analog $m(t)$ $\longrightarrow \boxed{\text{Sampling}} \longrightarrow m[n] \longrightarrow \boxed{\text{Quantization}} \longrightarrow$ digital m_n

Sirindhorn International Institute of Technology
Thammasat University
School of Information, Computer and Communication Technology

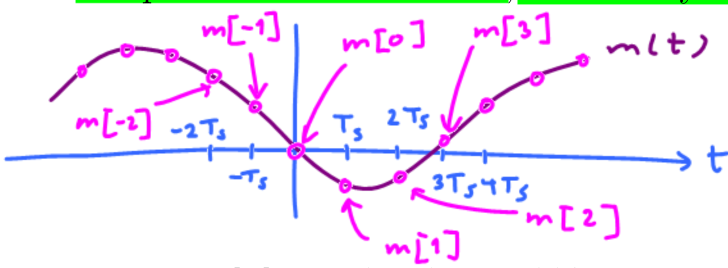
ECS332 2015/1 Part III.1 Dr.Prapun

6 Sampling, Reconstruction, and Pulse Modulation

6.1 Sampling

Definition 6.1. **Sampling** is the process of taking a (sufficient) number of discrete values of points on a waveform that will define the shape of wave form.

- The signal is sampled at a uniform rate, once every T_s seconds.



$$m[n] = m(nT_s) = m(t)|_{t=nT_s}.$$

- We refer to T_s as the **sampling period**, and to its reciprocal $f_s = 1/T_s$ as the **sampling rate**.

- The reverse process is called "reconstruction".

6.2. Sampling = loss of information? If not, how can we recover the original waveform back.

- The more samples you take, the more accurately you can define a waveform.

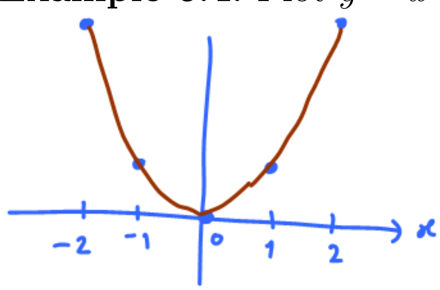
- Obviously, if the sampling rate is too low, you may experience distortion (aliasing).

[Version 1]

- The sampling theorem, to be discussed in the section, says that when the waveform is band-limited, if the sampling rate (f_s) is fast enough, we can reconstruct the waveform back and hence there is no loss of information.
  - This allows us to replace a continuous time signal by a discrete sequence of numbers.
  - Processing a continuous time signal is therefore equivalent to processing a discrete sequence of numbers.
  - In the field of communication, the transmission of a continuous time message reduces to the transmission of a sequence of numbers.

Example 6.3. Mathematical functions are frequently displayed as continuous curves, even though a finite number of discrete points was used to construct the graphs. If these points, or samples, have sufficiently close spacing, a smooth curve drawn through them allows us to interpolate intermediate values to any reasonable degree of accuracy. It can therefore be said that the continuous curve is adequately described by the sample points alone.

Example 6.4. Plot $y = x^2$.



step (1) is sampling

(1) "Evaluate" $y = x^2$ at many values of x

(2) "connect" the dots

step (2) is "reconstruction"

Example 6.5. Plot $g(t) = \sin(100\pi t)$. ← 50 Hz   sampling freq = 49 (f_s)
(See slides.)

[Version 2]

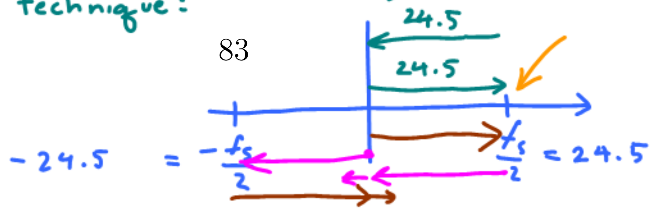
Theorem 6.6. Sampling Theorem: In order to (correctly and completely) represent an analog signal, the sampling frequency, f_s, must be at least twice the highest frequency component of the analog signal.

Example 6.7. In example 6.5, the frequency of the sine wave is 50 Hz. Therefore, we need the sampling frequency to be at least 100.

Example 6.8. Suppose the sampling frequency is 200 samples/sec. The analog signal should not have the frequency higher than 100 Hz.
(See slides)

Using the folding technique:

Back to Ex 6.5



$\sin x = \frac{e^{+jx} - e^{-jx}}{2j} = \frac{1}{2j} e^{+jx} - \frac{1}{2j} e^{-jx}$

$= -\frac{1}{2} j e^{jx} + \frac{1}{2} j e^{-jx}$

Definition 6.9.

