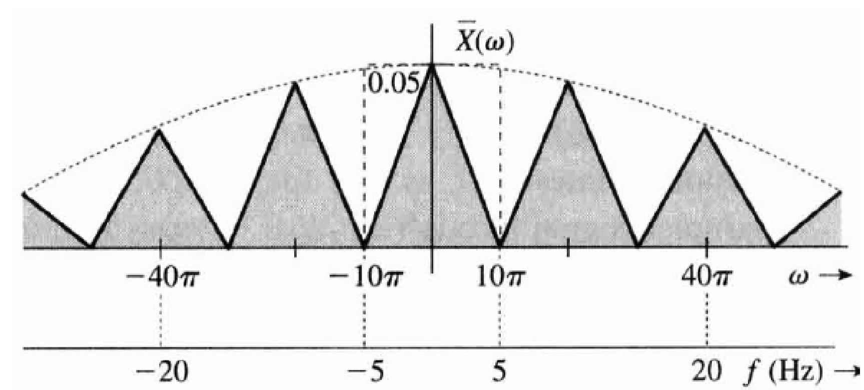
⇔







⇔



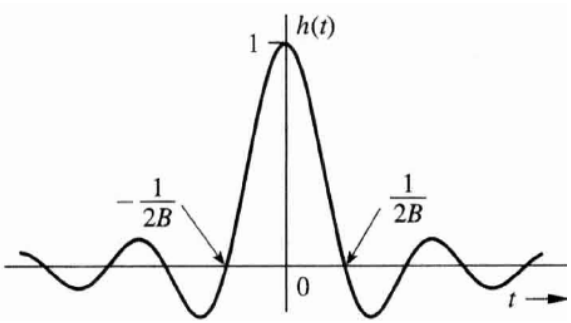
Ideal signal reconstruction

- The process of reconstructing a continuous-time signal $x(t)$ from its samples is also known as **interpolation**.
- As already mentioned, the filter we use for reconstruction must have gain T_s and bandwidth of any value between B and $(f_s - B) Hz$.
- A good choice is the middle value $\frac{f_s}{2} = \frac{1}{2T_s} Hz$ or $\frac{\pi}{T_s} rad/s$. This gives a filter with frequency response:

$$H(\omega) = T_s \text{rect}\left(\frac{\omega}{2\pi f_s}\right) = T_s \text{rect}\left(\frac{\omega T_s}{2\pi}\right)$$

- In time domain, the impulse response of the above filter is given by:

$$h(t) = \text{sinc}\left(\frac{\pi t}{T_s}\right)$$

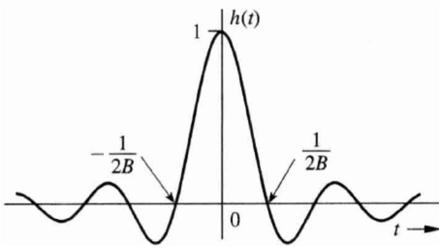


- For the Nyquist sampling rate, $T_s = \frac{1}{2B}$,

$$h(t) = \text{sinc}\,(2\pi B t)$$

Ideal signal reconstruction cont.

- Observe the interesting fact that $h(t) = 0$ for all Nyquist sampling instants $t = \pm\frac{n}{2B}$, except at $t = 0$.



- Each sample in $\bar{x}(t)$, being an impulse, generates a sinc pulse of height equal to the strength of the sample when is applied at the input of this filter (recall convolution with an impulse). Addition of all sinc impulses results in $x(t)$ (look at next slide).

- The filter output is:

$$x(t) = \sum_n x(nT_s)h(t - nT_s) = \sum_n x(nT_s)\mathrm{sinc}\left[\frac{\pi}{T_s}(t - nT_s)\right]$$

- In the case of Nyquist sampling rate $T_s = \frac{1}{2B}$. The above equation becomes:

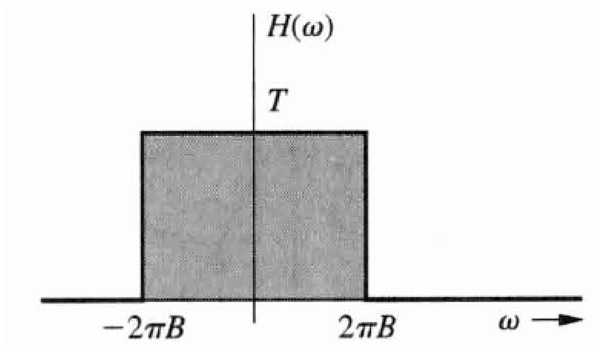
$$x(t) = \sum_n x(nT_s)h(t - nT_s) = \sum_n x(nT_s)\mathrm{sinc}(2\pi Bt - n\pi)$$

