# Lecture 1: Historical Notes and Overview

Dr. Geoffrey Li

Department of Electrical & Electronic Engineering
Imperial College London

# Course Information

**Contact:**

- Instructor: Dr. Geoffrey Li
- Office: EE 804
- Email: geoffrey.li@imperial.ac.uk
- Phone: +44 7709 517951

**Handouts:**

- Slides
- Problem sheets

**Grading:**

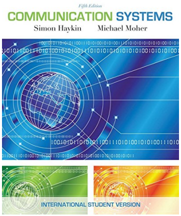
- Labs (30%)
- Mid-term exam (10%)
- Final 2-hour exam (60%)

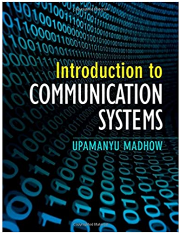
# Contents

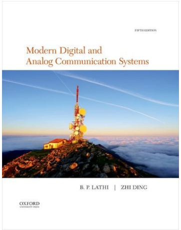
- Communication systems performance in the presence of noise
- About statistical aspects and impact of noise
- Main mathematical tools: Fourier transform, probability, and stochastic processes

# References

- [Haykin] S. Haykin and M. Moher, *Communication Systems*, 5th ed., Wiley, 2009.
- [Madhow] U. Madhow, *Introduction to Communication Systems*, Cambridge University Press, 2015.
- [Lathi] B. Lathi and Z. Ding, *Modern Digital and Analog Communication Systems*, 5th ed., Oxford University Press, 2018.



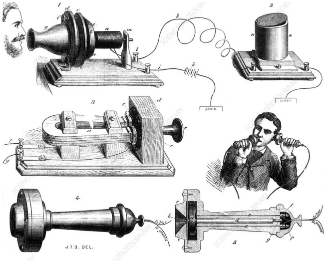



# Milestones in Communications: I

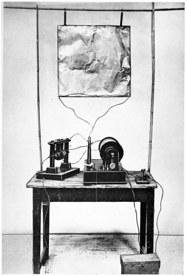
19th-century demonstration of the semaphore


Souvenir card for the Dover-Calais cable, 1854

- 1792, Chappe invented optical telegraph "semaphore" in France (mechanical method)
- 1837, First commercial telegraph service, Paddington station and West Drayton by William Cooke and Charles Wheatstone (telegraph/digital, wire)
- 1851, England connected to Europe by a cable between Dover, UK and Calais, France (submarine cable)
- 1864, Maxwell formulated the electromagnetic (EM) theory (predicted the existence of EM waves)



Bell's first telephone



Marconi's first radio transmitter

- 1875, Bell invented the telephone (transmit analog signal/speech, wired)
- 1887, Hertz demonstrated physical evidence of EM waves (made radio communication possible)
- 1890's-1900's, Marconi & Popov, long-distance radio telegraph (first wireless communication, telegraph)
- 1906, First radio broadcast (opera signal, wireless)

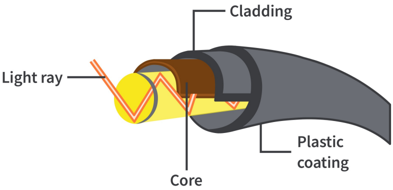
# Milestones in Communications: III

Edwin Howard Armstrong (1890-1954) was an American electrical engineer who invented frequency modulation (FM) radio and the superheterodyne receiver system.


John Logie Baird (1888-1946) was a Scottish electrical engineer who invented live television (TV) and color television systems.

- 1918, Armstrong invented superheterodyne radio receiver (and FM in 1933)
- 1920, First commercial radio station (500 stations by 1923)
- 1925, Baird demonstrated transmission of moving images (TV) in London
- 1928, First TV station by General Electric (GE) factory in Schenectady, New York
- 1928, Nyquist discovered sampling theorem at Bell Labs (will introduce in Lect. 10)
- 1948, Shannon established information theory at Bell Labs

# Milestones in Communications: IV



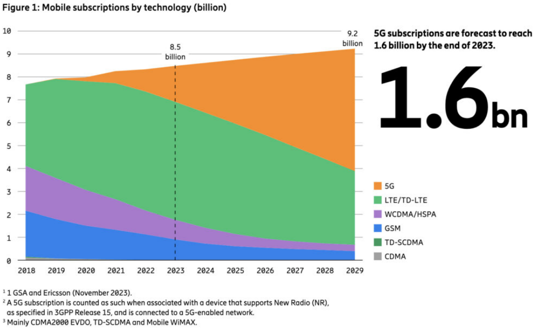

Parts of optical fiber: core, cladding, and plastic coating

Martin Cooper reenacting the first private handheld mobile-phone call, 2007

- 1966, Kuen Kao pioneered fiber-optical communications (Nobel Prize Winner)
- 1971, First public wireless network ALOHANET established at University of Hawaii
- 1973, First mobile phone by Martin Cooper, Motorola
- 1978, First cellular mobile phone system (1G) developed by AT&T Bell Labs

# Growth of Mobile Communications

- 1G: analog communications
  - AMPS
- 2G: digital communications
  - GSM
  - IS-95
- 3G: CDMA networks
  - WCDMA
  - CDMA2000
  - TD-SCDMA
- 4G: multi-antenna, multi-carrier
  - WIMAX
  - LTE
  - OFDMA
- 5G: high speed, low latency, massive connectivity
  - eMBB
  - mMTC
  - URLLC
- 6G: intelligent, conscious, secure, sustainable



## Transfer analog signal

- AM (525 – 1606.5 kHz)
- FM (87.5 – 108.0 MHz)
- Analog TV
  - Video 45 – 66.75/179.75 – 214.75 MHz
  - Audio 41.5 – 63.25/176.25 – 211.25 MHz
- 1G
  - AMPS (Advanced Mobile Phone System, 824 – 894 MHz) in America and Australia
  - TACS (Total Access Communication System, 890 – 950 MHz) in UK


Radio City Tower, Liverpool

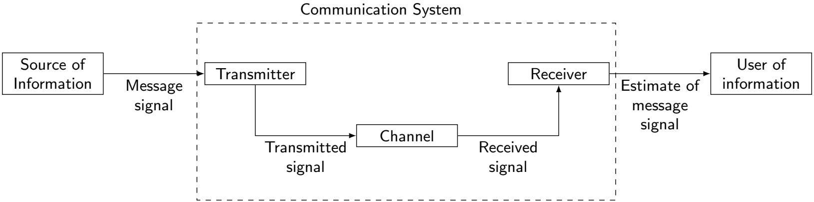
## Transfer digital signal

- Transfer of information in bits
- Digital TV, CDs, DVD
- Broadband, 2G – 5G, . . .

The data side of a DVD

# Communication Systems: Four Basic Elements

- Information source: voice, music, picture, video, ...
- Transmitter: converting information in the source into a form suitable for transmission over the channel
- Channel: the physical medium, introducing distortion, noise, and interference
- Receiver: reconstructing a recognizable form of the source signal




Block diagram of a communication system

# Noise

- Unwanted signals in a communication system
- External noise: interference from nearby channels, human-made noise, natural noise...
- Internal noise: thermal noise, random motion of electrons
- Noise limiting the performance of communication systems
- Signal-to-noise ratio (SNR) is a widely used metric (will discuss in Lect. 7)

$$\text{SNR} = \frac{\text{average signal power}}{\text{average noise power}}$$

# Transmitter

- converting the source signal into a form suitable for transmission over the channel
- including modulation and up-conversion

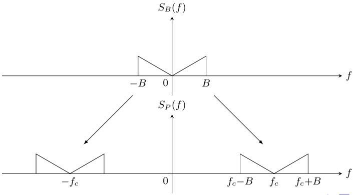
Modulation: changing some parameters of a carrier based on the source signal

- A carrier wave:

$$x(t) = A\cos(2\pi f_c t + \theta)$$

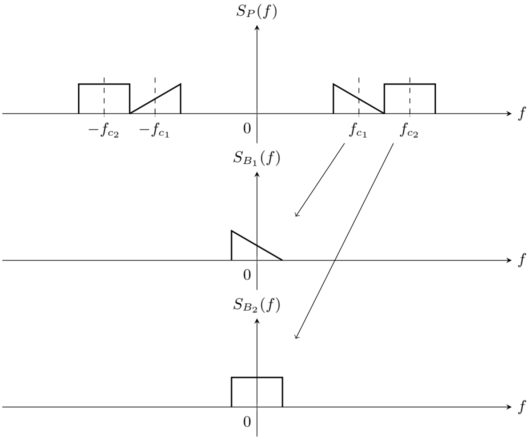
  - $f_c$: carrier frequency
  - $A$: amplitude
  - $\theta$: phase
- Analog: AM, FM, PM (*M: modulation*)
- Digital: ASK, FSK, PSK (*SK: shift keying*)

Up-conversion: converting modulated signal to final radio frequency (RF)



# Receiver

- reconstructing original message by down-conversion and demodulation
- no exact recovery due to noise and distortion
- degradation depending on the type of modulation and channel

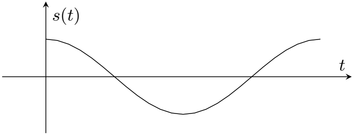


# Digital, Discrete, Analog Signals

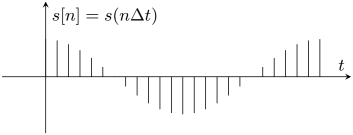
Analog signals: continuous in both time and amplitude (speech, image, video)

Discrete signals: discrete in time but continuous in amplitude (sampled version of continuous signal)

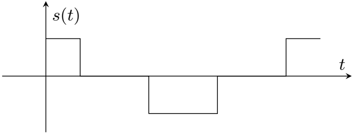
Digital signals: discrete in time and amplitude



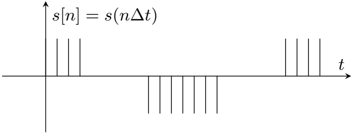

(a) Analog signal




(b) Discrete signal




(c) Digital signal




(d) Digital signal (seldom meet)

# Digital vs Analog Communications

- Transmitted signals: current, voltages, EM waves $\rightarrow$ always analog
- Digital vs analog: depending on how parameters of these waveforms are formed
- Digital systems: source signal $\rightarrow$ source messages, digital signal (such as binary, etc.), analog signal for channel transmission
- Analog systems: conceptually simple, directly converting analog signal for channel transmission
- Digital communication: more efficient and reliable; more sophisticated types
- Digital design: universal and modular, any signal can be converted to digital format
- Performance metric of digital communications: Bit Error Rate (BER) (will discuss more in Lect. 12)

# Note

# Lecture 2: Fourier Transform and Probability

Dr. Geoffrey Li

Department of Electrical & Electronic Engineering
Imperial College London

# Outline

- Fourier transform
  - Definition
  - Properties
- Probability
  - Definition
  - cdf and pdf
  - Mean and variance
- References
  - [Haykin] Chapter 5
  - [Lathi] Chapter 7 – 8

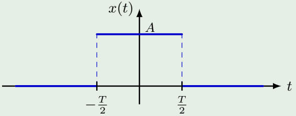
# Fourier Transform (FT) and Inverse Fourier Transform (IFT)

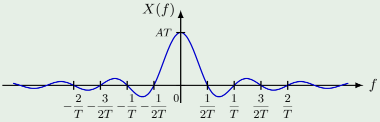
| | In $f$-Domain | In $\omega$-Domain |
|---|---|---|
| FT | $X(f) = \int_{-\infty}^{\infty} x(t)e^{-j2\pi ft}dt$ | $X(\omega) = \int_{-\infty}^{\infty} x(t)e^{-j\omega t}dt$ |
| IFT | $x(t) = \int_{-\infty}^{\infty} X(f)e^{j2\pi ft}df$ | $x(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} X(\omega)e^{j\omega t}d\omega$ |

## FT of rectangle function

$$\mathrm{rect}(x) \triangleq \begin{cases} 1, & -\frac{1}{2} < x < \frac{1}{2} \\ 0, & |x| \geq \frac{1}{2} \end{cases}, \quad \mathrm{sinc}(x) \triangleq \sin(\pi x)/\pi x$$

$$x(t) = A \cdot \mathrm{rect}\left(\frac{t}{T}\right) \Longleftrightarrow X(f) = AT \cdot \mathrm{sinc}(fT)$$



# Properties of the Fourier Transform

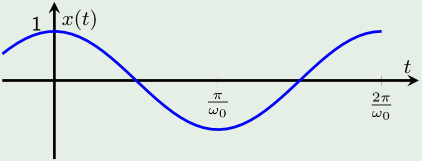
| Function | Fourier Transform |
|---|---|
| $a_1 x_1(t) + a_2 x_2(t)$ | $a_1 X_1(f) + a_2 X_2(f)$ |
| $x(t - t_0)$ | $X(f) e^{-j2\pi f t_0}$ |
| $x(at) \quad (a > 0)$ | $\frac{1}{a} X\left(\frac{f}{a}\right)$ |
| $x(-t)$ | $X(-f) = X^*(f)$ |
| $X(t)$ | $x(-f)$ |
| $x(t) e^{j\omega_0 t}$ | $X(f - f_0)$ |
| $x(t)\cos(\omega_0 t)$ | $\frac{1}{2} X(f - f_0) + \frac{1}{2} X(f + f_0)$ |
| $x(t)\sin(\omega_0 t)$ | $\frac{1}{2j} X(f - f_0) - \frac{1}{2j} X(f + f_0)$ |
| $\frac{d^n x(t)}{dt^n}$ | $(j2\pi f)^n X(f)$ |
| $\int_{-\infty}^{t} x(\tau) d\tau$ | $(j2\pi f)^{-1} X(f) + \frac{1}{2} X(0)\delta(f)$ |
| $\int_{-\infty}^{\infty} x_1(t-\tau) x_2(\tau) d\tau = \int_{-\infty}^{\infty} x_1(\tau) x_2(t-\tau) d\tau$ | $X_1(f) X_2(f)$ |
| $x_1(t) x_2(t)$ | $\int_{-\infty}^{\infty} X_1(f-\xi) X_2(\xi) d\xi = \int_{-\infty}^{\infty} X_1(\xi) x_2(f-\xi) d\xi$ |

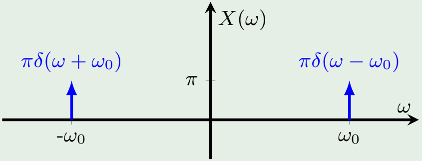
# Fourier Transform: Example 2

## Cosine wave

Find the Fourier transform of $x(t) = \cos\omega_0 t$:

$$x(t) = \frac{1}{2}(e^{-j\omega_0 t} + e^{j\omega_0 t}) \Longleftrightarrow X(\omega) = \pi\delta(\omega - \omega_0) + \pi\delta(\omega + \omega_0)$$



# Real and Symmetry

- $x(t)$ is real $\Leftrightarrow X(-\omega) = X^*(\omega)$
- $X(\omega)$ is real $\Leftrightarrow x(-t) = x^*(t)$
- $X(\omega)$ is real and $X(\omega) = X(-\omega) \Leftrightarrow x(t) = x(-t)$
- If $\mathcal{F}\{x(t)\} = X(\omega)$, then $\begin{cases} \mathcal{F}\{x(-t)\} = X(-\omega) \\ \mathcal{F}\{x^*(t)\} = X^*(-\omega) \end{cases}$

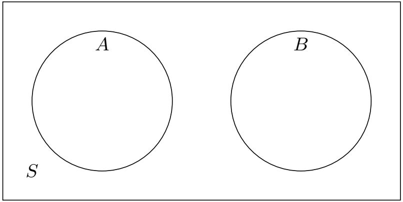
$$x(t) \longrightarrow \boxed{h(t)} \longrightarrow y(t) = x(t) * h(t)$$

$$X(f) \longrightarrow \boxed{H(f)} \longrightarrow Y(f) = X(f)H(f)$$

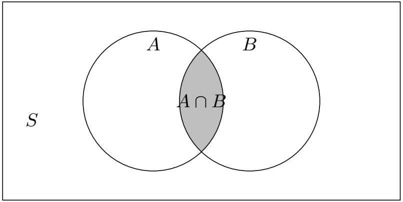
# Why Probability and Stochastic Process?

- Probability: the core mathematical tool for communication theory
- Stochastic model: widely used in the study of communication systems
- Random factors in a radio communication system
  - Message is random: no randomness, no information
  - Noise and interference are random
  - Many other factors are random (e.g., delay, phase, fading)
- Real-world applications
  - Stock market modelling, gambling, weather forecast, control systems, ...

- Sample space $S$: set of all possible outcomes of an experiment
- Event $A \subseteq S$: any subset of $S$




$A \cap B = \emptyset$

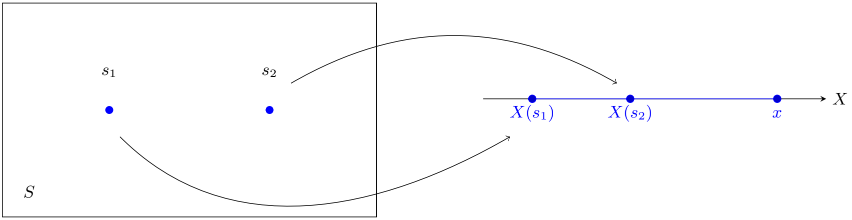



$A \cap B \neq \emptyset$

- Probability of an event: a non-negative number assigned to the event:
  1. The probability of the event that includes *all possible outcomes* is 1, i.e., $P(S) = 1$
  2. Probability of event $A$ is nonnegative, i.e., $P(A) \geq 0$
  3. For two exclusive events $A \cap B = \emptyset$, $P(A \cup B) = P(A) + P(B)$

**Example: prove $P(A) \leq P(B)$ if $A \subset B$**

- Random variable $X(s)$: a real-valued function defined on the set of all possible outcomes $S$




- $\{X \leq x\}$ : a subset of $S$ consisting of all outcomes $s$ such that $X(s) \leq x$
- Random variable $X(s)$ is often denoted by $X$ for simplicity.