(a) Given a sampling frequency, f_s, the **Nyquist frequency** is defined as $f_s/2$.

(b) Given the highest (positive-)frequency component $f_{\max}$ of an analog signal,

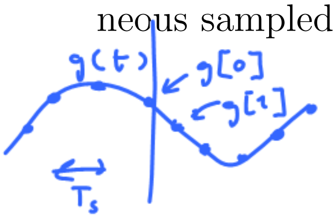
(i) the **Nyquist sampling rate** is $2f_{\max}$ (2B) and

(ii) the **Nyquist sampling interval** is $1/(2f_{\max})$.

6.10. Much more can be said about the result of performing the sampling process on a signal. Here we will use $g(t)$ to denote the signal under consideration. You may replace $g(t)$ below by $m(t)$ if you want to think of it as an analog message to be transmitted by a communication system. We use $g(t)$ here because the results provided here work in broader setting as well.

Definition 6.11. In **ideal sampling**, the (ideal instantaneous) sampled signal is represented by a train of impulses whose area equal the instantaneous sampled values of the signal

$g(t) \rightarrow \boxed{\text{sampling}} \rightarrow g[n]$



$$g_\delta(t) = \sum_{n=-\infty}^{\infty} g[n]\delta(t - nT_s).$$

6.12. The Fourier transform $G_\delta(f)$ of $g_\delta(t)$ can be found by first rewriting $g_\delta(t)$ as

$g[n] \equiv g(t)\big|_{t=nT_s}$

$$g_\delta(t) = \sum_{n=-\infty}^{\infty} g(nT_s)\delta(t - nT_s) = \sum_{n=-\infty}^{\infty} g(t)\delta(t - nT_s)$$

$g(t)\delta(t-c) = g(c)\delta(t-c)$

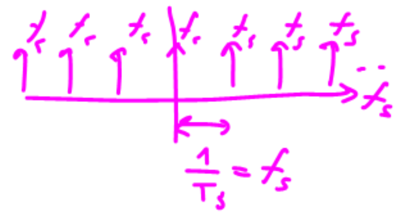
$$= g(t) \times \sum_{n=-\infty}^{\infty} \delta(t - nT_s)$$

Multiplication in the time domain corresponds to convolution in the frequency domain. Therefore,

$$G_\delta(f) = \mathcal{F}\{g_\delta(t)\} = G(f) * \mathcal{F}\left\{\sum_{n=-\infty}^{\infty} \delta(t - nT_s)\right\}.$$

For the last term, the Fourier transform can be found by applying what we found in Example 4.13[24]:

$$\sum_{n=-\infty}^{\infty} \delta\left(t - nT_s\right) \underset{\mathcal{F}^{-1}}{\overset{\mathcal{F}}{\rightleftharpoons}} f_s \sum_{k=-\infty}^{\infty} \delta\left(f - kf_s\right).$$



This gives

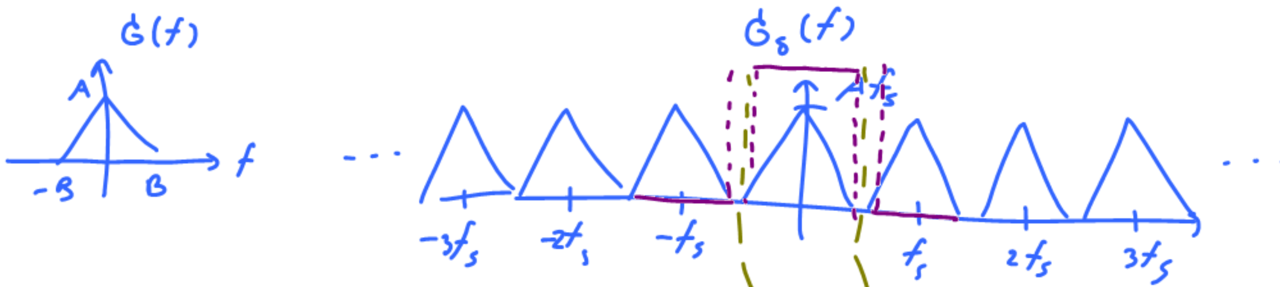
$$G_\delta\left(f\right) = G\left(f\right) * f_s \sum_{k=-\infty}^{\infty} \delta\left(f - kf_s\right) = f_s \sum_{k=-\infty}^{\infty} G\left(f\right) * \delta\left(f - kf_s\right).$$