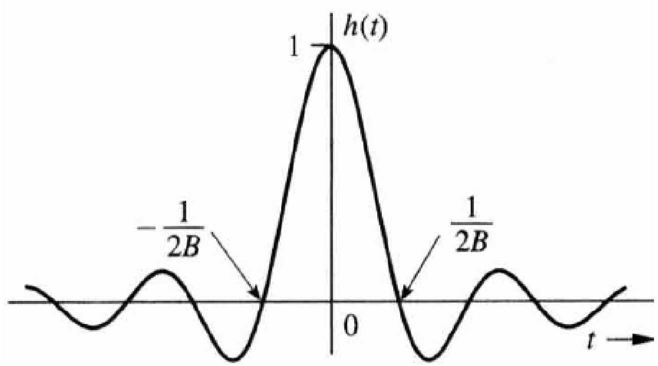
This is called the **interpolation formula**.

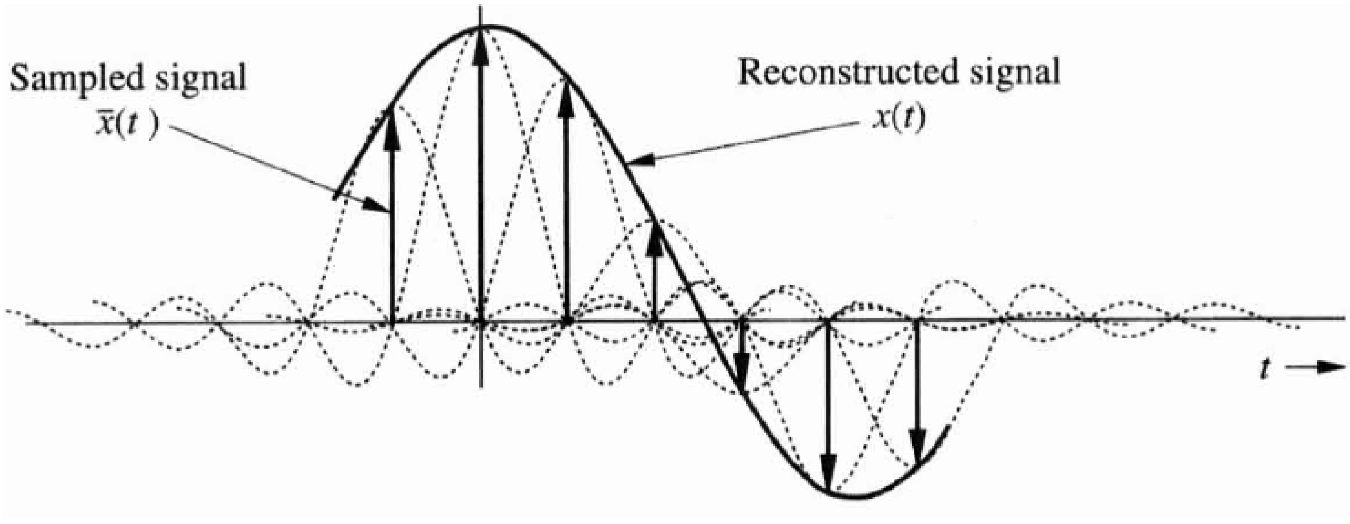
Imperial College London

Ideal signal reconstruction cont.

- The above analysis is depicted below.