### Example: coin toss

- $S = \{HH, HT, TH, TT\}$: all possible outcomes of two coin tosses
- Random variable $X$: number of heads in two coin tosses:
  $X(HH) = 2, X(HT) = X(TH) = 1, X(TT) = 0$; $\{X \leq 1\} = \{HT, TH, TT\}$

# cdf and pdf

- Cumulative distribution function (cdf) of a random variable:

$$F_X(x) = P(X \le x)$$

- Properties

$$F_X(-\infty) = 0, \quad F_X(\infty) = 1$$
$$F_X(x_1) \le F_X(x_2), \quad \text{if } x_1 \le x_2$$

- Probability density function (pdf):

$$f_X(x) = \frac{dF_X(x)}{dx}$$
$$F_X(x) = \int_{-\infty}^{x} f_X(y)dy$$
$$P(a < X \le b) = F_X(b) - F_X(a) = \int_a^b f_X(y)dy$$
$$f_X(x) = \frac{dF_X(x)}{dx} \ge 0 \quad \text{since } F_X(x) \text{ is non-decreasing}$$

# Mean and Variance

- Mean or expectation (corresponding to DC level of signals):

$$\mathbb{E}\{X\} = \mu_X = \int_{-\infty}^{\infty} x f_X(x) dx, \quad \mathbb{E}\{\cdot\} \text{ is the expectation operator}$$

- Variance (corresponding to power of zero-mean signals):

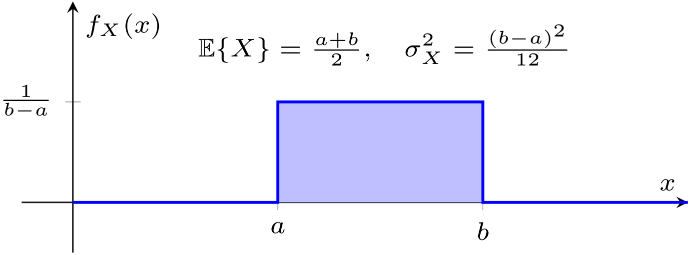
$$\sigma_X^2 = \mathbb{E}\{(X - \mu_X)^2\} = \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x) dx = \mathbb{E}\{X^2\} - \mu_X^2$$

## Evaluation of mean

If $Y = G(X)$, then

$$\mathbb{E}\{Y\} = \int_{-\infty}^{\infty} G(x) f_X(x) dx$$

# Uniform Distribution



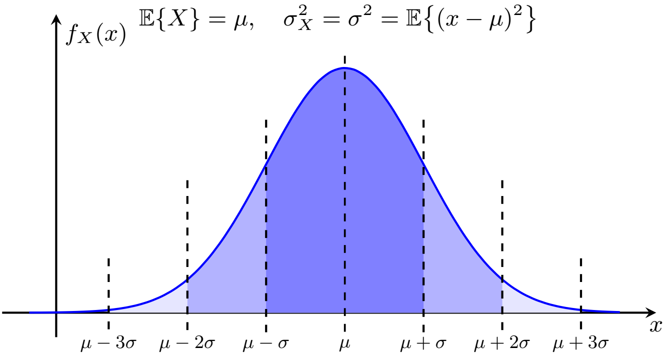

$$\mathbb{E}\{X\} = \frac{a+b}{2}, \quad \sigma_X^2 = \frac{(b-a)^2}{12}$$

$$f_X(x) = \begin{cases} \frac{1}{b-a}, & a \le x \le b \\ 0, & \text{elsewhere} \end{cases}$$

$$F_X(x) = \begin{cases} 0, & x < a \\ \frac{x-a}{b-a}, & a \le x \le b \\ 1, & x > b \end{cases}$$

$$X \sim U(a, b)$$

# Normal (Gaussian) Distribution




$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

$$F_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2} dy$$

$$X \sim \mathcal{N}(\mu, \sigma^2) \quad \text{(sometimes, } X \sim \mathcal{N}(\mu, \sigma)\text{)}$$

# Q Function and Error Function

- Q-function:

$$Q(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt, \quad x \geq 0$$

- Error function:

$$\mathrm{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty \exp(-t^2) dt$$

$$\mathrm{erf}(x) = 1 - \mathrm{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp(-t^2) dt$$

# Note

# Lecture 3: Random Variables and Stochastic Processes

Dr. Geoffrey Li

Department of Electrical & Electronic Engineering
Imperial College London

# Outline

- Random Variables
  - Joint distribution
  - Independence and uncorrelation
- Stochastic Processes
  - Mean and autocorrelation function
  - Gaussian process
  - Wide-sense stationary (WSS) process
- References
  - [Haykin] Chapter 5
  - [Lathi] Chapter 8

- Joint cdf for two random variables $X$ and $Y$

$$F_{XY}(x, y) = P(X \leq x, Y \leq y)$$

- Joint pdf

$$f_{XY}(x, y) = \frac{\partial^2 F_{XY}(x, y)}{\partial x \partial y}$$

- Properties
  1. $F_{XY}(\infty, \infty) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{XY}(u, v) du dv = 1$
  2. $f_X(x) = \int_{y=-\infty}^{\infty} f_{XY}(x, y) dy$, $f_Y(y) = \int_{x=-\infty}^{\infty} f_{XY}(x, y) dx$
  3. Independent: $f_{XY}(x, y) = f_X(x) f_Y(y)$
  4. Uncorrelated: $\mathbb{E}\{XY\} = \mathbb{E}\{X\}\mathbb{E}\{Y\}$ or $\mathbb{E}\{(X - \mu_X)(Y - \mu_Y)\}$

# Independent and Uncorrelated

- Independent $\Rightarrow$ uncorrelated
- Uncorrelated $\not\Rightarrow$ independent
- For jointly Gaussian random variables, uncorrelated $\iff$ independent

# Joint Distribution of $n$ random variables

- Joint cdf

$$F_{X_1 X_2 \dots X_n}(x_1 x_2 \dots x_n) = P(X_1 \le x_1, X_2 \le x_2, \dots X_n \le x_n)$$

- Joint pdf

$$f_{X_1 X_2 \dots X_n}(x_1 x_2 \dots x_n) = \frac{\partial^n F_{X_1 X_2 \dots X_n}(x_1 x_2 \dots x_n)}{\partial x_1 \partial x_2 \dots \partial x_n}$$

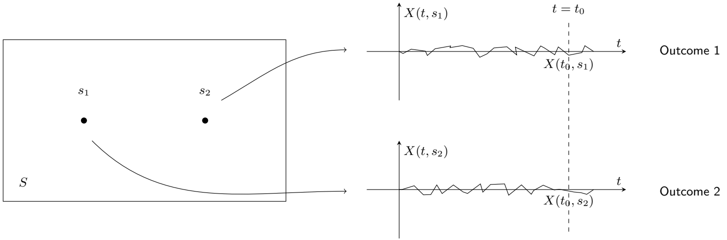
- Independent

$$F_{X_1 X_2 \dots X_n}(x_1 x_2 \dots x_n) = F_{X_1}(x_1) F_{X_2}(x_2) \dots F_{X_n}(x_n)$$
$$f_{X_1 X_2 \dots X_n}(x_1 x_2 \dots x_n) = f_{X_1}(x_1) f_{X_2}(x_2) \dots f_{X_n}(x_n)$$

- i.i.d. (independent and identically distributed)
  - Independent random variables with the same distribution (e.g., flipping $n$ coins)

# Stochastic Process

- Stochastic process $X(t,s)$: a collection of random variables over time. It represents the *evolution* of a random system.
- At a given time $t_0$, $X(t_0,s)$ is a random variable.
- At a sample outcome $s_j$, $X(t,s_j)$ is a deterministic function over time.
- Stochastic process $X(t,s)$ is often denoted by $X(t)$ for simplicity.
- Noise is often modelled as a Gaussian stochastic process.



# Statistics of Stochastic Processes

- Probability density function
  - 1st order: $f_X(x;t)$
  - 2nd order: $f_X(x_1, x_2; t_1, t_2)$
  - n-th order: $f_X(x_1, \ldots, x_n; t_1, \ldots, t_n)$
- Mean is usually a function of $t$:

$$\mu_X(t) = \mathbb{E}\{X(t)\} = \int_{-\infty}^{\infty} x f_X(x;t)dx$$

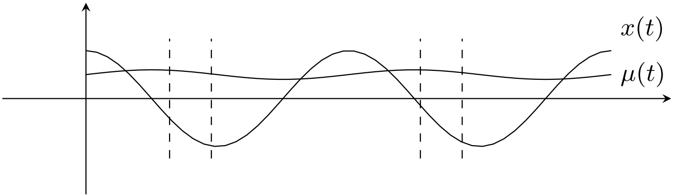
- Autocorrelation function is usually a function of $t_1$ and $t_2$. It measures the correlation between samples:

$$R_X(t_1, t_2) = \mathbb{E}\{X(t_1)X(t_2)\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 x_2 f_X(x_1, x_2; t_1, t_2)dx_1 dx_2$$

- A stochastic process is Gaussian if and only if the pdf $f_X(x;t)$ is Gaussian at any time $t_n$.

# Wide-Sense Stationary (WSS) Stochastic Processes

- A stochastic process is wide-sense stationary (WSS) if and only if:
  1. The mean is not a function of time:
  $$\mu_X(t) = \mu_X, \quad \forall t$$
  2. The autocorrelation function only depends on time difference:
  $$R_X(t+\tau, t) = R_X(\tau), \quad \forall t, \tau$$




- Noise and message signals are often modelled as WSS processes.

# Properties of Autocorrelation Function

For a real WSS process $X(t)$ with autocorrelation function $R_X(\tau)$:

1. $R_X(0) = \mathbb{E}\{X^2(t)\}$
2. $R_X(\tau)$ is an even function
$$R_X(\tau) = \mathbb{E}\big\{x(t+\tau)x(t)\big\} = \mathbb{E}\big\{x(t)x(t+\tau)\big\} = R_X(-\tau)$$
3. $R_X(\tau)$ takes maximum magnitude at $\tau = 0$ (Homework 1 Problem 5)
$$|R_X(\tau)| \leq R_X(0)$$

$R_X(\tau)$ can tell how predictable $X(t)$ is based on $X(t-\tau)$.

# Exercise

## WSS example

Show that sinusoidal wave with random phase in uniform distribution is WSS

$$X(t) = A\cos(\omega_c t + \Theta), \quad f_\Theta(\theta) = \frac{1}{2\pi}, \quad \theta \in [0, 2\pi).$$

## Note

# Lecture 4: More on Stochastic Processes

Dr. Geoffrey Li

Department of Electrical & Electronic Engineering
Imperial College London

# Outline

- Stochastic Processes
  - Passing through linear time-invariant (LTI) system
  - Power spectral density (PSD)
- References
  - [Haykin] Chapter 5
  - [Lathi] Chapter 8

# Stochastic Process Through a Linear Time-Invariant (LTI) System: I

$$X(t) \longrightarrow \boxed{\begin{array}{c}\text{Impulse Response}\\ h(t)\end{array}} \longrightarrow Y(t) = X(t) * h(t) = \int h(\tau)X(t-\tau)d\tau$$

- If $\mathbb{E}\{X(t)\}$ is finite for all $t$ and the system is bounded-input bounded-output (BIBO) stable, we have

$$\mu_Y(t) = \mathbb{E}\{Y(t)\} = \mu_X(t) * h(t)$$

$$\mu_X(t) \longrightarrow \boxed{h(t)} \longrightarrow \mu_Y(t)$$

# Stochastic Process Through an LTI System: II

- If $\mathbb{E}\{X^2(t)\}$ is finite for all $t$ and the system is BIBO stable, we have

$$R_Y(t,u) = h(t) * h(u) * R_X(t,u)$$

$$R_{XX}(t,u) \longrightarrow \boxed{h(u)} \xrightarrow{R_{XY}(t,u)} \boxed{h(t)} \longrightarrow R_{YY}(t,u)$$

# Stochastic Process Through an LTI System: III

- If $X(t)$ is real WSS, then

$$\mu_Y(t) = \mu_X \underbrace{\int_{-\infty}^{\infty} h(\tau)d\tau}_{H(0)} = \mu_X \cdot \text{DC response} = \text{constant}$$

$$\mu_X \longrightarrow \boxed{h(t)} \longrightarrow \mu_Y$$

$$R_Y(\tau) = \int_{-\infty}^{\infty} h(\tau_1)h(\tau_2)R_X(\tau - \tau_1 + \tau_2)d\tau_1 d\tau_2 = h(\tau) * h(-\tau) * R_X(\tau)$$

$$R_X(\tau) \longrightarrow \boxed{h(-\tau)} \longrightarrow \boxed{h(\tau)} \longrightarrow R_Y(\tau)$$

- For a complex WSS process $X(t)$, $R_X(\tau) = \mathbb{E}\{X(t+\tau)X^*(t)\}$

$$R_Y(\tau) = h(\tau) * h^*(-\tau) * R_X(\tau)$$

- If $X(t)$ is WSS, then $Y(t)$ is also WSS

$$\text{WSS Process} \longrightarrow \boxed{\text{LTI System}} \longrightarrow \text{WSS Process}$$

- If $X(t)$ is a Gaussian process, then $Y(t)$ is also a Gaussian process

$$\text{Gaussian Process} \longrightarrow \boxed{\text{LTI System}} \longrightarrow \text{Gaussian Process}$$

# Power Spectral Density (PSD): I

- PSD measures the distribution of power of a random process over its spectrum.
- PSD is defined only for WSS processes.

**Einstein-Wiener-Khintchine relation:**

The PSD of a wide sense stationary process is equal to the Fourier transform of its autocorrelation function:

$$S_X(f) = \int_{-\infty}^{\infty} R_X(\tau) e^{-j2\pi f \tau} d\tau \geq 0$$

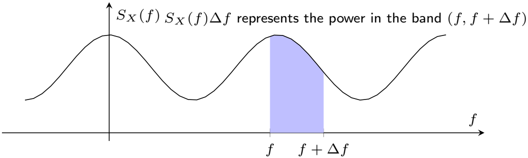
- The frequency content of a process depends on how rapidly the amplitude changes as a function of time (can be measured by the autocorrelation function).

$$R_X(\tau) = \int_{-\infty}^{\infty} S_X(f) e^{j2\pi f \tau} df$$

$$R_X(0) = \int_{-\infty}^{\infty} S_X(f) df$$

- $S_X(f)$ is real and nonnegative.

# Power Spectral Density (PSD): II




The average power of a random process $X(t)$

$$P = \mathbb{E}\{X^2(t)\} = R_X(0) = \int_{-\infty}^{\infty} S_X(f) df$$

$$X(t) \longrightarrow \boxed{\begin{array}{c}\text{Impulse Response}\\ h(t)\end{array}} \longrightarrow Y(t) = X(t) * h(t) = \int h(\tau)X(t-\tau)d\tau$$

- For a real or complex WSS process $X(t)$ goes through an LTI system:

$$S_Y(f) = |H(f)|^2 S_X(f)$$

# Exercise

## WSS process and random variable

Let $X(t)$ be a WSS process with autocorrelation function $R_X(\tau)$, and $\Theta$ be a random variable independent of $X(t)$ and uniformly distributed in $[0, 2\pi)$.

1. Prove $Z(t) = X(t)\cos(2\pi f_c t)$ is not WSS;
2. Prove $Y(t) = X(t)\cos(2\pi f_c t + \Theta)$ is WSS, and find the PSD of $Y(t)$.

# Note

# Lecture 5: Baseband and Passband Signals

Dr. Geoffrey Li

Department of Electrical & Electronic Engineering
Imperial College London

# Outline

- Energy and power
- Bandwidth
  - Real-valued signal and one-sided bandwidth
  - Complex-valued signal and two-sided bandwidth
- Additive white Gaussian noise (AWGN) channel
- Baseband and passband signals
- Upconversion and downconversion
- Representation of passband signals
- Hilbert transform and pre-envelope
- References
  - [Haykin] Chapter 2

# Energy and Power

Energy and power: two important concepts in communications

- How much power is needed to transmit a signal?
- How is the signal-to-noise ratio found?
- How much interference do signals create for each other?

Energy: the area under the squared magnitude of a signal

$$E = \int_{-\infty}^{\infty} |s(t)|^2 dt = \int_{-\infty}^{\infty} |S(f)|^2 df.$$

Power: time average of energy, evaluated over a period or large interval

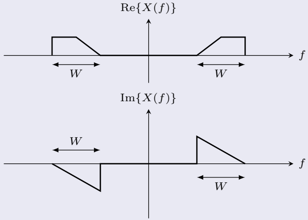
$$P = \frac{1}{T} \int_{-T/2}^{T/2} |s(t)|^2 dt.$$

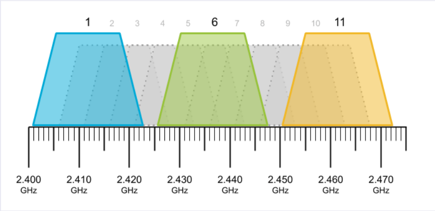
# Bandwidth: One-Sided

Bandwidth: the frequencies range of a signal. The way we express this range differs:

## One-sided bandwidth → real-valued signals

- Real-valued signals are real in the time domain. Often denoted as "real signals".
- One-sided bandwidth: the range of positive frequencies. For real-valued signals, the negative frequencies are simply mirror images and contain no extra information.
- Physical signals (e.g., light, sound, Wi-Fi) are real-valued.



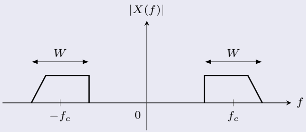




Wi-Fi Bands at 2.4 GHz

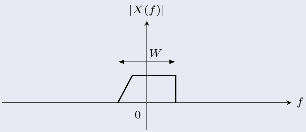
# Bandwidth: Two-Sided

## Two-sided bandwidth → complex-valued signals

- Complex-valued signals can take complex numbers. They describe the *complex envelope* of real-valued *passband* signals.
- Two-sided bandwidth: the range of negative-to-positive frequencies.
- Two-sided bandwidth of a complex-valued signal equals one-sided bandwidth of the corresponding real-valued passband signal.