Hence, we conclude that

$$\boxed{g_\delta\left(t\right) = \sum_{n=-\infty}^{\infty} g\left[n\right]\delta\left(t - nT_s\right) \underset{\mathcal{F}^{-1}}{\overset{\mathcal{F}}{\rightleftharpoons}} G_\delta\left(f\right) = f_s \sum_{k=-\infty}^{\infty} G\left(f - kf_s\right).} \quad (57)$$

6.13. As usual, we will assume that the signal $g(t)$ is band-limited to B Hz (($G(f) = 0$ for $|f| > B$)). In which case, the Fourier transform of the sampled signal is given by

For reconstruction,



6.14. Remarks:

(a) $G_\delta(f)$ is "periodic" (in the frequency domain) with "period" f_s.

- So, it is sufficient to look at $G_\delta\left(f\right)$ between $\pm\frac{f_s}{2}$

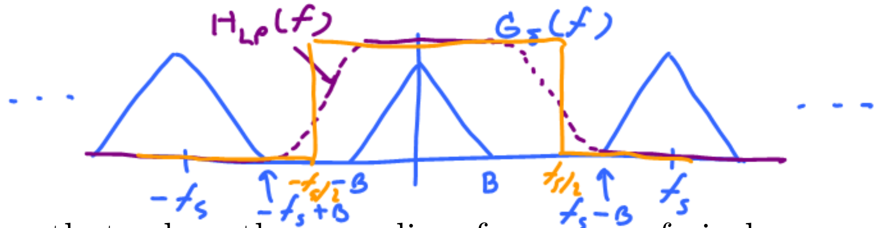
(b) The MATLAB script `plotspect` that we have been using to visualize magnitude spectrum also relies on sampled signal. Its frequency domain plot is between $\pm\frac{f_s}{2}$.

(c) Although this sampling technique is "ideal" because it involves the use of the δ-function. We can extract many useful conclusions.

(d) One can also study the discrete-time Fourier transform (DTFT) to look at the frequency representation of the sampled signal.

[24] We also considered an easy-to-remember pair and discuss how to extend it to the general case in 4.14.

6.2 Reconstruction



6.15. From (57), we see that when the sampling frequency f_s is large enough, the replicas of $G(f)$ will not overlap in the frequency domain. In such case, the original $G(f)$ is still intact and we can use a low-pass filter with gain T_s to recover $g(t)$ back from $g_\delta(t)$.

6.16. To prevent aliasing (the corruption of the original signal because its replicas overlaps in the frequency domain), we need

$f_s - B > B$

$f_s > 2B$ Nyquist sampling rate

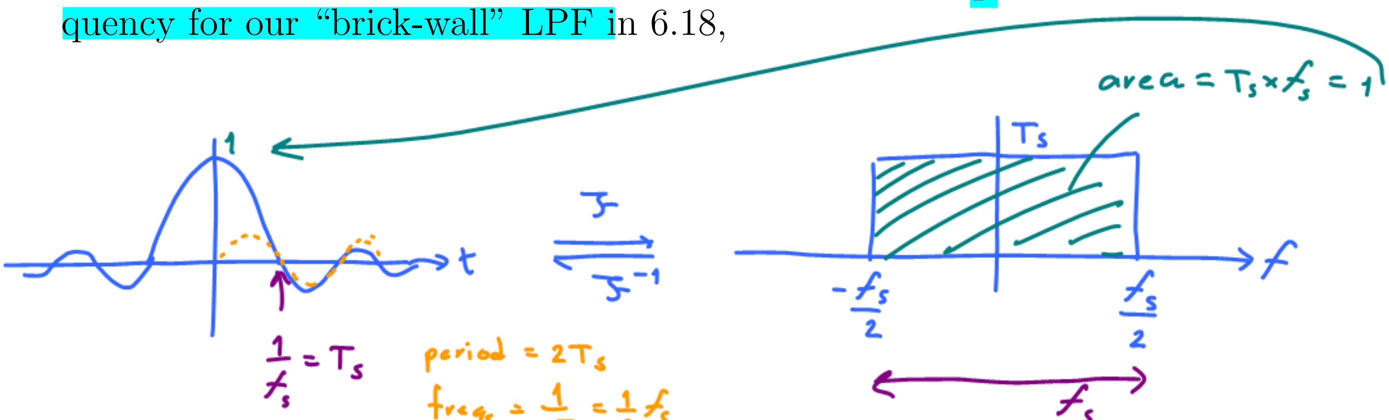
Theorem 6.17. A low-pass signal g whose spectrum is band-limited to B Hz ($G(f) = 0$ for $|f| > B$) can be reconstructed (interpolated) exactly (without any error) from its sample taken uniformly at a rate (sampling frequency/rate) $f_s > 2B$ Hz (samples per second).[5, p 302]