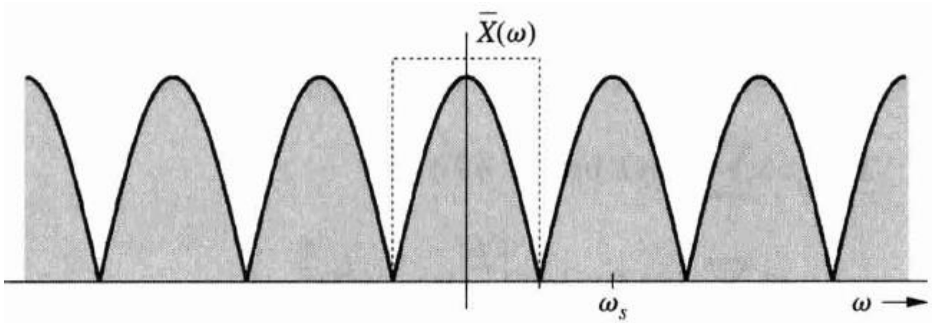


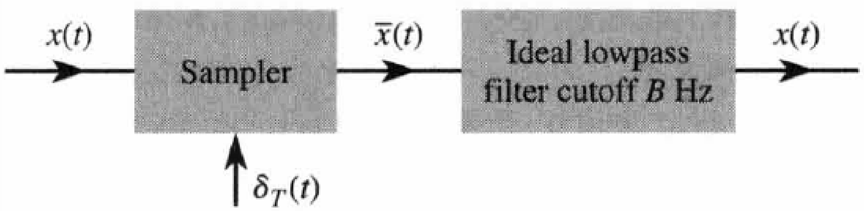


Imperial College London

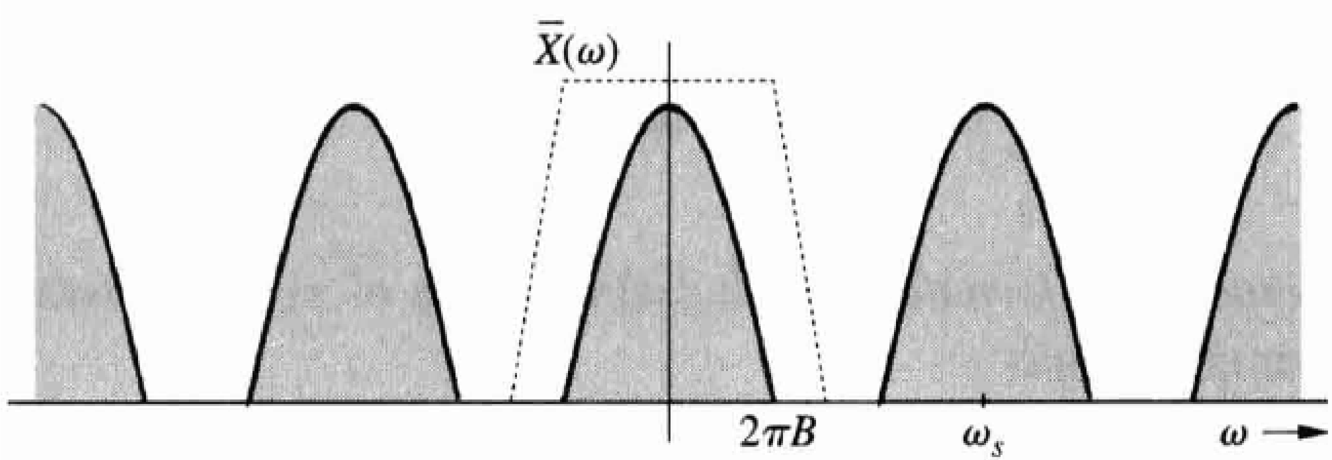
Practical signal reconstruction

- Ideal reconstruction system is therefore:





- In practise, we normally sample at higher frequency than Nyquist rate, so that we don't have to use an ideal lowpass filter (there isn't one anyway!).



Practical signal reconstruction cont.

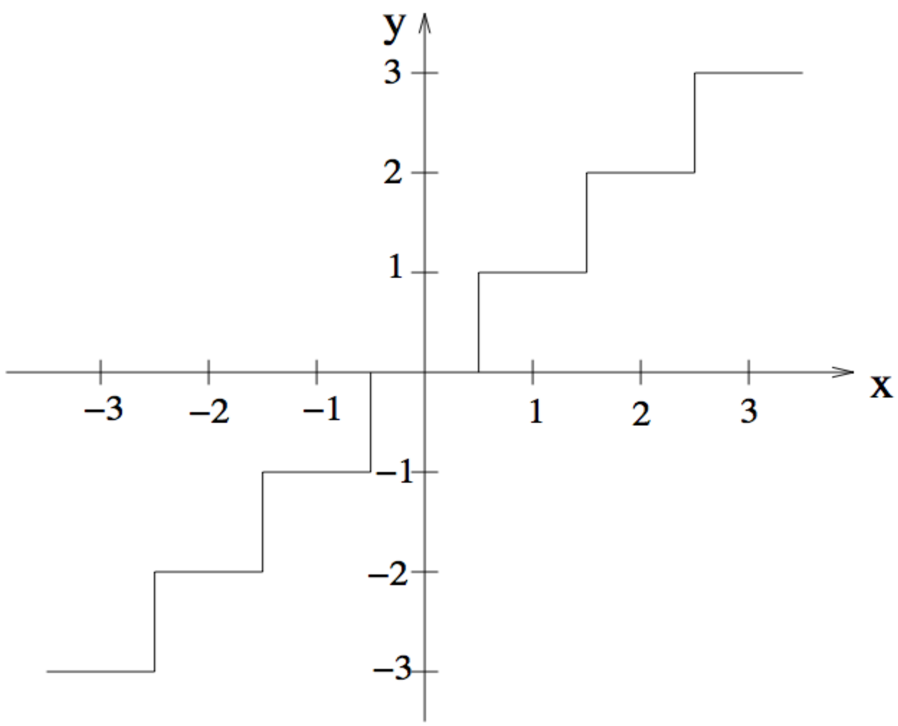
- The disappointing truth is that any filter that has a frequency response which is zero above a certain frequency is not realizable (Paley-Wiener criterion; for more information look at Lathi, Chapter 7, equation 7.61).
- The above implies that in practice it is impossible to perfectly reconstruct a bandlimited signal from its samples!
- However, a filter with gradual cutoff characteristics is easier to realize compared to an ideal filter.
- Furthermore, as the sampling rate increases, the recovered signal approaches the desired signal more closely.