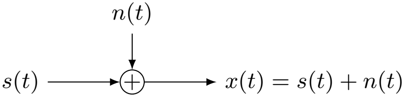

Passband signal




Complex-valued signal in the baseband

# Channels

The additive white Gaussian noise (AWGN) channel is considered in this module:



# Baseband and Passband Signals

## Baseband signals

- Power concentrated in a band around DC

$$U(f) \approx 0, \quad |f| > W$$

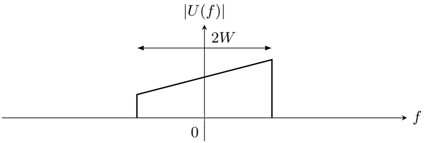
- Complex-valued in general, real-valued in special cases

## Passband signals

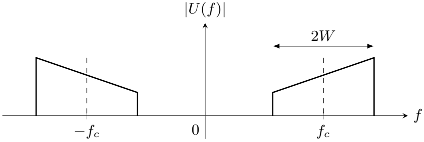
- Power concentrated around carrier frequency $f_c$ that is away from DC

$$U(f) \approx 0, \quad |f \pm f_c| > W, \quad f_c \gg W$$

- Always real-valued



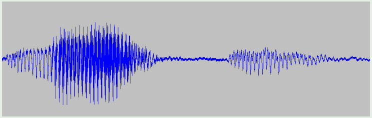

Baseband spectrum: not necessarily symmetric



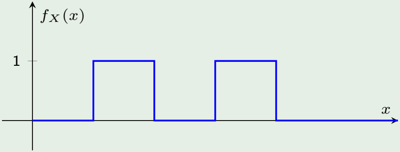

Passband spectrum: symmetric

## Real-valued baseband signals

- Speech and audio



- Two-level digital signal




We often want to send such signals over a passband channel (e.g., Wi-Fi channel with 20 MHz bandwidth).

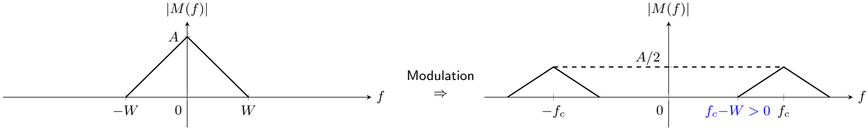
# Modulation: Baseband to Passband

Consider a real-valued baseband message signal $m(t)$ with bandwidth $W$.

Modulation: Translating to passband by multiplying a sinusoid at frequency $f_c \gg W$

$$u_p(t) = m(t)\cos 2\pi f_c t \longrightarrow U_p(f) = \frac{1}{2}\big(M(f-f_c) + M(f+f_c)\big)$$

$$v_p(t) = m(t)\sin 2\pi f_c t \longrightarrow V_p(f) = \frac{1}{2j}\big(M(f-f_c) - M(f+f_c)\big)$$



# I and Q components

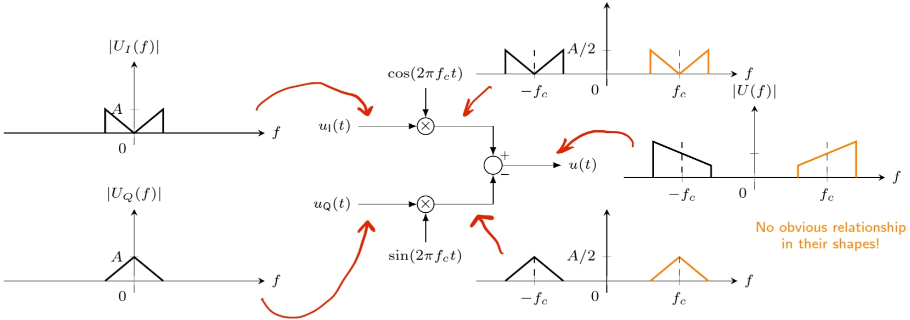
Can we modulate separately using cosine and sine carriers? Yes.

Sinusoids are rapidly varying but predictable (contain no info)

$$\underbrace{u(t)}_{\text{Passband signal}} = \underbrace{u_{\mathsf{I}}(t)}_{\text{In-phase (I) component}} \cos(2\pi f_c t) - \underbrace{u_{\mathsf{Q}}(t)}_{\text{Quadrature (Q) component}} \sin(2\pi f_c t)$$

Real baseband signals (contain all the info)

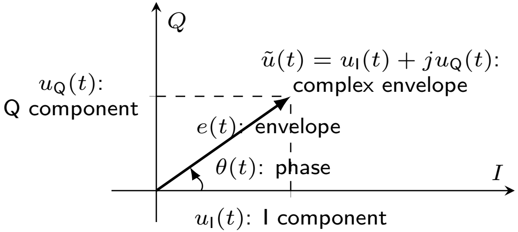
- $\cos(2\pi f_c t)$ and $\sin(2\pi f_c t)$ are rapidly varying but predictable (contain no information)
- $u_{\mathsf{I}}(t)$ and $u_{\mathsf{Q}}(t)$ are real baseband signals (contain all the information)
- How do we get back the I and Q components from the passband signal?
- Can any passband signal be decomposed into I and Q components?




Since $u_I(t)$ and $u_Q(t)$ are real, their spectra are conjugate symmetric

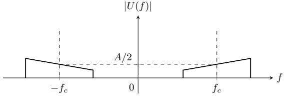
- Block diagram follows directly from the equation defining the modulated signal
- Happens at the transmitter

- Passband signal can be mapped to a pair of real baseband signals
- That is, passband modulation is two-dimensional
- We can also plot it on the complex plane

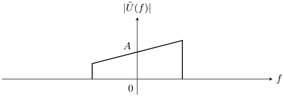


# Complex Envelope and Passband Signal

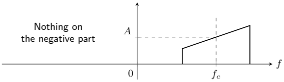
$$\underbrace{u(t)}_{\substack{\text{real-valued}\\ \text{passband signal}}} = \mathrm{Re}\{\underbrace{\tilde{u}(t)}_{\substack{\text{complex envelope}\\ \text{baseband signal}}} \exp\left(j2\pi f_c t\right)\}$$




$\mathcal{F}\{u(t)\}$




$\mathcal{F}\{\tilde{u}(t)\}$




$\mathcal{F}\{\tilde{u}(t)\exp(j2\pi f_c t)\}$

All information in a passband signal is contained in its complex envelope.

# Time Domain Expressions for a Passband Signal

## Passband signal expressions

- In I and Q components

$$u(t) = u_{\mathsf{I}}(t)\cos(2\pi f_c t) - u_{\mathsf{Q}}(t)\sin(2\pi f_c t)$$

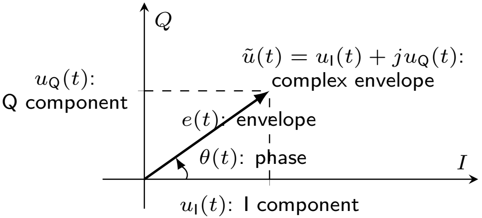
- In envelope and phase

$$u(t) = e(t)\cos\big(2\pi f_c t + \theta(t)\big)$$

- In complex envelope

$$u(t) = \mathrm{Re}\big\{\tilde{u}(t)\exp(j2\pi f_c t)\big\}$$

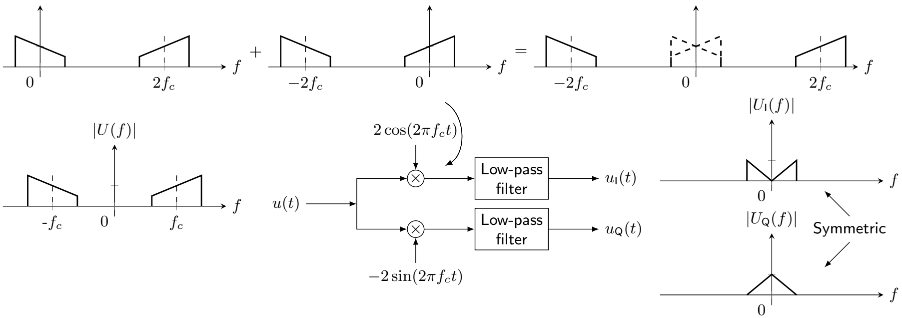
Starting from one representation, we can derive the rest based on the relations depicted in the figure.




$$e(t) = \sqrt{u_{\mathsf{I}}^2(t) + u_{\mathsf{Q}}^2(t)},$$
$$\theta(t) = \arg(u_{\mathsf{I}}(t) + ju_{\mathsf{Q}}(t))$$

Complex baseband signal

$$\tilde{u}(t) = u_{\mathsf{I}}(t) + ju_{\mathsf{Q}}(t) = e(t)\exp\big(j\theta(t)\big)$$




Receiver needs to be coherent: same phase and frequency of the copy of the carrier at the receiver as those of the incoming signal

# Hilbert Transform (HT)

- Hilbert transform of a signal $g(t)$: a linear transformation, defined as
$$\hat{g}(t) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{g(\tau)}{t-\tau} d\tau = g(t) * \frac{1}{\pi t}$$
- Inverse Hilbert transform
$$g(t) = -\frac{1}{\pi}\int_{\infty}^{\infty} \frac{\hat{g}(\tau)}{t-\tau} d\tau = -\hat{g}(t) * \frac{1}{\pi t}$$
- HT of $\hat{g}(t)$ is $-g(t)$: $\hat{\hat{g}}(t) = -g(t)$
- In the frequency domain, we have
$$\mathcal{F}\left\{\frac{1}{\pi t}\right\} = -j\text{sgn}(f) = \begin{cases} -j, & f > 0, \\ 0, & f = 0, \\ j, & f < 0 \end{cases}$$
$$\hat{G}(f) = \mathcal{F}\left(\frac{1}{\pi t}\right)G(f) = -j\text{sgn}(f)G(f), \quad \text{sgn}(f) = \begin{cases} +1, & f > 0, \\ 0, & f = 0, \\ -1, & f < 0 \end{cases}$$
- HT introduces a phase shift of -90 degrees for all positive frequencies of the input signal, and +90 degrees for all negative frequencies.

# Pre-Envelope

- Define the pre-envelope of a real signal $u(t)$ as the complex-valued function

$$u_+(t) = u(t) + j \underbrace{\hat{u}(t)}_{\mathcal{F}^{-1}\{-j\mathrm{sgn}(f)U(f)\}}$$

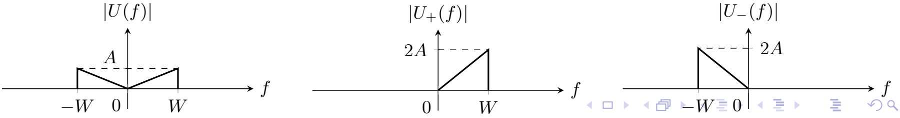
- Its Fourier transform:

$$U_+(f) = U(f) + \mathrm{sgn}(f)U(f) = \begin{cases} 2U(f), & f > 0, \\ U(0), & f = 0, \\ 0, & f < 0 \end{cases}$$

- Pre-envelope removes the negative frequency components
- Similarly define the pre-envelope for negative frequencies

$$u_-(t) = u(t) - j\hat{u}(t)$$

$$U_-(f) = U(f) - \mathrm{sgn}(f)U(f) = \begin{cases} 0, & f > 0, \\ U(0), & f = 0, \\ 2U(f), & f < 0 \end{cases}$$



## Example

Consider arbitrary complex-valued baseband signal $\tilde{u}(t)$, whose spectrum is limited to $[-W, +W]$. Define

$$u(t) = \mathrm{Re}\{\tilde{u}(t)\exp(j2\pi f_c t)\}$$

Show that u(t) is a real-valued passband signal concentrated around $\pm f_c$.

# Note

# Lecture 6: Noise

Dr. Geoffrey Li

Department of Electrical & Electronic Engineering
Imperial College London

# Outline

- What is noise?
- White noise and Gaussian noise
- Lowpass noise
- Bandpass noise
  - In-phase/quadrature representation
  - Phasor representation
- References
  - [Haykin] Chapter 5

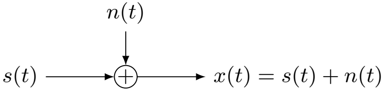
# Noise

Noise: unwanted waves disturbing the transmission of signals.

- Where does noise come from?
  - External sources: e.g., atmospheric, galactic noise, interference.
  - Internal sources: generated by communication devices themselves.
    - A basic limitaion on communication systems
    - Shot noise: usually in vacuum tubes or transistors
    - Thermal noise: caused by rapid and random motion of electrons due to thermal agitation
- Stationary and zero-mean Gaussian distributions.

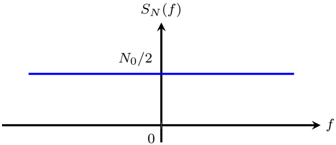
# White Noise

- The additive noise channel
  - $n(t)$ models all types of noise
  - Zero mean




- White noise
  - Flat PSD over *all frequencies*

$$S_N(f) = \frac{N_0}{2}, \quad -\infty < f < \infty$$

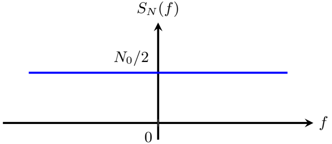


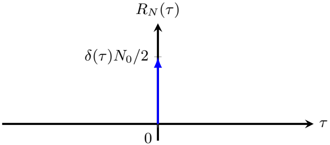

  - Half the power $\frac{N_0}{2}$ associated with positive frequencies and half with negative
  - The term white analogous to white light, indicating the shape of the PSD!
  - Defined for stationary noise
  - There are also non-stationary noises, but definitions are complicated

- *Infinite* bandwidth: a purely theoretic assumption, valid for flat PSD over the bandwidth of interest

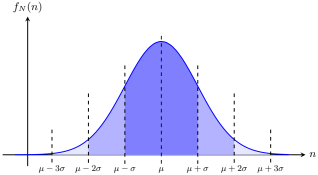
# White and Gaussian Noise

- White noise: shape of PSD is flat!
  - Autocorrelation function of $n(t)$: $R_N(\tau) = \mathcal{F}^{-1}\{S_N(f)\} = \frac{N_0}{2}\delta(\tau)$
  - Uncorrelated samples at different time instants
  - Also colored noise
  - White noise: either Gaussian or non-Gaussian




- Gaussian noise:
  - Gaussian distribution for a sample at any *time instant*
  - Colored or white Gaussian noise

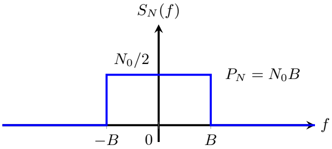



- White noise (shape of PSD) and Gaussian noise (distribution, CDF) are different concepts
- In communications, typically additive white Gaussian noise (AWGN)

# Ideal Low-pass White Noise

- White noise passing through an ideal low-pass filter of bandwidth $B$ (counting only positive part!)

$$S_N(f) = \begin{cases} \frac{N_0}{2}, & |f| \le B, \\ 0, & \text{otherwise} \end{cases}$$

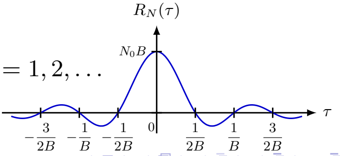



- By Einstein-Wiener-Khinchine relation, autocorrelation function

$$R_N(\tau) = \mathcal{F}^{-1}\{S_N(f)\} = N_0 B \mathrm{sinc}(2B\tau), \quad \mathrm{sinc}(x) = \frac{\sin \pi x}{\pi x} \quad \text{(Normalized definition)}$$

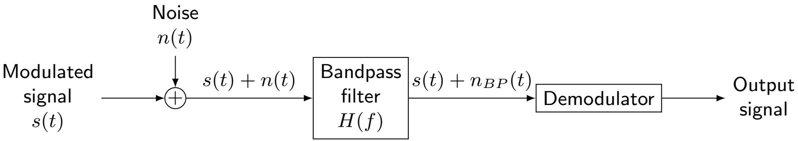
- Samples at Nyquist frequency $2B$ are uncorrelated, also independent if it is Gaussian

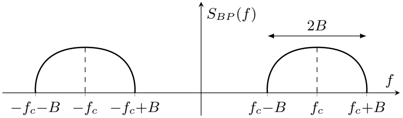
$$R_N(\tau) = 0, \quad \text{for } 2\pi B\tau = k\pi \text{ or } \tau = \frac{k}{2B},\ k = 1, 2, \ldots$$



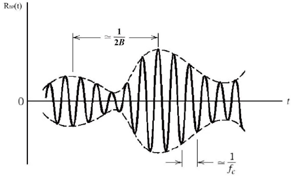
# WSS Bandpass Noise

Noise in a communication system with a bandpass filter of bandwidth $2B$ (counting only positive frequency, but both sides of $f_c$)




PSD

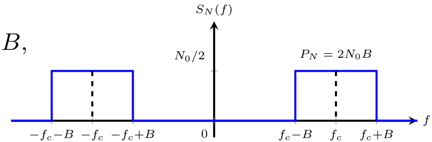



Autocorrelation

# Example

White noise with PSD of $N_0/2$ passing through an ideal band-pass filter

$$S_N(f) = \begin{cases} \frac{N_0}{2}, & |f - f_c| \le B \text{ or } |f + f_c| \le B, \\ 0, & \text{otherwise} \end{cases}$$

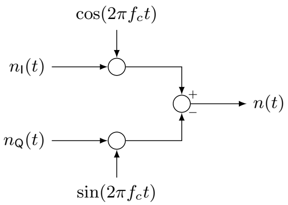


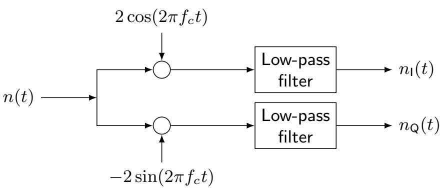
# Bandpass Noise

- $n(t)$ in canonical form:

$$n(t) = \underbrace{n_\mathsf{I}(t)}_{\text{real}}\cos(2\pi f_c t) - \underbrace{n_\mathsf{Q}(t)}_{\text{real}}\sin(2\pi f_c t)$$

  where $\mathbb{E}\{n_\mathsf{I}(t)\} = \mathbb{E}\{n_\mathsf{Q}(t)\} = 0$, $R_{N_\mathsf{I}}(\tau) = R_{N_\mathsf{Q}}(\tau)$, $R_{N_\mathsf{I}N_\mathsf{Q}}(\tau) = -R_{N_\mathsf{I}N_\mathsf{Q}}(-\tau)$
- $n_\mathsf{I}(t)$ and $n_\mathsf{Q}(t)$: fully representative of the band-pass noise
  - Given band-pass noise, one may extract in-phase and quadrature components (using LPF of bandwidth $B$).
  - Given the two components, one may generate band-pass noise. This is useful in computer simulation.