6.18. Ideal Reconstruction: Continue from 6.15. Assuming that $f_s > 2B$, the low-pass filter that we should use to extract $g(t)$ from $G_\delta(t)$ should be

$$H_{LP}(f) = \begin{cases} \overset{T_s}{1/f_s}, & |f| \le B, \\ \text{any}, & B < |f| < f_s - B, \\ 0, & |f| \ge f_s - B, \end{cases}$$

In particular, for "brick-wall" LPF, the cutoff frequency f_{cutoff} should be between B and $f_s - B$.

6.19. Reconstruction Equation: Suppose we use $\frac{f_s}{2}$ as the cutoff frequency for our "brick-wall" LPF in 6.18,



The impulse response of the LPF is $h_{LP}(t) = \operatorname{sinc}\left(2\pi\left(\frac{f_s}{2}\right)t\right) = \operatorname{sinc}(\pi f_s t)$.

$\operatorname{sinc}(2\pi(\frac{1}{2}f_s)t)$

The output of the LPF is

$$g_r(t) = g_\delta(t) * h_{LP}(t) = \left(\sum_{n=-\infty}^{\infty} g[n]\delta(t - nT_s)\right) * h_{LP}(t)$$
$$= \sum_{n=-\infty}^{\infty} g[n]h_{LP}(t - nT_s) = \sum_{n=-\infty}^{\infty} g[n]\operatorname{sinc}(\pi f_s(t - nT_s)).$$

When $f_s > 2B$, this output will be exactly the same as $g(t)$:

$$\boxed{g(t) = \sum_{n=-\infty}^{\infty} g[n]\operatorname{sinc}(\pi f_s(t - nT_s))} \quad (58)$$

Reconstruction equation.

- This formula allows perfect reconstruction the original continuous-time function from the samples.
- At the sampling instants $t = nT_s$, all sinc functions are zero at these times save one, and that one yields $g(nT_s)$ which is the correct values.
- Note that at time t between the sampling instants, $g(t)$ is interpolated by summing the contributions from all the sinc functions.
- The LPF is often called an interpolation filter, and its impulse response is called the interpolation function.

Example 6.20.

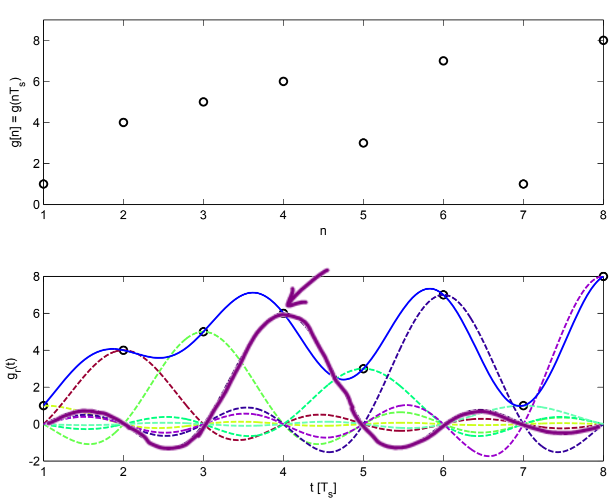


Figure 33: Application of the reconstruction equation

Theorem 6.21. Sampling theorem for uniform periodic sampling: If a signal $g(t)$ contains no frequency components for $|f| \geq B$, it is completely described by instantaneous sample values uniformly spaced in time with sampling period $T_s \leq \frac{1}{2B}$. In which case, $g(t)$ can be exactly reconstructed from its samples $(\ldots, g[-2], g[-1], g[0], g[1], g[2], \ldots)$ by the reconstruction equation (58).

Example 6.22. We now return to the sampling of the cosine function (sinusoid).