Scalar quantisation

- The sampling theorem allows us to represent signals by means of samples.
- In order to store these samples we need to convert the real values $x[n]$ into a format which fits the memory model of a computer.
- The process that maps the real line to a countable set is called quantisation.
- Quantisation is irreversible. This means that quantisation introduces approximation errors.
- Usually the number of quantisation levels N is a power of 2, i.e., $N = 2^R$ so that each symbol can be expressed using a stream of bits.

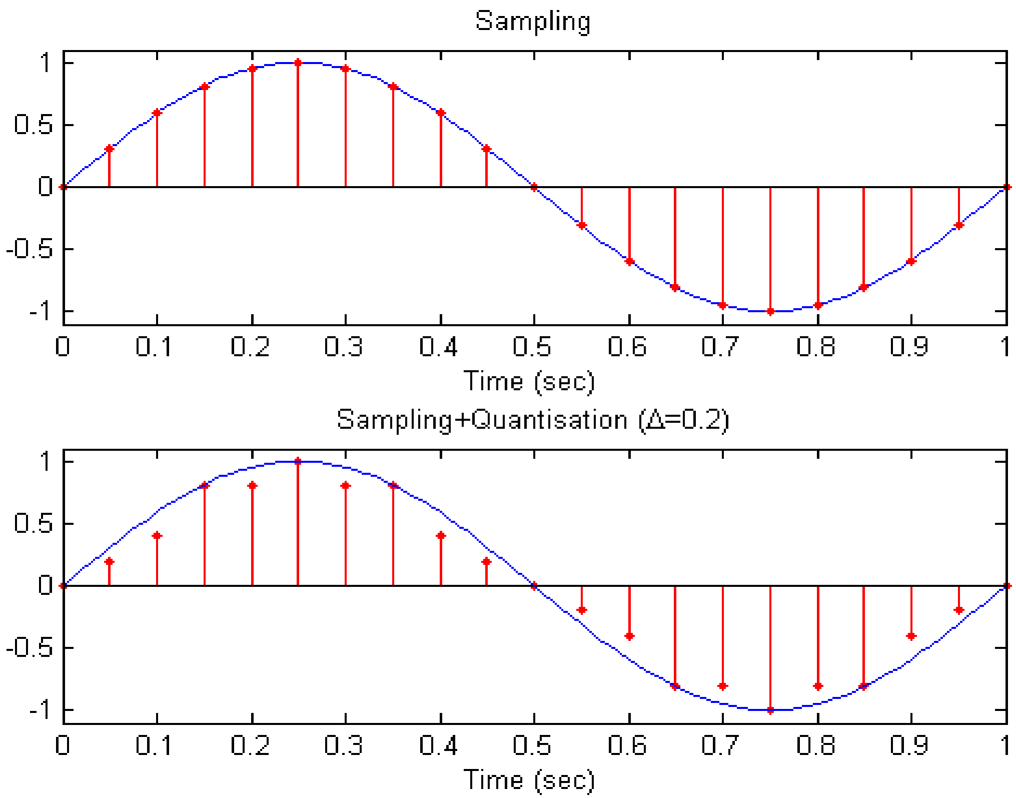
Scalar quantisation cont.

- Typically each sample is quantized independently (scalar quantisation).
- When the quantisation intervals have equal width we say that quantisation is ‘uniform’.



Scalar quantisation cont.

- Sampling followed by quantisation.