# Properties of Baseband Noise

- If noise $n(t)$ is Gaussian, then so are $n_\mathsf{I}(t)$ and $n_\mathsf{Q}(t)$
- $R_N(\tau) = R_{N_\mathsf{I}}(\tau)\cos(2\pi f_c \tau) + R_{N_\mathsf{Q}}(\tau)\sin(2\pi f_c \tau)$
  - $R_N(\tau) = R_{N_\mathsf{I}}(\tau)\cos(2\pi f_c \tau)$ if $n_\mathsf{I}(t)$ and $n_\mathsf{Q}(t)$ are uncorrelated
- $n_\mathsf{I}(t)$ and $n_\mathsf{Q}(t)$ have the same variance (i.e., same power) as $n(t)$
- Both in-phase and quadrature components have the same PSD:

$$S_{N_I}(f) = S_{N_Q}(f) = \begin{cases} S_N(f - f_c) + S_N(f + f_c), & |f| \leq B \\ 0, & \text{otherwise} \end{cases}$$

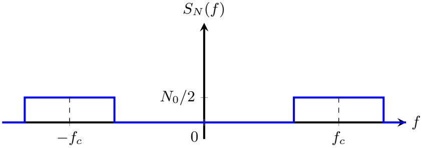
# Noise Power: Narrowband White Noise

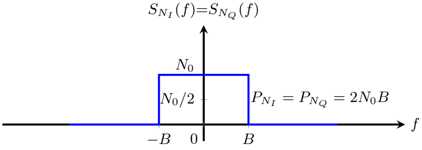
- For ideally filtered narrowband white noise, the PSDs of $n_{\mathsf{I}}(t)$ and $n_{\mathsf{Q}}(t)$:

$$S_{N_I}(f) = S_{N_Q}(f) = \begin{cases} N_0, & |f| \leq B, \\ 0, & \text{otherwise} \end{cases}$$

- The average power in *each* of the baseband waveforms $n_{\mathsf{I}}(t)$ and $n_{\mathsf{Q}}(t)$ is identical to the average power in the bandpass noise waveform $n(t)$.
- For ideally filtered narrowband noise, the variance of $n_{\mathsf{I}}(t)$ and $n_{\mathsf{Q}}(t)$:

$$P_{N_I} = P_{N_Q} = 2N_0B$$



# Phasor Representation

Bandpass noise in alternative form:

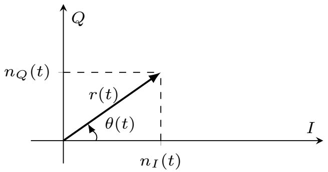
$$n(t) = n_{\mathsf{I}}(t)\cos(2\pi f_c t) - n_{\mathsf{Q}}(t)\sin(2\pi f_c t) = r(t)\cos\big(2\pi f_c t + \phi(t)\big)$$

- $r(t) = \sqrt{n_{\mathsf{I}}^2(t) + n_{\mathsf{Q}}^2(t)}$: the envelope of the noise following Rayleigh distribution

$$f(r) = \frac{r}{\sigma^2}\exp(-\frac{r^2}{2\sigma^2})$$

- $\phi(t) = \arg(n_{\mathsf{I}}(t) + j n_{\mathsf{Q}}(t))$: the phase of the noise following uniform distribution

$$f(\theta) = \begin{cases} \frac{1}{2\pi}, & 0 \le x < 2\pi, \\ 0, & \text{otherwise} \end{cases}$$



# Summary

- White noise: constant PSD over an infinite bandwidth
- Gaussian noise: Gaussian distribution for any sample
- Bandpass noise:
  - In-phase and quadrature components $n_\mathsf{I}(t)$ and $n_\mathsf{Q}(t)$ are low-pass random processes.
  - The same PSD for $n_\mathsf{I}(t)$ and $n_\mathsf{Q}(t)$
  - $\mathbb{E}\{n_\mathsf{I}^2(t)\} = \mathbb{E}\{n_\mathsf{Q}^2(t)\} = \mathbb{E}\{n^2(t)\}$

# Note

# Lecture 7: Noise Performance of Double Sideband (DSB)

Dr. Geoffrey Li

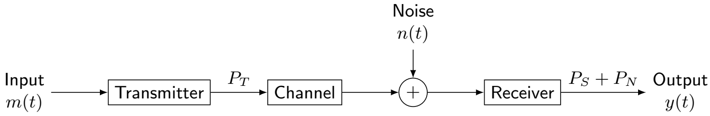
Department of Electrical & Electronic Engineering
Imperial College London

# Outline

- SNR of baseband analogue transmission
- Noise in Double Sideband-Suppressed Carrier (DSB-SC)
- SNR of DSB-SC
- References
    - [Haykin] Chapter 6

# Noise in Analog Communication Systems

- How do various analog modulation schemes perform in the presence of noise?
- Which scheme performs best?
- How can we measure its performance?



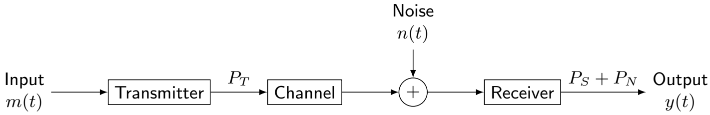
# Signal-to-Noise Ratio

Signal-to-noise ratio (SNR) at the output of the receiver:

$$\text{SNR} \triangleq \frac{\text{average power of message signal at the receiver output}}{\text{average power of noise at the receiver output}} = \frac{P_S}{P_N}$$

- Normally expressed in decibels (dB): $\text{SNR}_{\text{dB}} = 10\log_{10}(\text{SNR})$
- Managing the wide range of power levels

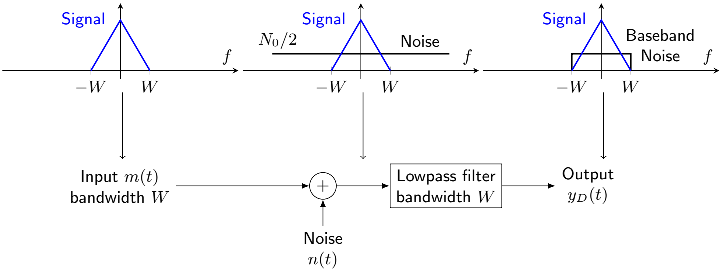
# Transmitted Power




- Higher transmitted power $P_T \Rightarrow$ higher received power $P_S \Rightarrow$ higher SNR
- Limited by: equipment capability, battery life, cost, government restrictions, interference with other channels, . . .
- For a fair comparison between different modulation schemes, $P_T$ should be the same & so is PSD of noise
- Baseband SNR, $\text{SNR}_{\text{baseband}}$: calibrate and compare the SNR values we obtain

# A Baseband Communication System

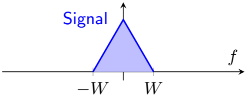
- No modulation
- Suitable for transmission over wires
- Transmit power $P_T$ = message power $P$
- Unit channel gain or no propagation loss $P_S = P_T = P$
- Results can be extended to bandpass systems



# Baseband SNR

- Average signal (= message) power

$$P = \text{area under the triangular curve}$$

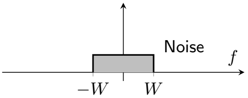



- Noise power
  - AWGN power spectral density

$$\text{PSD} = N_0/2$$

  - average noise power at receiver

$$P_N = \text{area under the straight line} = 2WN_0/2 = WN_0$$



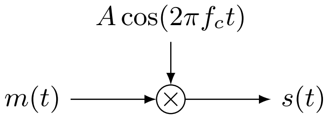

- SNR at receiver output:

$$\text{SNR}_{\text{baseband}} = \frac{P_T}{N_0 W}$$

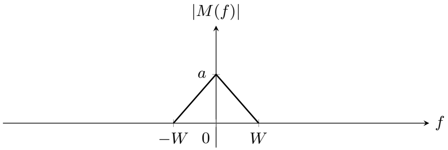
- Improve SNR by:
  - increasing the transmitted power $P_T \uparrow$
  - making the channel/receiver less noisy $N_0 \downarrow$

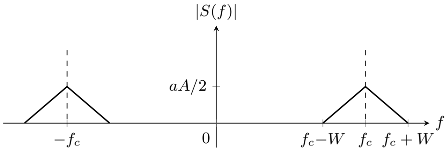
# Double Sideband-Suppressed Carrier (DSB-SC) Modulation

$$s(t) = m(t)A\cos(2\pi f_c t)$$



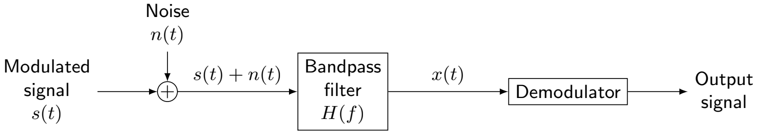

- $A$: amplitude of the carrier
- $f_c$: carrier frequency
- $m(t)$: message signal with $\mathbb{E}\{|m(t)|^2\} = P$



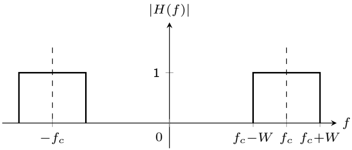



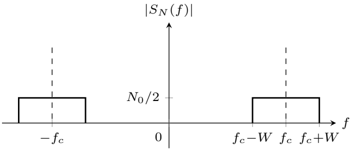
# DSB-SC Receiver with Noise

- Receiver




- Received signal
  - Bandpass filter

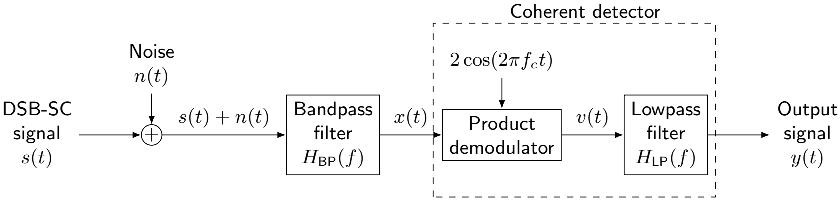



  - Received and filtered noisy signal:

$$
\begin{aligned}
x(t) &= s(t) + n(t) \\
&= Am(t)\cos(2\pi f_c t) + n_{\mathsf{I}}(t)\cos(2\pi f_c t) - n_{\mathsf{Q}}(t)\sin(2\pi f_c t) \\
&= \big(Am(t) + n_{\mathsf{I}}(t)\big)\cos(2\pi f_c t) - n_{\mathsf{Q}}(t)\sin(2\pi f_c t)
\end{aligned}
$$

- Synchronous detection = Product detection = Coherent detection
- "Detection" and "demodulation" are used interchangeably

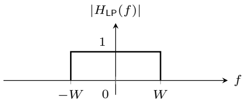



- Multiply $x(t)$ with $2\cos(2\pi f_c t)$:

$$\begin{aligned} v(t) &= 2\cos(2\pi f_c t)\cdot x(t) \\ &= \big(Am(t)+n_{\mathsf{I}}(t)\big)\big(\cos(4\pi f_c t)+1\big) - n_{\mathsf{Q}}(t)\sin(4\pi f_c t) \\ &= \big(Am(t)+n_{\mathsf{I}}(t)\big) + \underbrace{\big(Am(t)+n_{\mathsf{I}}(t)\big)\cos(4\pi f_c t) - n_{\mathsf{Q}}(t)\sin(4\pi f_c t)}_{\text{High frequency around } \pm 2f_c} \end{aligned}$$

- LP filter output

$$y(t) = Am(t) + n_{\mathsf{I}}(t)$$



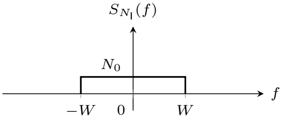

- Signal power at the receiver output

$$P_S = \mathbb{E}\{A^2 m^2(t)\} = A^2 \mathbb{E}\{m^2(t)\} = A^2 P$$

- Noise power

$$S_{N_{\mathsf{I}}}(f) = \begin{cases} S_N(f+f_c) + S_N(f-f_c) = N_0, & |f| \leq W \\ 0, & \text{otherwise} \end{cases}$$

$$P_N = \int_{-W}^{W} N_0 df = 2N_0 W$$



# Comparison

- SNR at the receiver output:

$$\text{SNR}_{\text{DSB-SC}} = \frac{P_S}{P_N} = \frac{A^2 P}{2N_0 W}$$

- Transmit power:

$$P_T = \mathbb{E}\{A^2 m^2(t)\cos^2(2\pi f_c t)\} = \frac{1}{T}\int_0^T A^2 \cos^2(2\pi f_c t)\mathbb{E}\{m^2(t)\}dt$$
$$= \frac{1}{T}\int_0^T A^2 \cos^2(2\pi f_c t) P dt = \frac{A^2 P}{2}$$

- Compare with baseband transmission $\text{SNR}_{\text{DSB-SC}} = \frac{A^2 P}{2N_0 W} = \frac{P_T}{N_0 W} = \text{SNR}_{\text{baseband}}$

- Conclusion: DSB-SC system has the same SNR performance as a baseband system

# Note

# Lecture 8: Noise Performance of Single Sideband (SSB) and Conventional Amplitude Modulation (AM)

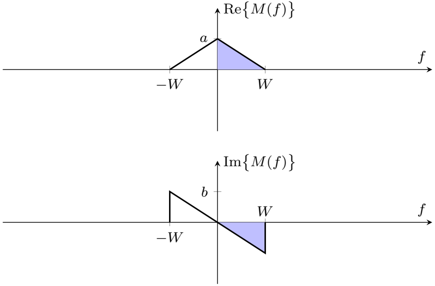
Dr. Geoffrey Li

Department of Electrical & Electronic Engineering
Imperial College London

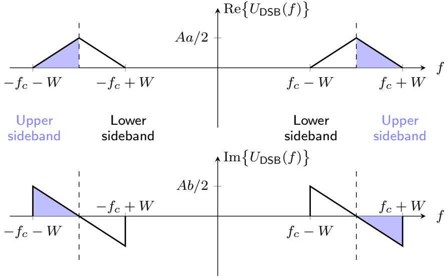
# Outline

- Noise in SSB
- Noise in standard AM
  - coherent detection (of theoretic interest only)
  - envelope detection
- SNR of SSB and AM
- References
  - [Haykin] Chapter 6

# Double Sideband to Single Sideband



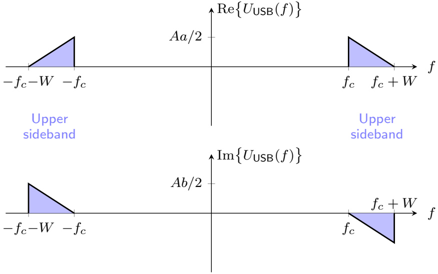

Real message signal has conjugate symmetric spectrum



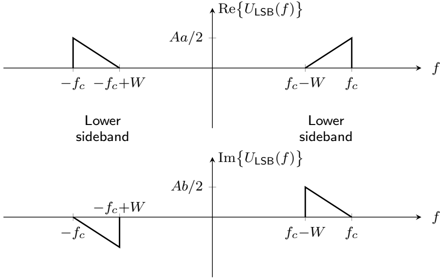

DSB signal spectrum

One side band is enough to reconstruct the message!

# Single Sideband (SSB)



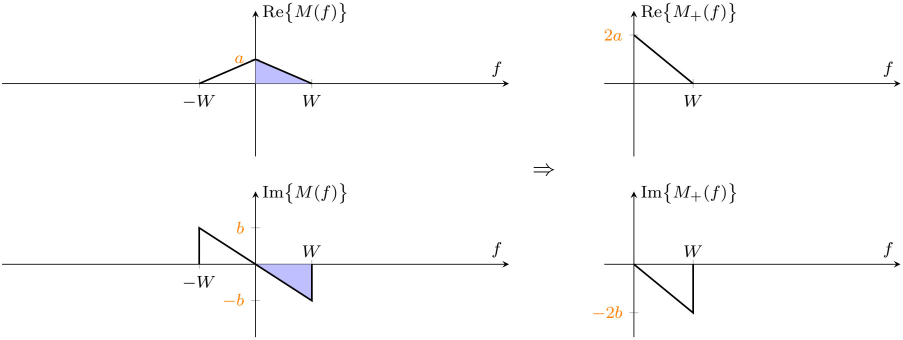

Upper sideband signaling (high freq. side)




Lower sideband signaling (lower freq. side)

Message can be recovered by moving SSB components left and right by $f_c$, and low pass filtering (just like DSB).

# Upper Sideband SSB Modulated Signal: I

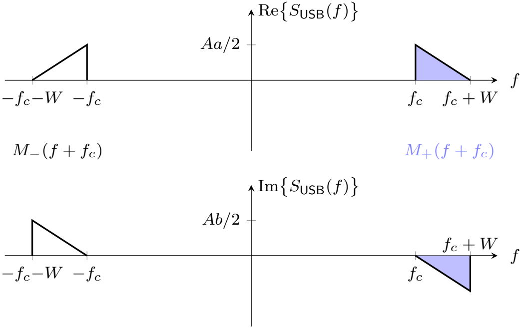



$$m(t) \Longleftrightarrow M(f), \quad m_+(t) = m(t) + j\hat{m}(t) \Longleftrightarrow M_+(f) = \begin{cases} 2M(f), & f > 0, \\ M(0), & f = 0, \\ 0, & f < 0 \end{cases}$$

Pre-envelope removes the negative frequency components.