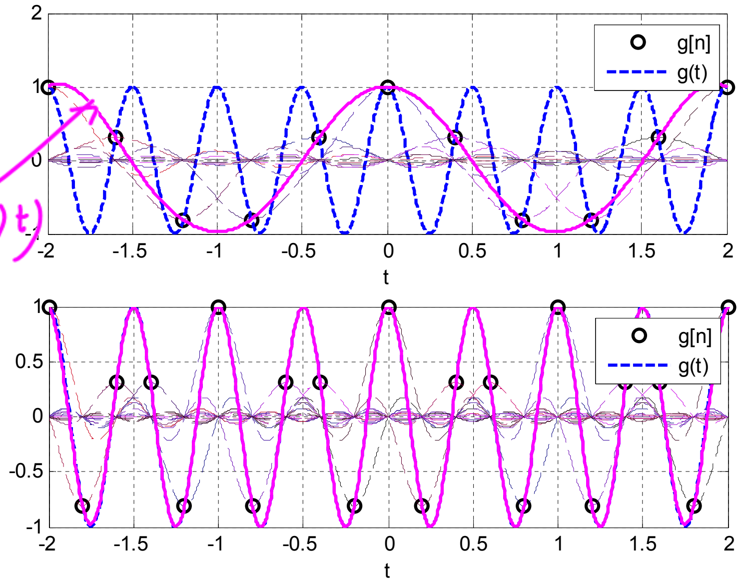
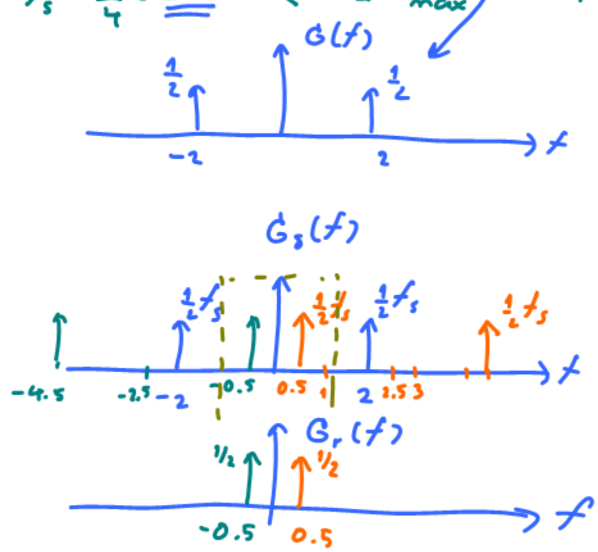


Figure 34: Reconstruction of the signal $g(t) = \cos(2\pi 2 t)$ by its samples $g[n]$. The upper plot uses $T_s = 0.4$. The lower plot uses $T_s = 0.2$.

$\cos(2\pi(0.5)t)$

$f_s = \frac{10}{4} = 2.5 < 2 \times f_{max} = 2 \times 2 = 4$



6.23. Remarks:

- Need a lot of $g[n]$ for the reconstruction.
- Practical signals are time-limited.
  - Filter the message as much as possible before sampling.

6.24. The possibility of $f_s = 2B$:

- If the spectrum $G(f)$ has no impulse (or its derivatives) at the highest frequency B, then the overlap is still zero as long as the sampling rate is greater than or equal to the Nyquist rate, that is, $f_s \geq 2B$.
- If $G(f)$ contains an impulse at the highest frequency $\pm B$, then $f_s = 2B$ would cause overlap. In such case, the sampling rate f_s must be greater than $2B$ Hz.

Theorem 6.21. Sampling theorem for uniform periodic sampling: If a signal $g(t)$ contains no frequency components for $|f| \geq B$, it is completely described by instantaneous sample values uniformly spaced in time with sampling period $T_s \leq \frac{1}{2B}$. In which case, $g(t)$ can be exactly reconstructed from its samples $(\ldots, g[-2], g[-1], g[0], g[1], g[2], \ldots)$ by the reconstruction equation (58).

Example 6.22. We now return to the sampling of the cosine function (sinusoid).

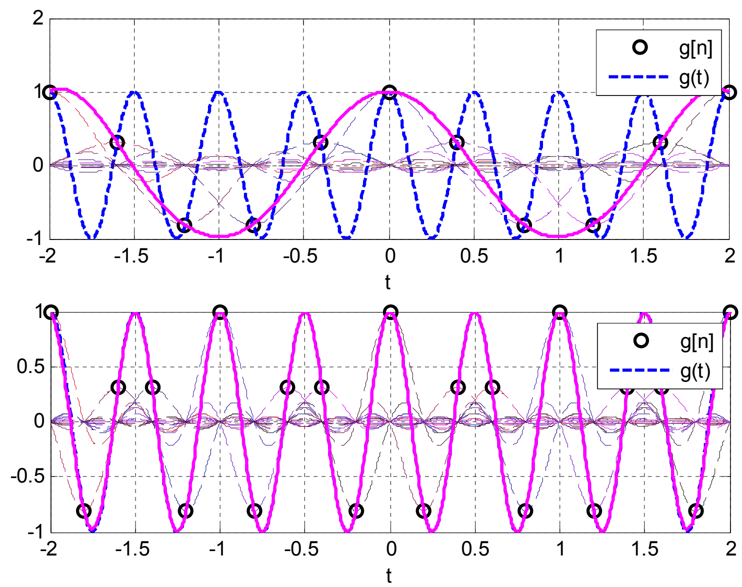


Figure 34: Reconstruction of the signal $g(t) = \cos(2\pi(2)t)$ by its samples $g[n]$. The upper plot uses $T_s = 0.4$. The lower plot uses $T_s = 0.2$.