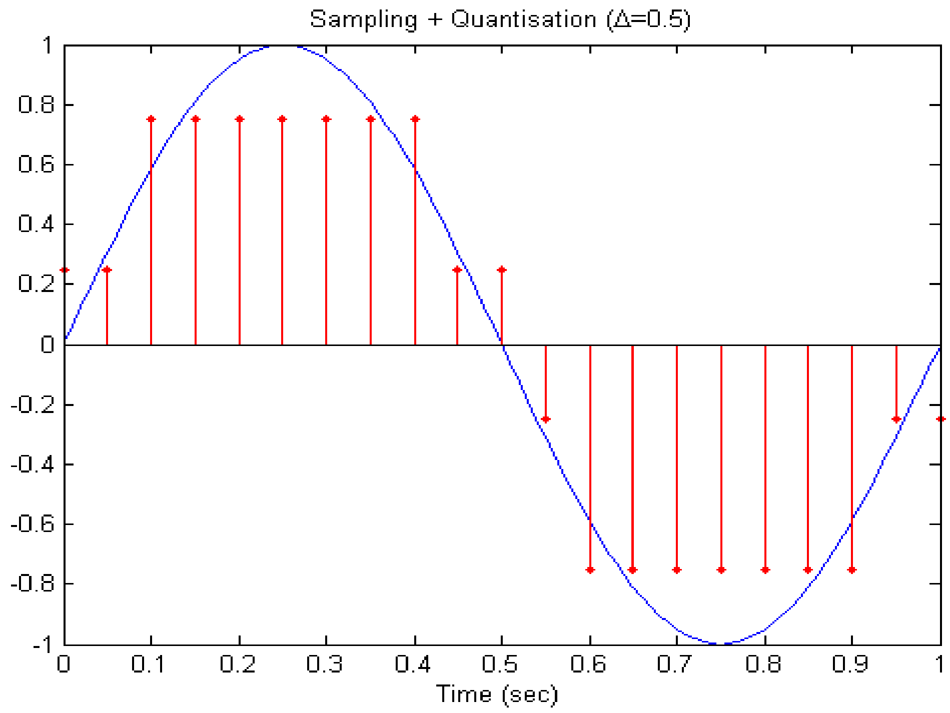


Quantisation examples

- A few numbers (in absence of compression)
  - Low-quality audio signal (e.g., voice)
    - Frequency from $300Hz$ to $3400Hz$, we assume $B = 4kHz$.
    - Sampling Frequency $f_s = 2B = 8kHz$.
    - Each sample is quantized with a 8-bit quantiser. The signal-to-noise-ratio due to quantisation error is $SNR \approx 48dB$.
    - This produces $\frac{64kbits}{s}$.
  - High Fidelity Audio Signals (e.g., CD)
    - Bandwidth $15Hz$, we assume $B = 22.5kHz$.
    - Sampling frequency $f_s = 2B = 44.1kHz$.
    - 16-bit quantiser ($SNR \approx 96dB$).
    - This produces $\approx \frac{706kbits}{s}$.

Quantisation error

- The figure depicts a 2-bit quantiser.

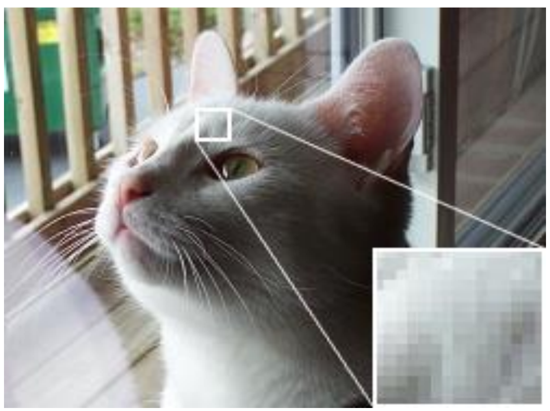


- Low-rate quantization introduces correlation between the quantisation error and the original signal.
- This correlation leads to audible artifacts in audio signals or artificial edges in images.

Imperial College London

Quantisation error cont.

- In the example below the quantized image has been allocated a number of intensity levels which is not large enough. Therefore, the quantized image does not look “real” to the human eye; instead it looks like a segmented image.



Original

Quantized

- Images have been taken from http://en.wikipedia.org/wiki/Dither.

For your own interest: Who discovered the sampling theorem?

- The sampling theorem is usually known as the Shannon Sampling Theorem due to Claude E. Shannon's paper "A mathematical theory of communication" in 1948.
- The minimum required sampling rate f_s (i.e. $2B$) is known as the Nyquist sampling rate or Nyquist frequency because of H. Nyquist's work on telegraph transmission in 1924 with K. Küpfmüller.
- The first formulation of the sampling theorem which was applied to communication is probably due to a Russian scientist by the name of V. A. Kotelnikov in 1933.
- However, mathematicians already knew about this in a different form and called this the interpolation formula. E. T. Whittaker published the paper "On the functions which are represented by the expansions of the interpolation theory" back in 1915.