# Upper Sideband SSB Modulated Signal: II

Move positive frequency part $m_+(t)$ to $f_c$




Move negative frequency part $m_-(t)$ to $-f_c$

$$s(t) = \frac{A}{4} m_+(t) \exp(j2\pi f_c t) + \frac{A}{4} m_-(t) \exp(-j2\pi f_c t) \quad \text{(is it real?)}$$

$$S(f) = \frac{A}{4} M_+(f - f_c) + \frac{A}{4} M_-(f + f_c)$$

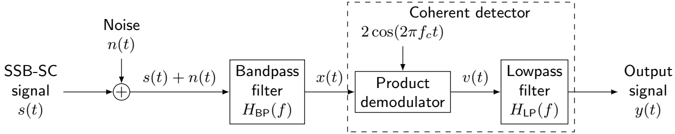
# Upper Sideband SSB Modulation

$$
\begin{aligned}
s(t) &= \frac{A}{4} m_+(t) \exp(j2\pi f_c t) + \frac{A}{4} m_-(t) \exp(-j2\pi f_c t) \\
&= \frac{A}{4} \big(m(t) + j\hat{m}(t)\big) \exp(j2\pi f_c t) + \frac{A}{4} \big(m(t) - j\hat{m}(t)\big) \exp(-j2\pi f_c t) \\
&= \frac{A}{2} m(t) \cos(2\pi f_c t) - \frac{A}{2} \hat{m}(t) \sin(2\pi f_c t) \quad \text{(it is real!)}
\end{aligned}
$$

- I component: the message $m(t)$
- Q component: its Hilbert transform $\hat{m}(t)$
- $m(t)$ and $\hat{m}(t)$ have the same power $P$ (why?)
- Transmission (signal) power: $P_T = A^2 P/4$, $P = \mathbb{E}\{m^2(t)\}$

# Noise in Upper Sideband SSB

- Received signal $x(t) = s(t) + n(t)$, apply a band-pass filter on the upper sideband
- Still denote by $n_I(t)$ the upper-sideband in-phase noise (different from the double-sideband noise in DSB)

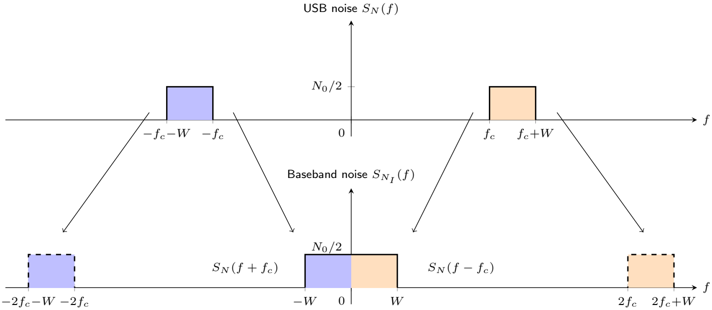



- Demodulation output (after low-pass filtering)

$$y(t) = \left(\frac{A}{2}m(t) + n_I(t)\right)$$

# Noise Power

Noise power of $n_I(t)$ = power of band-pass noise $= N_0 W$ (halved compared to DSB)




$$S_{N_I}(f) = \begin{cases} S_N(f+f_c) + S_N(f-f_c) = N_0/2, & |f| \leq W \\ 0, & \text{otherwise} \end{cases}$$

$$P_N = \int_{-W}^{W} \frac{N_0}{2} df = N_0 W \quad \text{(halved compared to DSB)}$$

# Output SNR

- Signal power at the receiver output:

$$P_S = \mathbb{E}\{(A/2)^2 m^2(t)\} = \frac{A^2}{4}\mathbb{E}\{m^2(t)\} = \frac{A^2 P}{4} \quad \text{(1/4 of DSB)}$$

- SNR at the receiver output:

$$\mathsf{SNR}_{\mathsf{SSB}} = \frac{P_S}{P_N} = \frac{A^2 P}{4N_0 W}$$

- Transmit power: $P_T = A^2 P/4$ (halved compared to DSB)

- Conclusion: SSB has the same SNR performance as DSB-SC and baseband systems, but only requires half the bandwidth!

# DSB, SSB, and standard AM

- DSB:

$$s_{\mathsf{DSB}}(t) = m(t)A\cos(2\pi f_c t)$$

- SSB:

$$s_{\mathsf{SSB}}(t) = \frac{A}{2}m(t)\cos(2\pi f_c t) - \frac{A}{2}\hat{m}(t)\sin(2\pi f_c t)$$

- Standard AM:

$$s_{\mathsf{AM}}(t) = \big(A + m(t)\big)\cos(2\pi f_c t)$$

To ensure non-coherent demodulation,

$$A \geq m_p = \max|m(t)| \quad \text{or} \quad \text{modulation index } \mu = \frac{m_p}{A} \leq 1$$

Coherent detection is sometimes referred to as synchronous recovery.

# Standard AM: Coherent Detection

- Pre-detection signal:

$$\begin{aligned}
x(t) &= \big(A + m(t)\big)\cos(2\pi f_c t) + n(t) \\
&= \big(A + m(t)\big)\cos(2\pi f_c t) + n_I(t)\cos(2\pi f_c t) - n_Q(t)\sin(2\pi f_c t) \\
&= \big(A + m(t) + n_I(t)\big)\cos(2\pi f_c t) - n_Q(t)\sin(2\pi f_c t)
\end{aligned}$$

- Multiply with $2\cos(2\pi f_c t)$:

$$y(t) = \big(A + m(t) + n_I(t)\big)\big(1 + \cos(4\pi f_c t)\big) - n_Q(t)\sin(4\pi f_c t)$$

- LPF

$$\tilde{y} = A + m(t) + n_I(t)$$

# Output SNR

- Signal power at the receiver output:

$$P_S = \mathbb{E}\{m^2(t)\} = P$$

- Noise power:

$$P_N = 2N_0W \quad \text{(same as DSB)}$$

- SNR at the receiver output:

$$\text{SNR}_{\text{AM}} = \frac{P}{2N_0W}$$

- Transmitted power

$$P_T = \frac{A^2}{2} + \frac{P}{2} = \frac{A^2 + P}{2} \quad \text{(with } A^2 \text{ term!)}$$

- SNR of a baseband system with the same transmitted power:

$$\mathsf{SNR}_{\mathsf{baseband}} = \frac{P_T}{N_0 W} = \frac{A^2 + P}{2N_0 W}$$

- Thus

$$\mathsf{SNR}_{\mathsf{AM}} = \frac{P}{2N_0 W} = \frac{P}{A^2 + P}\frac{A^2 + P}{2N_0 W} = \frac{P}{A^2 + P}\mathsf{SNR}_{\mathsf{baseband}}$$

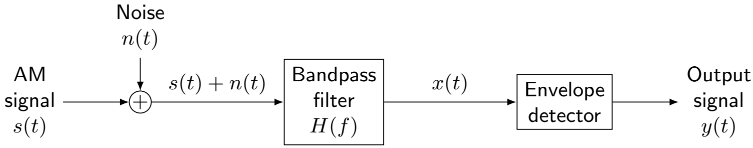
- Note

$$\frac{P}{A^2 + P} < 1$$

- Conclusion: performance of standard AM with coherent detection is worse than a baseband system.

# Non-coherent Receiver

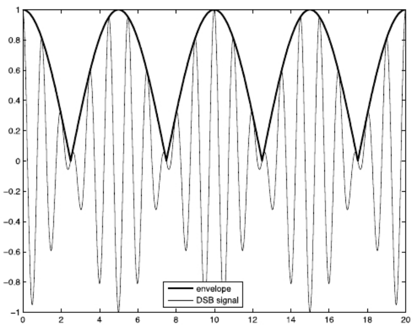
- The envelope (or magnitude of complex envelope) does not depend on carrier phase
- Extracting the envelope of a passband signal does not require carrier sync
- Can we recover the message from the envelope?



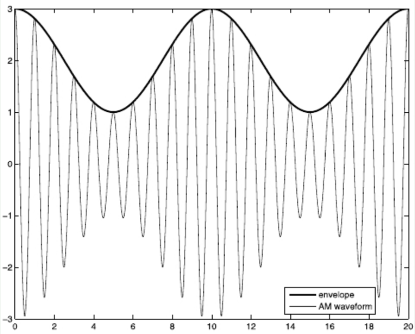

Model of AM envelope detector

# What does the envelope tell us?

## Example: sinusoidal message waveform




DSB modulated signal




DSB signal + strong carrier component

- Envelope = message magnitude
- Envelope detection loses message sign

- Envelope = message + DC
- Envelope detector + DC block recovers message info

# Threshold Effect

- Small noise case:
  - Almost same performance as coherent detection (assume $\mu \leq 1$ or $m_P \leq A$)
- Large noise case:
  - Information is lost!
  - Threshold effect: below some carrier-to-noise ratio level (very low $A$), performance of envelope detector deteriorates very rapidly (not the case in coherent detection)

# Summary

| (De-) modulation format | Output SNR | Transmitted power | Baseband reference SNR | Output SNR / reference SNR |
|---|---|---|---|---|
| AM coherent detection | $\frac{P}{2N_0W}$ | $\frac{A^2+P}{2}$ | $\frac{A^2+P}{2N_0W}$ | $\frac{P}{A^2+P} < 1$ |
| DSB-SC coherent detection | $\frac{A^2P}{2N_0W}$ | $\frac{A^2P}{2}$ | $\frac{A^2P}{2N_0W}$ | 1 |
| SSB coherent detection | $\frac{A^2P}{4N_0W}$ | $\frac{A^2P}{4}$ | $\frac{A^2P}{4N_0W}$ | 1 |
| AM envelope detection (small noise) | $\frac{P}{2N_0W}$ | $\frac{A^2+P}{2}$ | $\frac{A^2+P}{2N_0W}$ | $\frac{P}{A^2+P} < 1$ |
| AM envelope detection (large noise) | Poor | $\frac{A^2+P}{2}$ | $\frac{A^2+P}{2N_0W}$ | Poor |

- $A$: carrier amplitude
- $P$: power of message signal
- $N_0$: single-sided PSD of noise
- $W$: message bandwidth

# Note

# Lecture 9: Frequency Modulation (FM)

Dr. Geoffrey Li

Department of Electrical & Electronic Engineering
Imperial College London

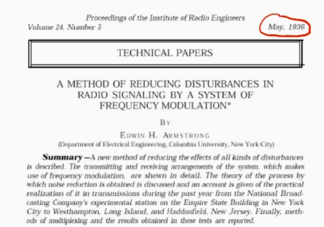
# Outline

- FM system: model, bandwidth, etc.
- Demodulation and output SNR
- References
  - [Haykin] Chapter 6

# Invention of FM

Edwin Howard Armstrong:

- Invented wideband FM
- Patented the regenerative circuit in 1914
- Presented *A Method of Reducing Disturbances in Radio Signalling by a System of Frequency Modulation* on 6 Nov. 1935
- This is the first paper to describer FM radio before the New York section of the Institute of Radio Engineers (now IEEE)
- Committed suicide in 1954



*Proceedings of the Institute of Radio Engineers*
*Volume 24, Number 5* — *May, 1936*

TECHNICAL PAPERS

A METHOD OF REDUCING DISTURBANCES IN RADIO SIGNALING BY A SYSTEM OF FREQUENCY MODULATION*

BY

EDWIN H. ARMSTRONG
(Department of Electrical Engineering, Columbia University, New York City)

**Summary**—*A new method of reducing the effects of all kinds of disturbances is described. The transmitting and receiving arrangements of the system, which makes use of frequency modulation, are shown in detail. The theory of the process by which noise reduction is obtained is discussed and an account is given of the practical realization of it in transmissions during the past year from the National Broadcasting Company's experimental station on the Empire State Building in New York City to Westhampton, Long Island, and Haddonfield, New Jersey. Finally, methods of multiplexing and the results obtained in these tests are reported.*

# FM vs AM

Fundamental difference:

- AM: information contained in the signal amplitude
  - additive noise: corrupting the modulated signal directly
- FM: information contained in the signal frequency
  - additive noise: affecting the signal by changing the frequency of the modulated signal
  - consequently, affected less by noise than AM

# Frequency and Phase

- A carrier waveform

$$s(t) = A\cos \underbrace{\theta_i(t)}_{\text{instantaneous phase angle}}$$

- Constant frequency

$$s(t) = A\cos(2\pi f t + \theta) \Rightarrow \theta_i(t) = 2\pi f t + \theta$$

$$\frac{d\theta_i(t)}{dt} = 2\pi f \Rightarrow f = \frac{1}{2\pi}\frac{d\theta_i(t)}{dt}$$

- Generalization: instantaneous frequency

$$f_i(t) = \frac{1}{2\pi}\frac{d\theta_i(t)}{dt}, \quad \theta_i(t) = 2\pi\int_0^t f_i(\tau)d\tau + \theta_i(0)$$

# FM and PM

- Instantaneous frequency: varied linearly with message

$$f_i(t) = f_c + \underbrace{k_f}_{\text{frequency sensitivity of the modulator}} m(t)$$

- Instantaneous phase angle

$$\theta_i(t) = 2\pi \int_0^t f_i(\tau)d\tau + \theta_i(0) = 2\pi f_c t + 2\pi k_f \int_0^t m(\tau)d\tau, \quad (\theta_i(0) = 0)$$

- FM signal:

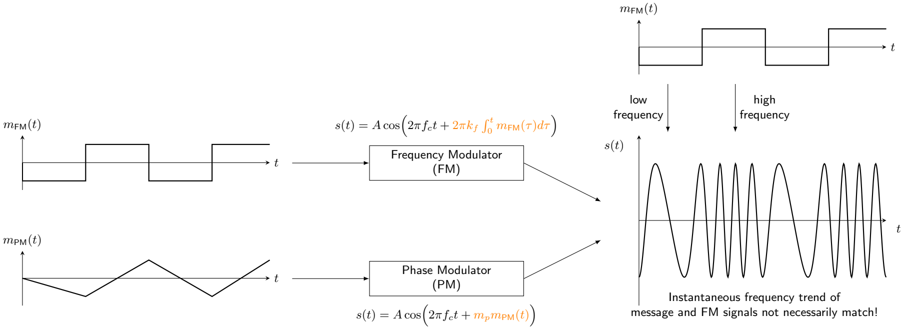
$$s(t) = A\cos\Big(2\pi f_c t + 2\pi k_f \int_0^t m(\tau)d\tau\Big) + \theta_i(0)$$

- Phase modulation (PM) signal:

$$s(t) = A\cos\big(2\pi f_c t + k_p m(t)\big)$$

- FM or PM signals:
  - constant envelope
  - non-linear function of the message signal, $m(t)$

# PM-FM Equivalence




- FM signal $=$ PM signal with the modulating signal $\int_0^t m(\tau)d\tau$
- Similar properties for PM and FM
- Focusing on FM

# FM Basics

- $m_p = \max|m(t)|$: peak message amplitude

$$f_c - k_f m_p \le f_i(t) \le f_c + k_f m_p$$

- Frequency deviation: deviation of $f_i(t)$ from the carrier frequency, $f_c$,

$$\Delta f = k_f m_p$$

- Deviation ratio/Modulation index:

$$\beta = \frac{\Delta f}{W}$$

  - $W$: message bandwidth

**Sinusoidal message $m(t) = A_m \cos(2\pi f_m t)$**

# Bandwidth of FM

- Deviation ratio/Modulation index:

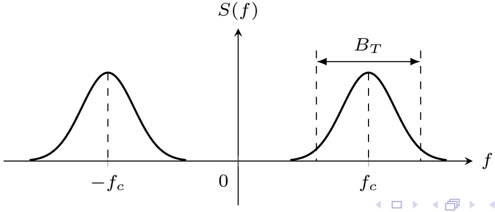
$$\beta = \frac{\Delta f}{W}$$

  - $W$: message bandwidth
  - $\beta$ small: narrowband FM
  - $\beta$ large: wideband FM

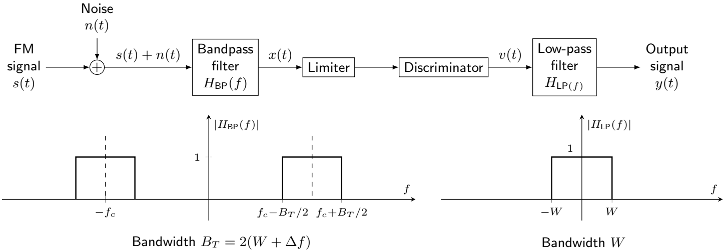
- Carson's rule of thumb: Transmission bandwidth of FM

$$B_T = 2W(\beta + 1) = 2(\Delta f + W)$$

  - $\beta \ll 1 \Rightarrow B_T \approx 2W$ (as in AM)
  - $\beta \gg 1 \Rightarrow B_T \approx 2\Delta f$



# FM Receiver




- Bandpass filter: removes signals outside bandwidth of $f_c \pm B_T/2$
  - predetection noise at the receiver is bandpass with a bandwidth of $B_T$
- FM signal with a *constant envelope*
  - use a limiter to remove any amplitude variations
- Discriminator: a device with instantaneous amplitude proportional to instantaneous frequency
  - recovering the message signal
- Final baseband low-pass filter: a bandwidth of $W$
  - removing out-of-band noise

# Linear Argument at High SNR

- FM: nonlinear modulation and demodulation, no superposition principle
- For *high SNR*, noise and message signals are approximately independent of each other:

$$\text{Output} \approx \text{Message} + \text{Noise (i.e., no other nonlinear terms)}$$
$$y(t) \approx k_f m(t) + n_0(t)$$

(will show)