6.23. Remarks:

- Need a lot of $g[n]$ for the reconstruction.
- Practical signals are time-limited.
  - Filter the message as much as possible before sampling.

6.24. The possibility of $f_s = 2B$:

- If the spectrum $G(f)$ has no impulse (or its derivatives) at the highest frequency B, then the overlap is still zero as long as the sampling rate is greater than or equal to the Nyquist rate, that is, $f_s \geq 2B$.
- If $G(f)$ contains an impulse at the highest frequency $\pm B$, then $f_s = 2B$ would cause overlap. In such case, the sampling rate f_s must be greater than $2B$ Hz.

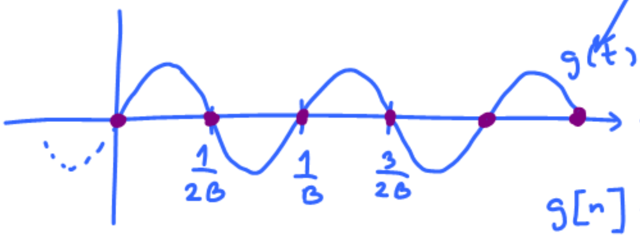
freq. of the sine wave

period = 1/B

Example 6.25. Consider a sinusoid $g(t) = \sin(2\pi(B)t)$. This signal is bandlimited to B Hz, but all its samples are zero when uniformly taken at a rate $f_s = 2B$, and $g(t)$ cannot be recovered from its (Nyquist) samples. Thus, for sinusoids, the condition of $f_s > 2B$ must be satisfied.

$f_s = 2B$

$T_s = \frac{1}{2B}$



for all

$g[n] \equiv 0 \ \forall n \Rightarrow g_\delta(t) = \sum_n \overbrace{g[n]}^{0}\delta(t-nT_s) \equiv 0$

$\Downarrow$

$G_\delta(f) \equiv 0 \ \forall f$

$\Downarrow$

$G_r(f) \equiv 0$

$\Downarrow$

$g_r(t) \equiv 0$

Let's check with our formula (57) for $G_\delta(f)$. First, recall that

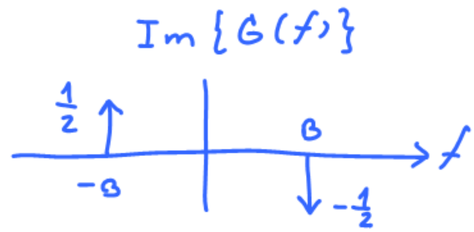
$$\sin x = \frac{e^{jx} - e^{-jx}}{2j} = \frac{1}{2j}e^{jx} - \frac{1}{2j}e^{-jx}.$$

Therefore,

$$g(t) = \sin(2\pi(B)t) = \frac{1}{2j}e^{j2\pi(B)t} - \frac{1}{2j}e^{-j2\pi(B)t} = -\frac{1}{2}j\,e^{j2\pi(B)t} + \frac{1}{2}j\,e^{j2\pi(-B)t}$$

and

$G(f) = -\frac{1}{2}j\,\delta(f-B) + \frac{1}{2}j\,\delta(f-(-B))$



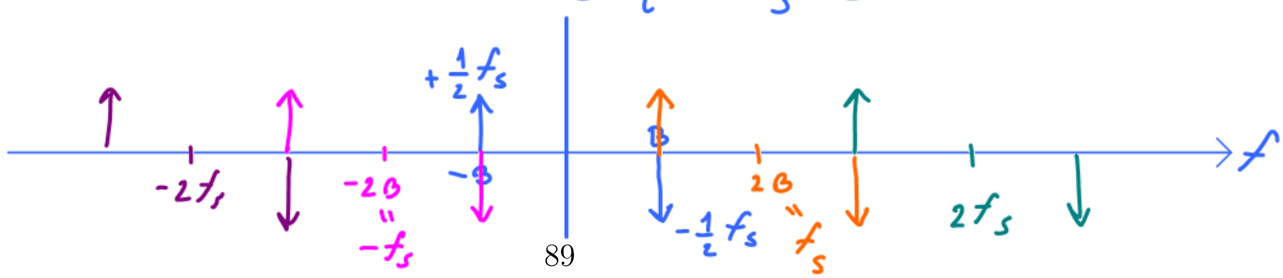
Note that $G(f)$ is pure imaginary. So, it is more suitable to look at the plot of its imaginary part. (We do not look at its magnitude plot because the information about the sign is lost. We also do not consider the real part because we know that it is 0.)