# Phase Noise at High SNR

$$\begin{aligned} x(t) &= A\cos\big(2\pi f_c t + \phi(t)\big) + n_I(t)\cos(2\pi f_c t) + n_Q(t)\sin(2\pi f_c t) \\ &= A\cos\big(2\pi f_c t + \phi(t)\big) + r(t)\cos\big(2\pi f_c t + \psi(t)\big) \end{aligned}$$

$$\phi(t) = 2\pi k_f \int_0^t m(\tau)d\tau$$

- Instantaneous phase of the resultant phasor:

$$\theta(t) = \phi(t) + \arctan\left(\frac{r(t)\sin\big(\psi(t)-\phi(t)\big)}{A + r(t)\cos\big(\psi(t)-\phi(t)\big)}\right) \quad \big(\arctan(x) \approx x \text{ if } |x| \ll 1\big)$$

$$\approx \phi(t) + \frac{r(t)}{A}\sin\big(\psi(t)-\phi(t)\big)$$

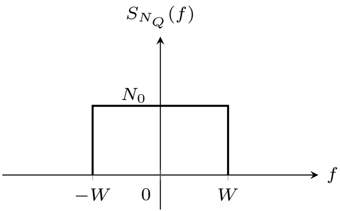
# Phase Noise at High SNR

Discriminator output:

$$v(t) = \frac{1}{2\pi}\frac{d\theta(t)}{dt} \approx k_f m(t) + n_d(t)$$

where the *additive* noise term is

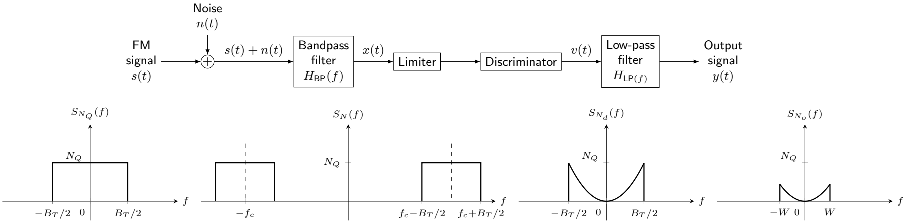
$$\begin{aligned} n_d(t) &= \frac{1}{2\pi A}\frac{d}{dt}r(t)\sin\big(\psi(t)-\phi(t)\big) \quad \text{(HW1.3)} \\ &\approx \frac{1}{2\pi A}\frac{d}{dt}r(t)\sin\big(\psi(t)\big) \\ &= \frac{1}{2\pi A}\frac{d}{dt}n_Q(t) \end{aligned}$$




with PSD

$$S_{N_0}(f) = \left(\frac{1}{2\pi A^2}\right)(2\pi f)^2 S_{N_Q}(f) = \frac{f^2}{A^2}N_0, \quad |f| \le W \quad \text{(HW1.8)}$$

# Noise PSD




- $S_{N_Q}(f)$: PSD of $n_Q(t)$ of narrowband noise $n(t)$
- $S_{N_d}(f)$: PSD of $n_d(t)$ at the discriminator output
- $S_{N_o}(f)$: PSD of $n_o(t)$ at the receiver output

# Noise Power

- Average noise power at the receiver output

$$P_N = \int_{-W}^{W} S_{N_0}(f)df = \int_{-W}^{W} \frac{f^2}{A^2} N_0 df = \frac{2N_0W^3}{3A^2}$$

- Average noise power at the output of an FM receiver

$$\propto \frac{1}{\text{carrier power } A^2}$$

- $A \uparrow \Rightarrow$ noise $\downarrow$, called the noise-quieting effect

# Output SNR

- $P_S = k_f^2 P$, $P_N = 2N_0 W^3/3A^2$

$$\text{SNR}_{\text{FM}} = \frac{P_S}{P_N} = \frac{3A^2 k_f^2 P}{2N_0 W^3}$$

- For baseband transmission,

$$\text{SNR}_{\text{baseband}} = \frac{P_T}{N_0 W} = \frac{A^2/2}{N_0 W} = \frac{A^2}{2N_0 W}$$

- $P_T = A^2/2$, $\beta = k_f m_p / W$

$$\text{SNR}_{\text{FM}} = \frac{3k_f^2 P}{W^2}\text{SNR}_{\text{baseband}} = 3\beta^2 \frac{P}{m_p^2}\text{SNR}_{\text{baseband}}$$
$$\propto \beta^2 \text{SNR}_{\text{baseband}} \text{(could be much higher than AM)}$$

- Valid for large carrier power
- $\text{SNR}_{\text{FM}}$: quadratically increasing with $\beta$

# Threshold Effect

- A more pronounced threshold effect than AM envelope detector
- Threshold point at

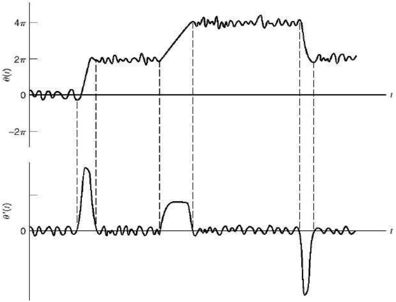
$$\text{Carrier-to-noise ratio:} \quad \rho = \frac{A^2}{2N_0 B_T} \approx 10, \quad B_T = 2W(\beta+1)$$

- FM receiver breaks (i.e., significantly deteriorates) at $\rho < 10$
- Analyzed by *S. O. Rice* (very complicated!), the noise in FM receiver is called "click noise" or "Rice noise"

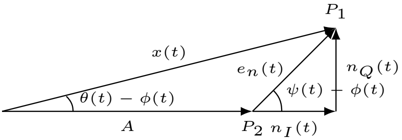
# Qualitative Discussion

As noise changes randomly, point $P_1$ wanders around $P_2$

- High SNR: change of angle is small
- Low SNR: $P_1$ occasionally sweeps around origin, resulting in changes of $2\pi$ in a short time

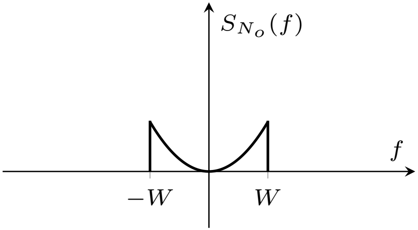



Phase noise

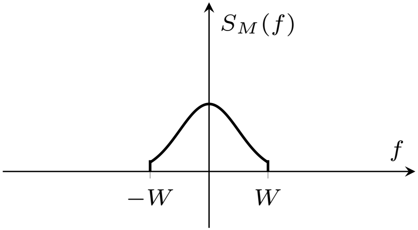



Phasor diagram of the FM carrier and noise signals

# Improve Output SNR

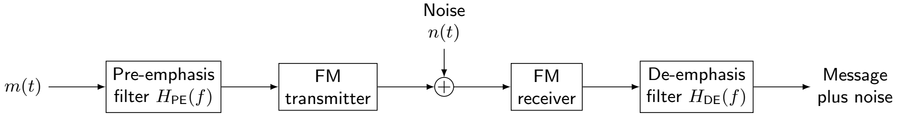



Noise PSD




Message PSD

- Noise PSD at detector output $\propto$ square of frequency
- Message PSD typically decays towards the ends of its band




- $H_{\mathrm{PE}}(f)$: artificially emphasizes high frequency components of the message prior to modulation (before noise is introduced)
- $H_{\mathrm{DE}}(f)$: de-emphasizes high frequency components at the receiver, and restore the original PSD of the message
- In theory, $H_{\mathrm{PE}}(f) \propto f$, $H_{\mathrm{DE}}(f) \propto 1/f$
- This can improve output SNR by around 13 dB

- Assumptions:
  - single-tone modulation $m(t) = A_m \cos(2\pi f_m t)$
  - message bandwidth $W = f_m$
  - for AM system, modulation index $\mu = m_p/A = A_m/A = 1$, $m_p = \max|m(t)| = A_m$
  - for FM system, modulation index $\beta = \Delta f/W = 5$, $\Delta f = k_f m_p = k_f A_m$ (used in commercial FM transmission with $\Delta f = 75$ kHz and $W = 15$ kHz)

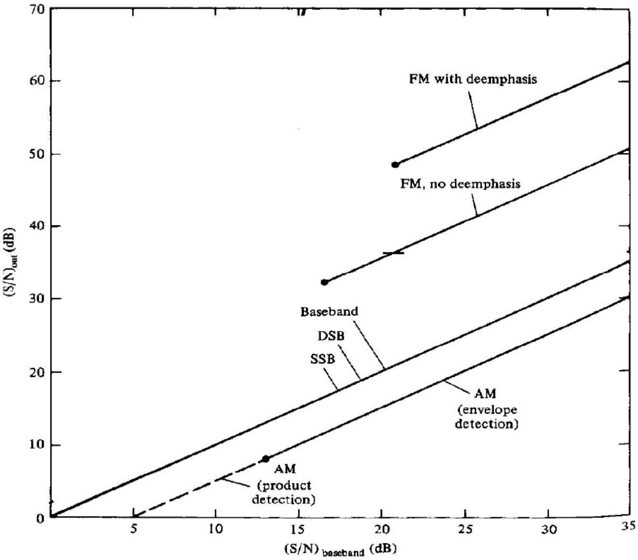
- SNR expressions for various modulation schemes

$$\mathsf{SNR}_{\mathsf{DSB\text{-}SC}} = \mathsf{SNR}_{\mathsf{baseband}} = \mathsf{SNR}_{\mathsf{SSB}}$$

$$\mathsf{SNR}_{\mathsf{AM}} = \frac{P}{A^2 + P}\mathsf{SNR}_{\mathsf{baseband}} = \frac{\mu^2}{2 + \mu^2}\mathsf{SNR}_{\mathsf{baseband}} \leq \frac{1}{3}\mathsf{SNR}_{\mathsf{baseband}}$$

$$\mathsf{SNR}_{\mathsf{FM}} = \frac{3\beta^2}{2}\mathsf{SNR}_{\mathsf{baseband}} = \underbrace{\frac{75}{2}}_{15.7\mathrm{dB}}\mathsf{SNR}_{\mathsf{baseband}} \quad \text{(without pre/de-emphasis)}$$

# Performance of Analog Systems



# Conclusions

- (Full) AM:
  - SNR: 4.8 dB worse than a baseband system
  - transmission bandwidth: $B_T = 2W$
- DSB:
  - SNR: identical to a baseband system
  - transmission bandwidth: $B_T = 2W$
- SSB:
  - SNR: again identical
  - transmission bandwidth: $B_T = W$
- FM:
  - SNR: 15.7 dB better than a baseband system
  - transmission bandwidth: $B_T = 2(\beta + 1)W = 12W$
  - with pre- and de-emphasis, SNR is increased by 13 dB with the same bandwidth

# Note

# Lecture 10: Digital Representation of Signals

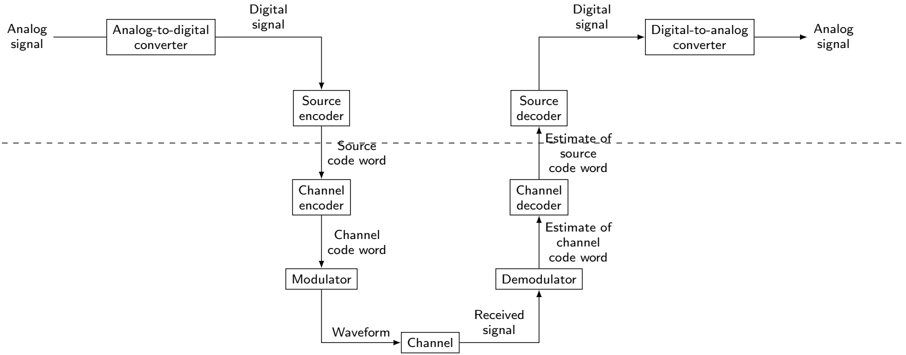
Dr. Geoffrey Li

Department of Electrical & Electronic Engineering
Imperial College London

# Outline

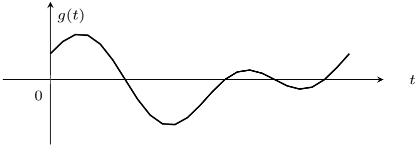
- Digital communication
- Quantization (A/D) and noise
- Pulse-Coded Modulation (PCM)
- Companding and expanding
- References
  - [Haykin] Chapter 7

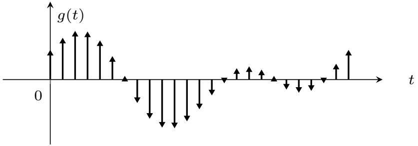
# Digital Communication



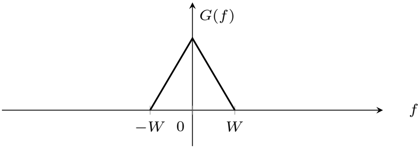
# Why Digital?

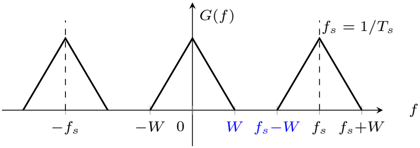
- More immune to channel noise by using channel coding
- Repeaters along the transmission path: error-correction and retransmission
- Representing different analog sources using digital signals, a uniform format
- Easily processing using microprocessors and VLSI (e.g., digital signal processors, FPGA)
- More and more things are digital . . .




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




No overlapping in $G(f)$ if $W \leq f_s - W$, namely $f_s \geq 2W$

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

## Sampling theorem

For distortionless recovery, sampling rate $f_s \geq 2W$ for a (real) signal with bandwidth $W$. The Nyquist frequency is

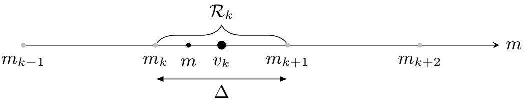
$$f_{\mathrm{N}} = 2W$$

- Quantization: transforming the sample amplitude into a discrete amplitude taken from a finite set of possible amplitudes
- The more levels, the better approximation
- No need to transmit exact values
- Memoryless and instantaneous quantization: quantization at time $t$ is independent of other samples
- Quantizers: uniform or nonuniform

Continuous sample $m$ $\longrightarrow$ $\boxed{\text{Quantizer } g(\cdot)}$ $\longrightarrow$ Discrete sample $v_k$

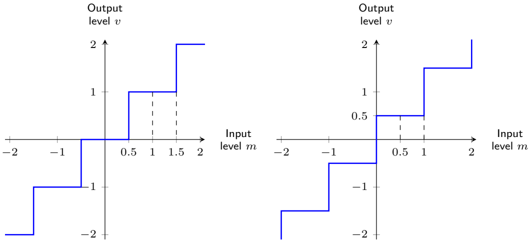
$$v_k = g(m)$$

# Quantization: II



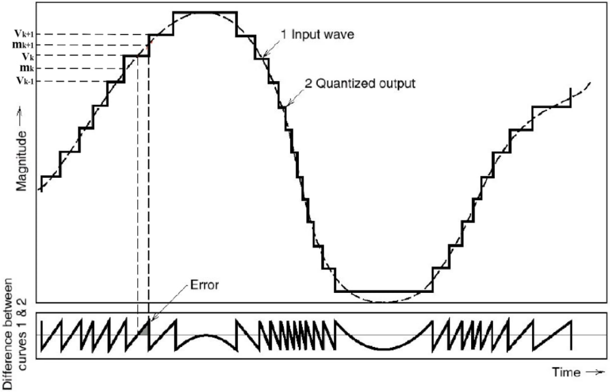

$$v_k = \frac{m_k + m_{k+1}}{2}, \quad \Delta = m_{k+1} - m_k$$

- Decision levels: $m_1, \ldots, m_L$
- Decision region: $\mathcal{R}_k = (m_k, m_{k+1}]$, $k = 1, \ldots, L$
- Reconstruction levels: $v_k$, $k = 1, \ldots, L$
- Quantizer output $v_k$ represents decision region $\mathcal{R}_k$
- Mapping $v = g(m)$ is the quantizer characteristic



# Quantization Noise

- Error between the input and the output signals



# Variance of Quantization Noise

- $\Delta$ : gap between quantizing levels (of a uniform quantizer)
- $q$ : quantization error = random variable within the range

$$-\frac{\Delta}{2} \leq q \leq \frac{\Delta}{2}$$

- For small $\Delta$, $q$ is uniform:

$$f_Q(q) = \begin{cases} \frac{1}{\Delta}, & -\frac{\Delta}{2} \leq q \leq \frac{\Delta}{2} \\ 0, & \text{otherwise} \end{cases}$$

- Quantization noise variance

$$P_N = \mathbb{E}\{q^2\} = \int_{-\infty}^{\infty} q^2 f_Q(q) dq = \frac{1}{\Delta} \int_{-\Delta/2}^{\Delta/2} q^2 dq = \frac{1}{\Delta} \left( \frac{\Delta^3}{24} - \frac{(-\Delta)^3}{24} \right) = \frac{\Delta^2}{12}$$

# Quantization Gap $\Delta$

For $n$-bit quantization:

- Maximum number of quantizing levels: $L = 2^n$
- Maximum peak-to-peak dynamic range: $2^n \Delta$
- Power of the message signal:

$$P = \mathbb{E}\{m^2(t)\} \xlongequal{\text{periodic}} \frac{1}{T}\int_{-T/2}^{T/2} |m(t)|^2 dt$$

- Maximum magnitude: $m_p = \max|m(t)|$
- Full load quantizer:

$$2m_p = 2^n \Delta \quad \text{or} \quad \Delta = 2^{-(n-1)} m_p$$

# SNR

- SNR at the quantizer output:

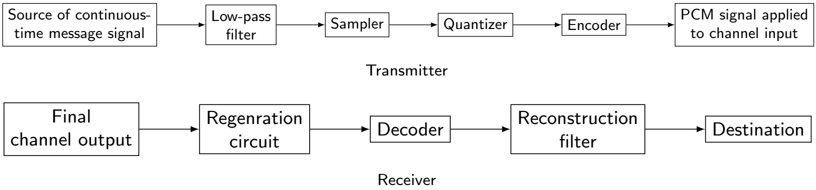
$$\text{SNR}_\text{o} = \underbrace{\frac{P_S}{P_N} = \frac{P}{\Delta^2/12}}_{\text{by quantizer principle}} = \frac{3P}{m_p^2} 2^{2n}$$

$$\text{SNR}_\text{o}(\text{dB}) = 10 \log_{10}(2^{2n}) + 10 \log_{10}\left(\frac{3P}{m_p^2}\right) = 6n + 10 \log_{10}\left(\frac{3P}{m_p^2}\right)$$

- An extra bit in the encoder $\Leftrightarrow$ 6 dB more to the output SNR
- Recognize the tradeoff between SNR and $n$ (i.e., rate, or bandwidth)

# Example

# Pulse-Coded Modulation (PCM)



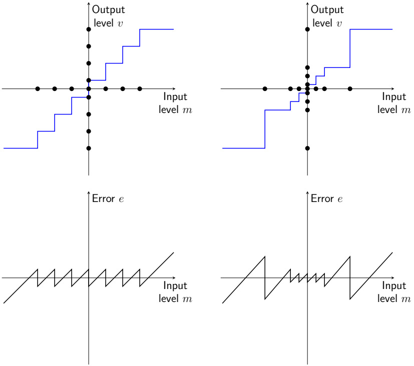

- Low-pass filter is applied to prevent aliasing
- Sample the message signal above Nyquist rate
- Quantize each sample
- Encode discrete amplitudes into a binary codeword
- PCM: not modulation in usual sense; type of Analog-to-Digital Converter

# Problem With Uniform Quantization

- SNR: adversely affected by peak-to-average power ratio
- More often with small signals than large signals
- More quantization levels for smaller amplitudes

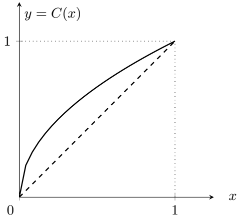
# Solution: Nonuniform Quantization

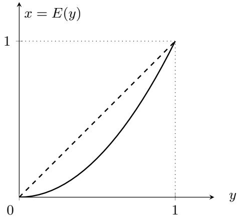
Nonuniform quantization: quantization levels of variable spacing, denser at small signal amplitudes, broader at large amplitudes.




Uniform    Nonuniform

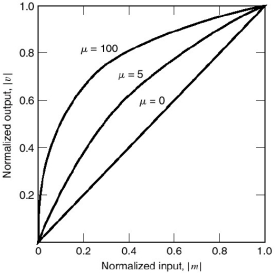
- A practical (and equivalent) solution to nonuniform quantization:
    - compress the signal
    - quantize it (using a uniform quantizer)
    - transmit it
    - expand it




- Companding/Expanding: pre-emphasising/de-emphasising as in FM
- Ideal compression and expansion: exactly inverse of each other
- Exact SNR gain: depending on the exact form of the compression

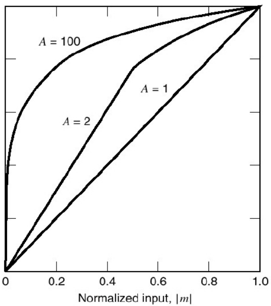
# Compander Standards: $\mu$-Law




- $\mu$-Law (North America and Japan, typical $\mu = 255$)

$$y = F(x) = \mathrm{sgn}(x)\frac{\log\big(1+\mu|x|\big)}{\log(1+\mu)}, \quad |x| < 1$$

$$F^{-1}(y) = \mathrm{sgn}(y)\frac{(1+\mu)^{|y|}-1}{\mu}$$




- A-Law (in most countries of the world; typical A = 87.6)

$$y = F(x) = \begin{cases} \operatorname{sgn}(x)\frac{A|x|}{1+\log(A)}, & |x| < \frac{1}{A} \\ \operatorname{sgn}(x)\frac{1+\log(A|x|)}{1+\log(A)}, & \frac{1}{A} \le |x| \le 1 \end{cases}$$

$$x = F^{-1}(y) = \begin{cases} \operatorname{sgn}(y)\frac{|y|(1+\log(A))}{A}, & |y| < \frac{1}{1+\log(A)} \\ \operatorname{sgn}(y)\frac{\exp(|y|(1+\log(A))-1)}{A}, & \frac{1}{1+\log(A)} \le |y| \le 1 \end{cases}$$

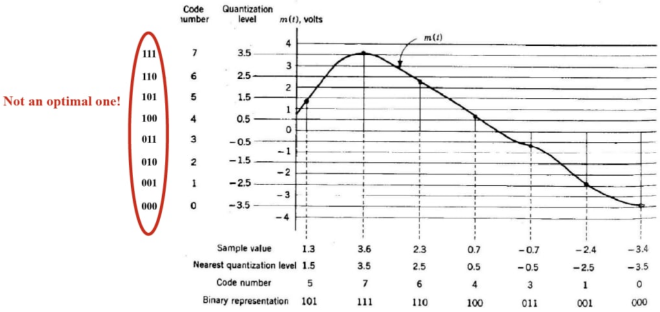
# Applications of PCM & Variants

- Speech:
  - PCM: Voice signal is sampled at 8 kHz, quantized into 256 levels (8 bits). Thus, a telephone PCM signal requires 64 kbps (need to reduce bandwidth requirements).
  - DPCM (differential PCM): quantize the difference between consecutive samples; can save 8 to 16 kbps. ADPCM (Adaptive DPCM) can go further down to 32 kbps.
  - Delta modulation: 1-bit DPCM with oversampling; has even lower symbol rate (e.g., 24 kbps).

- Audio CD: 16-bit PCM at 44.1 kHz sampling rate.

- MPEG audio coding: 16-bit PCM at 48 kHz sampling rate compressed to a rate as low as 16 kbps.

# PCM Process


Not an optimal one!
111
110
101
100
011
010
001
000
Code number
Quantization level
m(t), volts
m(t)
7
6
5
4
3
2
1
0
3.5
2.5
1.5
0.5
−0.5
−1.5
−2.5
−3.5
4
3
2
1
0
−1
−2
−3
−4
Sample value 1.3 3.6 2.3 0.7 −0.7 −2.4 −3.4
Nearest quantization level 1.5 3.5 2.5 0.5 −0.5 −2.5 −3.5
Code number 5 7 6 4 3 1 0
Binary representation 101 111 110 100 011 001 000

# Summary

Digitization of signals requires

- Sampling: sampled at the Nyquist frequency $2W$
- Quantization: link between analog waveforms and digital representation
  - SNR (under high-resolution assumption)

$$\mathrm{SNR_o(dB)} = 6n + 10\log_{10}\left(\frac{3P}{m_p^2}\right)$$

  - Companding to improve SNR

PCM: a common method of representing audio signals

- "Pulse coded modulation": a simple source coding technique (i.e, method of digitally representing analog information)
- More advanced source coding (compression) techniques in information theory

# Note

# Lecture 11: Matched Filter

Dr. Geoffrey Li

Department of Electrical & Electronic Engineering
Imperial College London

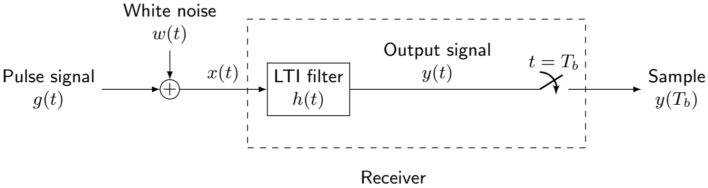
# Outline

- Matched filter
  - ✓ Impulse response
  - ✓ Maximum peak SNR
- Binary baseband communication
  - ✓ Distribution of noise
  - ✓ Decision rule
  - ✓ Error cases and probabilities
- References
  - ✓ [Haykin] Chapter 8

# Analog and Digital Communications

- Analog communication systems: reproducing transmitted waveform accurately
  - ✓ Signal-to-noise ratio (SNR) to assess the quality of the system
- Digital communication systems: recovering the transmitted symbol correctly
  - ✓ Probability of error or bit-error rate (BER) at the receiver to assess the quality of the system

# Matched Filter: I




$$x(t) = g(t) + w(t), \quad 0 \le t \le T_b$$
$$y(t) = x(t) * h(t)$$

- $w(t)$ : white noise with zero mean and PSD $N_0/2$
- Detecting whether pulse presents with known pulse shaping $g(t)$
- Goal: designing a receive (linear) filter to minimize the effect of noise
  - ✓ Optimizing the receive filter $h(t)$

# Matched Filter: II

- Filter output:

$$y(t) = x(t) * h(t) = g(t) * h(t) + w(t) * h(t) = g_\circ(t) + n(t)$$
$$g_\circ(t) = g(t) * h(t) = \int_{-\infty}^{\infty} g(\tau)h(t-\tau)d\tau$$
$$n(t) = w(t) * h(t)$$

- We want instantaneous power of signal component $g_\circ(t)$ at time $t = T_b$ as large as possible compared to noise component $n(t)$

- Maximize peak signal-to-noise ratio

$$\eta = \frac{|g_\circ(T_b)|^2}{\mathbb{E}\{n^2(T_b)\}} = \frac{\text{instantaneous power}}{\text{average power}}$$

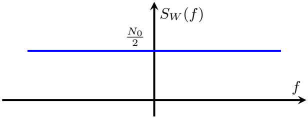
# Matched Filter Derivation: I

- Noise

$$n(t) = w(t) * h(t)$$

$$S_N(f) = \underbrace{S_W(f)}_{\text{AWGN}} \underbrace{S_H(f)}_{\text{receive filter}} = \frac{N_0}{2}|H(f)|^2$$

$$\mathbb{E}\{n^2(t)\} = \int_{-\infty}^{\infty} S_N(f)df = \frac{N_0}{2}\int_{-\infty}^{\infty}|H(f)|^2 df$$




- Signal

$$g_o(t) = g(t) * h(t) = \int_{-\infty}^{\infty} H(f)G(f)e^{j2\pi f t}df$$

$$G_o(f) = H(f)G(f)$$

$$|g_o(T_b)|^2 = \left|\int_{-\infty}^{\infty} H(f)G(f)e^{j2\pi f T_b}df\right|^2$$

- Find $h(t)$ to maximize peak SNR

$$\eta = \frac{\left| \int_{-\infty}^{\infty} \overbrace{H(f)}^{\phi_1(x)} \overbrace{G(f)e^{j2\pi f T_b}}^{\phi_2^*(x)} df \right|^2}{\frac{N_0}{2} \int_{-\infty}^{\infty} |H(f)|^2 df}$$

- Schwartz's inequality: Consider two energy signals $\phi_1(x)$ and $\phi_2(x)$,

$$\left| \int_{-\infty}^{\infty} \phi_1(x)\phi_2(x)^* dx \right|^2 \le \int_{-\infty}^{\infty} |\phi_1(x)|^2 dx \int_{-\infty}^{\infty} |\phi_2(x)|^2 dx,$$

where equality holds if and only if $\phi_1(x) = k\phi_2(x)$ for an arbitrary constant $k$.

# Matched Filter Derivation: III

Let $\phi_1(f) = H(f)$ and $\phi_2(f) = G^*(f)e^{-j2\pi f T_b}$,

$$\left|\int_{-\infty}^{\infty} H(f)G(f)e^{j2\pi f T_b} df\right|^2 \leq \int_{-\infty}^{\infty} |H(f)|^2 df \int_{-\infty}^{\infty} |G(f)|^2 df$$

$$\eta = \frac{\left|\int_{-\infty}^{\infty} H(f)G(f)e^{j2\pi f T_b} df\right|^2}{\frac{N_0}{2}|H(f)|^2} \leq \frac{2}{N_0}\int_{-\infty}^{\infty} |G(f)|^2 df,$$

where $\eta_{\max} = \frac{2}{N_0}\int_{-\infty}^{\infty} |G(f)|^2 df$ occurs if $H_{\text{opt}}(f) = kG^*(f)e^{-j2\pi f T_b}$.

$$G^*(f) \Leftrightarrow g^*(-t)$$

$$G^*(f)e^{-j2\pi f T_b} \Leftrightarrow g^*\big(-(t-T_b)\big) = g^*(T_b - t)$$

Hence,

$$h_{\text{opt}}(t) = kg^*(T_b - t) = kg(T_b - t)$$

# Properties of Matched Filters

- Impulse response is

$$h_{\text{opt}}(t) = kg(T_b - t)$$

  ✓ $T_b$ : symbol period
  ✓ $g(t)$ : transmitter pulse shape
  ✓ $k$ : gain
  ✓ scaled, time-reversed and shifted version of $g(t)$
  ✓ duration and shape determined by pulse shape $g(t)$

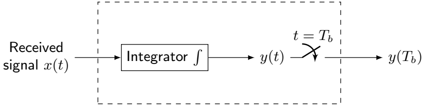
- Maximum peak SNR

$$\eta_{\max} = \frac{2}{N_0} \underbrace{\int_{-\infty}^{\infty} |G(f)|^2 df}_{E} = \frac{2}{N_0} \underbrace{\int_{-\infty}^{\infty} |g(t)|^2 dt}_{E} = \frac{2E}{N_0} = \text{SNR}$$

  ✓ independent of pulse shape $g(t)$
  ✓ proportional to signal energy (energy per bit) $E$
  ✓ inversely proportional to noise power spectral density

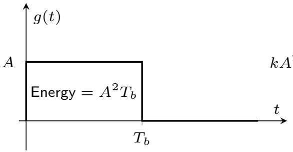
# Matched Filter for Rectangular Pulse

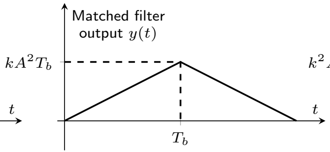
- Matched filter for rectangular pulse shape
  - ✓ matched filter: a rectangular pulse of same duration
- Integrate and dump circuit

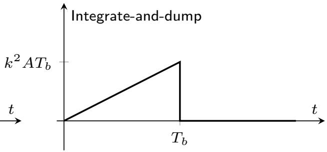


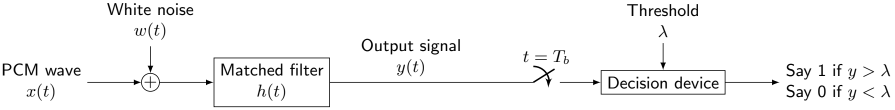

$$h(t) = kg(T_b - t), \quad y(t) = \int_0^{T_b} x(\tau)h(T_b - \tau)d\tau = kA\int_0^{T_b} x(\tau)d\tau$$

$$y(T_b) = kT_bA \text{ if no noise}$$






- For binary PCM with on-off signaling:
  - ✓ $0 \to 0$ and $1 \to A$ with bit duration $T_b$
- Assumptions:
  - ✓ AWGN channel with double-sided noise PSD of $N_0/2$
  - ✓ Rectangular matched filter (set $kT_b = 1$ for simplicity)
- Effect of additive noise: symbol 1 may be mistaken for 0, and vice versa $\Rightarrow$ bit errors
- The probability of a bit error?

# Distribution of Noise

- After the matched filter, the pre-detection signal:

$$Y = y(T_b) = \frac{1}{T_b}\int_0^{T_b} x(t)dt = s + \underbrace{\frac{1}{T_b}\int_0^{T_b} w(t)dt}_{\text{noise } N}$$

  ✓ $s$ : binary-valued function (either $0$ or $A$ volts)
  ✓ $N$ : zero-mean additive white Gaussian noise with variance:

$$\sigma^2 = \frac{N_0}{2T_b}$$

- PDF of Gaussian random variable $N$:

$$p_N(n) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{n^2}{2\sigma^2}\right) = \mathcal{N}(0, \sigma^2)$$

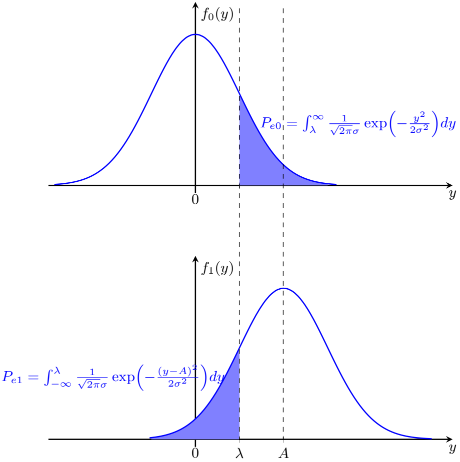
# Decision

- If a symbol 0 was transmitted, $Y = N$
  - ✓ $Y \sim \mathcal{N}(0, \sigma^2)$, $f_0(y) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{y^2}{2\sigma^2}\right)$
- If a symbol 1 was transmitted, $Y = A + N$
  - ✓ $Y \sim \mathcal{N}(A, \sigma^2)$, $f_1(y) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(y-A)^2}{2\sigma^2}\right)$
- Use $\lambda$ as the decision threshold:
  - ✓ choose symbol 0 if $y < \lambda$
  - ✓ choose symbol 1 if $y > \lambda$

# Errors

Two cases of decision error:

- Case I: Symbol 0 was transmitted, but symbol 1 is decided (with probability $P_{e0}$)
- Case II: Symbol 1 was transmitted, but symbol 0 is decided (with probability $P_{e1}$)



# Error Cases

Case I:

$$\text{Prob(error|symbol 0 was transmitted)} \times \text{Prob(symbol 0 was transmitted)}$$
$$P_{\mathsf{I}} = P_{e0} \times p_0$$

- $p_0$ : a priori probability of transmitting a symbol 0
- $P_{e0}$ : conditional probability of error if symbol 0 was transmitted

$$P_{e0} = \int_{\lambda}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp\Big(-\frac{y^2}{2\sigma^2}\Big) dy$$

Case II:

$$\text{Prob(error|symbol 1 was transmitted)} \times \text{Prob(symbol 1 was transmitted)}$$
$$P_{\mathsf{II}} = P_{e1} \times p_1$$

- $p_1$ : a priori probability of transmitting a symbol 1
- $P_{e1}$ : conditional probability of error if symbol 1 was transmitted

$$P_{e1} = \int_{-\infty}^{\lambda} \frac{1}{\sigma\sqrt{2\pi}} \exp\Big(-\frac{(y-A)^2}{2\sigma^2}\Big) dy$$

# Note

# Lecture 12: Quadrature Amplitude Modulation (QAM)

Dr. Geoffrey Li

Department of Electrical & Electronic Engineering
Imperial College London

# Outline

- Binary baseband communication (continued)
  - ✓ Optimal threshold
  - ✓ Average probability of error
- Digital baseband modulation
  - ✓ Waveform, bandwidth, symbol duration, and rate
- Digital passband modulation
  - ✓ QAM modulation and demodulation
- References
  - ✓ [Haykin] Chapter 8, 9

# Optimal Threshold

- Total error probability:

$$P_e(\lambda) = P_{\mathsf{I}} + P_{\mathsf{II}} = p_0 P_{e0} + p_1 P_{e1}$$
$$= (1-p_1)\int_{\lambda}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp\Big(-\frac{y^2}{2\sigma^2}\Big)dy + p_1 \int_{-\infty}^{\lambda} \frac{1}{\sigma\sqrt{2\pi}} \exp\Big(-\frac{(y-A)^2}{2\sigma^2}\Big)dy$$

- Optimal threshold: setting $dP_e(\lambda)/d\lambda = 0$, then

$$-(1-p_1)\frac{1}{\sigma\sqrt{2\pi}} \exp\Big(-\frac{\lambda^2}{2\sigma^2}\Big) + p_1 \frac{1}{\sigma\sqrt{2\pi}} \exp\Big(-\frac{(\lambda-A)^2}{2\sigma^2}\Big) = 0 \Rightarrow$$
$$\lambda_{\text{opt}} = -\frac{\sigma^2}{A}\log\frac{p_1}{1-p_1} + \frac{A}{2}$$

- Equal symbol probability ($p_0 = p_1 = 0.5$):

$$\lambda_{\text{opt}} = \frac{A}{2}, \quad P_{e0} = P_{e1}$$

- Unequal symbol probability: if $p_0 > p_1$, then

$$\lambda_{\text{opt}} > \frac{A}{2}, \quad P_{e0} < P_{e1}$$

# Calculation of $P_e$ for $p_0 = p_1 = 0.5$

- Define a new variable of integration

$$z \triangleq \frac{y}{\sigma} \Rightarrow dy = \sigma dz$$

  ✓ When $y = A/2$, $z = A/2\sigma$
  ✓ When $y = \infty$, $z = \infty$

- Then

$$P_{e0} = \frac{1}{\sigma\sqrt{2\pi}} \int_{A/2\sigma}^{\infty} e^{-z^2/2} \sigma dz = \frac{1}{\sqrt{2\pi}} \int_{A/2\sigma}^{\infty} e^{-z^2/2} dz = Q\Big(\frac{A}{2\sigma}\Big)$$

where

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_{x}^{\infty} \exp\Big(-\frac{t^2}{2}\Big) dt$$

$$P_e = P_{e_0} = P_{e_1} = Q\Big(\frac{A}{2\sigma}\Big)$$

# Probability of Error

$$P_e = Q\left(\frac{A}{2\sigma}\right)$$

- Energy of a pulse is:

$$E = A^2 T_b$$

- We transmit a pulse only half of the time on average $\Rightarrow$ Average energy per bit ($E_b$):

$$E_b = \frac{A^2 T_b}{2}$$

- Noise variance:

$$\sigma^2 = \frac{N_0}{2T_b} \qquad \text{(from slide \textbf{??})}$$

- Probability of error in terms of energy per bit and noise PSD:

$$P_e = Q\left(\sqrt{\frac{E_b}{N_0}}\right)$$

# Example

## Example 1

$$A/\sigma = 7.4(17.4 \text{ dB}) \Rightarrow P_e = 10^{-4}$$

For a transmission rate of $10^5$ bits/sec, there will be an error every 0.1 seconds on the average.