$G_\delta(f) = \sum_k f_s G(f-kf_s)$

$\mathrm{Re}\{G_\delta(f)\} = \sum_k f_s \overbrace{\mathrm{Re}\{G(f-kf_s)\}}^{0} = 0$

$\mathrm{Im}\{G_\delta(f)\} = \sum_k f_s \mathrm{Im}\{G(f-kf_s)\}$

$\mathrm{Im}\{G_\delta(f)\} \equiv 0$



$G_\delta(f) = \mathrm{Re}\{G_\delta(f)\} + j\,\mathrm{Im}\{G_\delta(f)\} \equiv 0 + j0 \equiv 0$ ✓

6.26. A maximum of $2B$ independent pieces (samples/symbols) of information per second can be transmitted, errorfree, over a noiseless channel of bandwidth B Hz [4, p 260].

- Start with $2B$ pieces of information per second. Denote the sequence of such information by m_n.
- Construct a signal $m(t)$ whose (Nyquist) sample values $m[n] = m\left(n\frac{1}{2B}\right)$ agrees with m_n by the reconstruction equation (58).

6.27. A bandpass signal whose spectrum exists over a frequency band $f_c - \frac{B}{2} < |f| < f_c + \frac{B}{2}$ has a bandwidth B Hz. Such a signal is also uniquely determined by samples taken at above the Nyquist frequency $2B$. The sampling theorem is generally more complex in such case. It uses two interlaced sampling trains, each at a rate of $f_s > B$ samples per second (known as second-order sampling). [5, p 304]

6.3 Analog Pulse Modulation

In Section 6.1 we saw that continuous bandlimited signals can be represented by a sequence of discrete samples. Moreover, in Section 6.2, we saw that the continuous signal can be reconstructed if the sampling rate is sufficiently high.

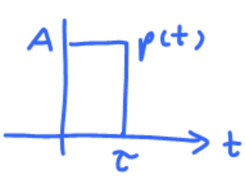
6.28. Because the sequence $m[n]$ completely contains the information about $m(t)$, in this section, instead of trying to send $m(t)$, we consider transmitting the message in the form of pulse modulation.

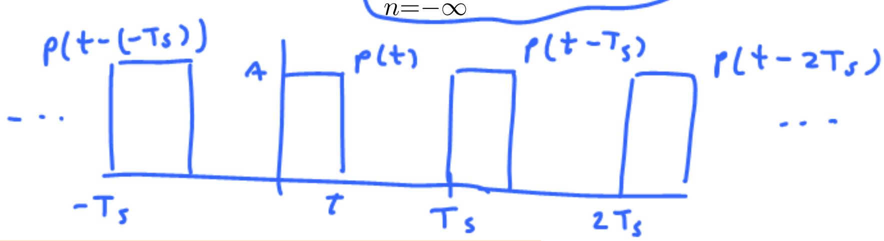
Definition 6.29. In analog pulse modulation, some attribute of a pulse varies continuously in one-to-one correspondence with a sample value.

- Example of a pulse:

- Three attributes can be readily varied: amplitude, width, and position.
- These lead to pulse-amplitude modulation (PAM), pulse-width modulation (PWM), and pulse-position modulation (PPM) as illustrated in Figure 35.

Definition 6.30. Unmodulated pulse train: $\sum_{n=-\infty}^{\infty} p(t - nT_s)$



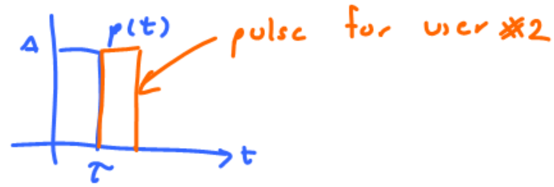


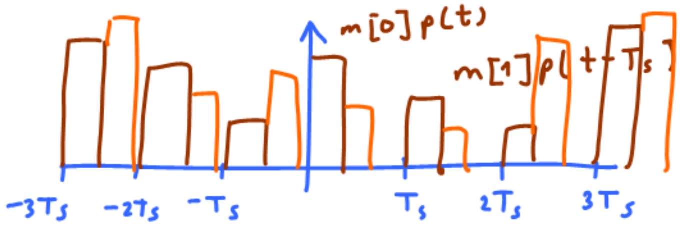
Definition 6.31. In ***Pulse-Amplitude Modulation*** (PAM), the sample values modulate the amplitude of a pulse train:

$$x_{\text{PAM}}(t) = \sum_{n=-\infty}^{\infty} m[n]\, p(t - nT_s)$$

Pulse-modulated signal

Example 6.32.





6.33. One advantage of using pulse modulation is that it permits the simultaneous transmission of several signals on a **time-sharing** basis.

- When a pulse-modulated signal occupies only a part of the channel time, we can transmit several pulse-modulated signals on the same channel by interweaving them.
- One User: TDM (time division multiplexing).
  - Transmit/multiplex multiple streams of information simultaneously.
- Multiple Users: TDMA (time division multiple access).

6.34. Frequency-Domain Analysis of PAM:

$$x_{\text{PAM}}(t) = \sum_{n=-\infty}^{\infty} m[n]\, p(t - nT_s) = \sum_{n=-\infty}^{\infty} m[n]\, p(t) * \delta(t - nT_s)$$
$$= p(t) * \left(\sum_{n=-\infty}^{\infty} m[n]\, \delta(t - nT_s) \right) = p(t) * m_\delta(t)$$

Therefore,

$$X_{\text{PAM}}(f) = P(f)\, M_\delta(f)\,.$$

6.35. Figure 35 compares different types of analog pulse modulation.

Definition 6.36. ***Pulse-Width Modulation*** (PWM): A PWM waveform consists of a sequence of pulses with the width of the nth pulse is proportional to the value of $m[n]$.

Ch 6 $\rightarrow$ Sampling $\leftarrow$ 6.1

Reconstruction $\leftarrow$ 6.2

$m(t)$ cont.-time signal (analog)

$\downarrow$ sampling $\uparrow$ reconst.

$m[n]$ discrete-time signal

(analog) pulse modulation $\leftarrow$ 6.3

↳ PAM

↳ Amplitude

$m[n]$ discrete-time signal

Ch 7 $\rightarrow$ pulse shaping for PAM

$$e^{-j2\pi \underline{\underline{f}}\, nT_s} = e^{-j\,2\pi\,(nT_s)\,f} = \cos\left(2\pi\,(nT_s)\,f\right) - j\sin\left(2\pi(nT_s)f\right)$$

$$e^{jx} = \cos(x) + j\sin(x)$$

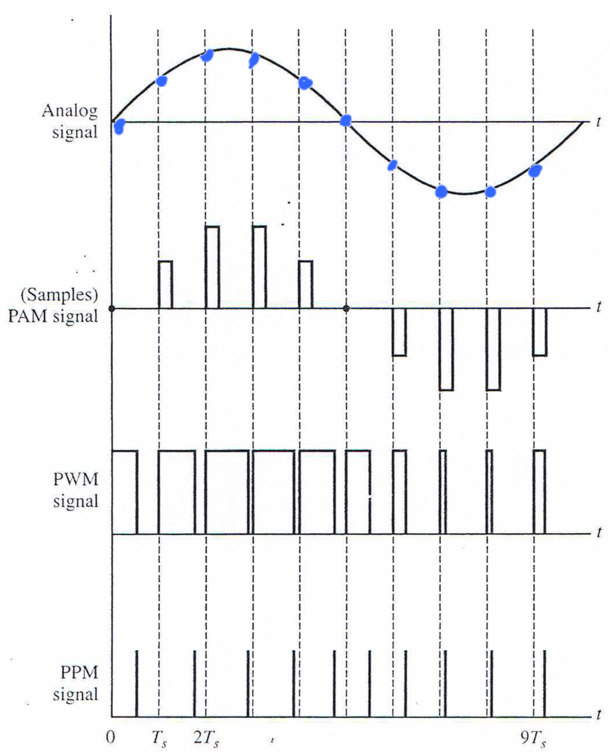


Figure 35: Illustration of PAM, PWM, and PPM.

- Seldom used in modern communications systems.
- Used extensively for DC motor control in which motor speed is proportional to the width of the pulses . Since the pulses have equal amplitude, the energy in a given pulse is proportional to the pulse width.

Definition 6.37. ***Pulse-Position Modulation*** (PPM): A PPM signal consists of a sequence of pulses in which the pulse displacement from a specified time reference is proportional to the sample values of the information-bearing signal.

- Have a number of applications in the area of ultra-wideband communications.

6.38. Pulse-modulation scheme are really baseband coding schemes, and they yield baseband signal.