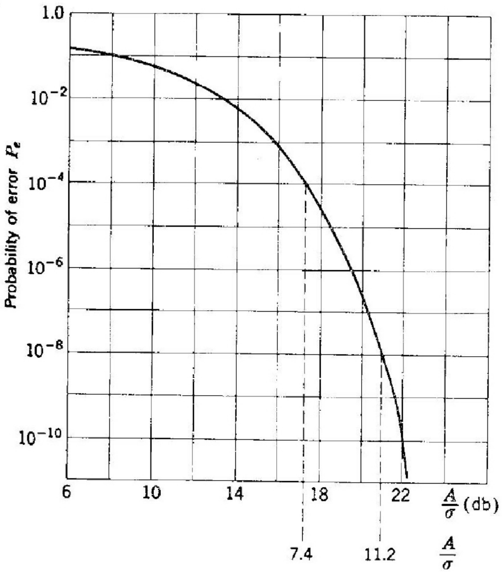
## Example 2

$$A/\sigma = 11.2(21 \text{ dB}) \Rightarrow P_e = 10^{-8}$$

For a transmission rate of $10^5$ bits/sec, there will be an error every 17 mins on the average.

Enormous increase in reliability by a relatively small increase in SNR (if that is affordable).

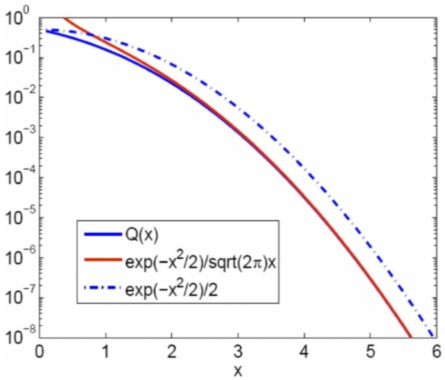
# Probability of Bit Error



# Q-Function

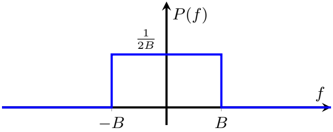
- Upper bounds and good approximations
- For $x \geq 0$, we have

$$Q(x) \leq \begin{cases} \frac{1}{\sqrt{2\pi}x} e^{-x^2/2}, & \text{(often used for large } x\text{)} \\ \frac{1}{2} e^{-x^2/2}, & \text{(only good for small } x\text{)} \end{cases}$$



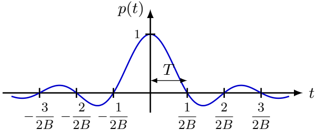
- A baseband waveform with bandwidth $B$

$$P(f) = \begin{cases} \frac{1}{2B}, & |f| \le B \\ 0, & \text{otherwise} \end{cases}$$




- Corresponding time-domain waveform or modulation pulse

$$p(t) = \int_{-B}^{B} \frac{1}{2B} e^{j2\pi f t} df$$
$$= \frac{\sin(2\pi Bt)}{2\pi Bt} = \mathrm{sinc}(2Bt)$$




$T = 1/2B$ for ISI-free transmission

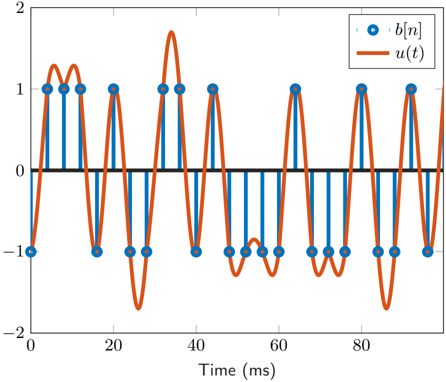
- Even if the *modulation pulse may be different*, we still regard the following to be true:
  - ✓ Symbol duration $T = \frac{1}{2B}$, or symbol rate $R = \frac{1}{T} = 2B$
  - ✓ For a modulated waveform with symbol duration $T$, bandwidth is $B = \frac{1}{2T} = \frac{R}{2}$

# Baseband Modulation

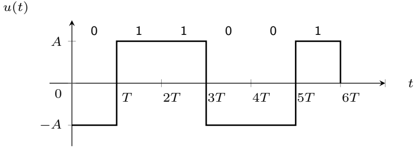
- Baseband modulation, linearly modulated waveform

$$u(t) = \sum_n b[n]p(t - nT)$$

✓ $\{b[n]\}$ : the sequence of symbols
✓ $p(t)$ : the modulating pulse



- Example: symbols $\{-1, +1\}$ and $p(t) = \begin{cases} 1, & 0 \leq t \leq T, \\ 0, & \text{otherwise} \end{cases}$


- Baseband → passband: $s(t) = u(t)\cos(2\pi f_c t)$

- For the $n$th symbol interval, $nT \leq t \leq (n+1)T$, we have

$$s(t) = \begin{cases} \cos(2\pi f_c t), & b[n] = +1, \\ \cos(2\pi f_c t + \pi), & b[n] = -1 \end{cases}$$

- Binary antipodal modulation switches the phase of the carrier between $0$ and $\pi$, hence it is called binary phase-shift keying (BPSK)

# PSK Signal $\rightarrow$ Constellations

- We can modulate both I and Q components (BPSK modulates only the I-component):

$$s(t) = u_{\mathsf{I}}(t)\cos(2\pi f_c t) - u_{\mathsf{Q}}(t)\sin(2\pi f_c t),$$

where

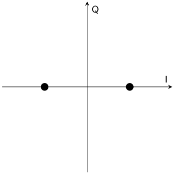
$$u_{\mathsf{I}}(t) = \sum_n b_{\mathsf{I}}[n]p(t-nT), \quad u_{\mathsf{Q}}(t) = \sum_n b_{\mathsf{Q}}[n]p(t-nT)$$


- If $b_{\mathsf{I}}[n]$ and $b_{\mathsf{Q}}[n]$ take values from $\{-1,+1\}$, for the $n$th symbol interval, $nT \le t \le (n+1)T$:

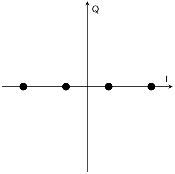
$$s(t) = \begin{cases} +\cos(2\pi f_c t) - \sin(2\pi f_c t) = \sqrt{2}\cos(2\pi f_c t + \pi/4), & \text{if } b_{\mathsf{I}}[n] = +1, b_{\mathsf{Q}}[n] = +1, \\ +\cos(2\pi f_c t) + \sin(2\pi f_c t) = \sqrt{2}\cos(2\pi f_c t - \pi/4), & \text{if } b_{\mathsf{I}}[n] = +1, b_{\mathsf{Q}}[n] = -1, \\ -\cos(2\pi f_c t) - \sin(2\pi f_c t) = \sqrt{2}\cos(2\pi f_c t + 3\pi/4), & \text{if } b_{\mathsf{I}}[n] = -1, b_{\mathsf{Q}}[n] = +1, \\ -\cos(2\pi f_c t) + \sin(2\pi f_c t) = \sqrt{2}\cos(2\pi f_c t + 3\pi/4), & \text{if } b_{\mathsf{I}}[n] = -1, b_{\mathsf{Q}}[n] = -1 \end{cases}$$

- Modulation switches the phase among $\pm\pi/4, \pm 3\pi/4$, called quadrature phase-shift keying (QPSK), 4-PSK, or 4-QAM

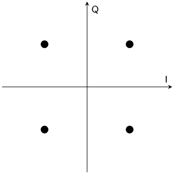
$$s(t) = u_{\mathsf{I}}(t)\cos(2\pi f_c t) - u_{\mathsf{Q}}(t)\sin(2\pi f_c t),$$



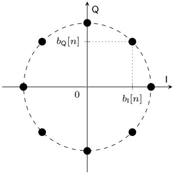

BPSK/2PAM



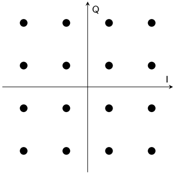

4PAM




QPSK/4PSK/4QAM



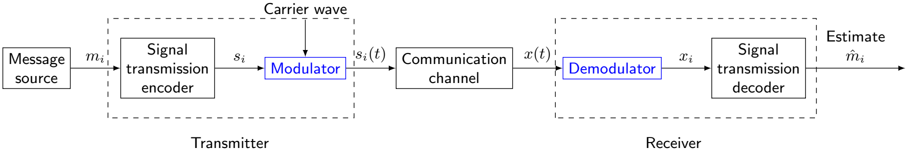

8PSK



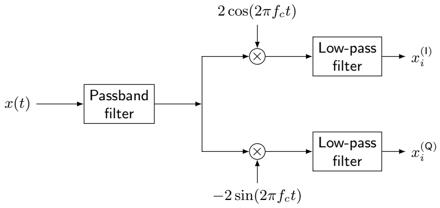

16QAM

- ASK and PSK: special cases of QAM
- FSK: not a special case

# QAM-Demodulation




- Coherent (synchronous) demodulation/detection
  - ✓ Use a band-pass filter (BPF) to reject out-of-band noise
  - ✓ Multiply the incoming waveform with a cosine and a sine of the carrier frequency
  - ✓ Use a low-pass filter (LPF)
  - ✓ Require carrier regeneration (both frequency and phase synchronization using a phase-locked loop)



# Coherent Detection of QAM

QAM signal:

$$s(t) = u_{\mathsf{I}}(t)\cos(2\pi f_c t) - u_{\mathsf{Q}}(t)\sin(2\pi f_c t),$$

Received signal:

$$x(t) = s(t) + n(t) = u_{\mathsf{I}}(t)\cos(2\pi f_c t) - u_{\mathsf{Q}}(t)\sin(2\pi f_c t) + n(t)$$

After passband filter,

$$\begin{aligned}\hat{x}(t) &= u_{\mathsf{I}}(t)\cos(2\pi f_c t) - u_{\mathsf{Q}}(t)\sin(2\pi f_c t) + n_{\mathsf{I}}(t)\cos(2\pi f_c t) - n_{\mathsf{Q}}(t)\sin(2\pi f_c t)\\ &= \big(u_{\mathsf{I}}(t) + n_{\mathsf{I}}(t)\big)\cos(2\pi f_c t) - \big(u_{\mathsf{Q}}(t) + n_{\mathsf{Q}}(t)\big)\sin(2\pi f_c t)\end{aligned}$$

Outputs of coherent detector:

$$\begin{aligned}x_{\mathsf{I}}(t) &= u_{\mathsf{I}}(t) + n_{\mathsf{I}}(t)\\ x_{\mathsf{Q}}(t) &= u_{\mathsf{Q}}(t) + n_{\mathsf{Q}}(t)\end{aligned}$$

# Note

# Lecture 13: ASK, PSK, FSK and Coherent Detection

Dr. Geoffrey Li

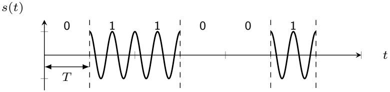
Department of Electrical & Electronic Engineering
Imperial College London

# Outline

- Amplitude-shift keying (ASK)
- Phase-shift keying (PSK)
- Frequency-shift keying (FSK)
- Coherent detection
  - ✓ BER of ASK, PSK, and FSK
- Minimum-shift keying (MSK)
- References
  - ✓ [Haykin] Chapter 9

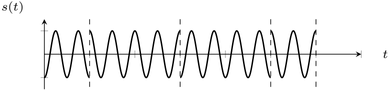
- Amplitude-shift keying (ASK)

$$s(t) = \begin{cases} A\cos(2\pi f_c t), & \text{if transmitting "1"} \\ 0, & \text{otherwise} \end{cases}$$



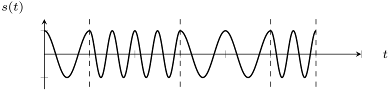

- Phase-shift keying (PSK)

$$s(t) = \begin{cases} A\cos(2\pi f_c t), & \text{if transmitting "1"} \\ A\cos(2\pi f_c t + \pi), & \text{otherwise} \end{cases}$$




- Frequency-shift keying (FSK)

$$s(t) = \begin{cases} A\cos(2\pi f_0 t), & \text{if transmitting "0"} \\ A\cos(2\pi f_1 t), & \text{if transmitting "1"} \end{cases}$$



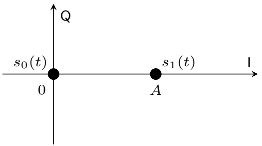
# Coherent ASK Detection

- Amplitude shift keying (ASK) = on-off keying (OOK)

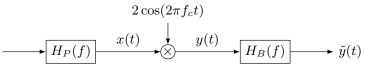
$$s_0(t) = 0, \quad s_1(t) = A\cos(2\pi f_c t)$$

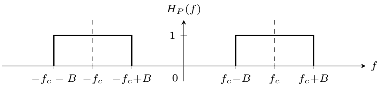
or

$$s(t) = A(t)\cos(2\pi f_c t), \quad A(t) \in \{0, A\}$$

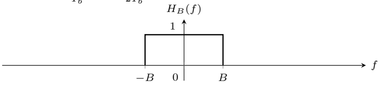



- Coherent detection




$$2B = \frac{1}{T_b}, \; B = \frac{1}{2T_b}$$



# Coherent Demodulation

- Pre-detection signal:

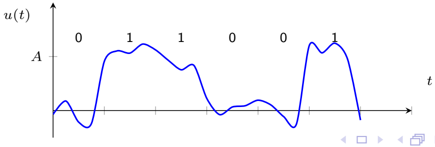
$$
\begin{aligned}
x(t) &= s(t) + n(t) \\
&= A(t)\cos(2\pi f_c t) + n_{\mathsf{I}}(t)\cos(2\pi f_c t) - n_{\mathsf{Q}}(t)\sin(2\pi f_c t) \\
&= \big(A(t) + n_{\mathsf{I}}(t)\big)\cos(2\pi f_c t) - n_{\mathsf{Q}}(t)\sin(2\pi f_c t)
\end{aligned}
$$

- After multiplication with $2\cos(2\pi f_c t)$ :

$$
\begin{aligned}
y(t) &= \big(A(t) + n_{\mathsf{I}}(t)\big)2\cos^2(2\pi f_c t) - n_{\mathsf{Q}}(t)2\sin(2\pi f_c t)\cos(2\pi f_c t) \\
&= \big(A(t) + n_{\mathsf{I}}(t)\big)\big(1 + \cos(4\pi f_c t)\big) - n_{\mathsf{Q}}(t)\sin(4\pi f_c t)
\end{aligned}
$$

- After low-pass filtering:

$$
\tilde{y}(t) = A(t) + n_{\mathsf{I}}(t)
$$

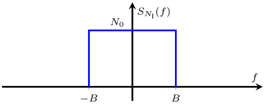


# Bit-Error Rate (BER) of ASK

- PSD of $n_\mathsf{I}(t)$: $N_0$
- For equiprobable transmission of 0s and 1s: decision threshold $\lambda = A/2$
- Probability of error:

$$P_{e,\mathsf{ASK}} = Q\Big(\frac{A}{2\sigma}\Big) \quad \text{(from Lect. 11)}$$

- Transmission energy for a pulse: $E = A^2 T_b/2$
- Average energy per bit: $E_b = A^2 T_b/4$
- Noise variance: $\sigma^2 = N_0 \times 2B = N_0/T_b$




- Probability of error in terms of energy per bit and noise PSD

$$P_{e,\mathsf{ASK}} = Q\Big(\sqrt{\frac{E_b}{N_0}}\Big)$$

## ASK Modulation System with Coherent Demodulation

- Carrier Amplitude $A = 0.7$ V
- Standard Deviation of White Gaussian Noise $\sigma = 0.125$ V
- Symbols "0" and "1" with equal probability

What is BER?

# PSK

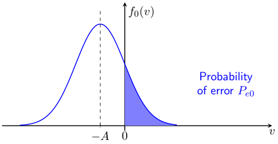
- PSK:

$$s(t) = A(t)\cos(2\pi f_c t), \quad A(t) \in \{-A, A\}$$

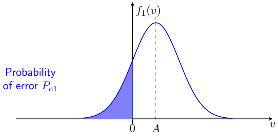
- Use coherent detection to eventually get detection signal:

$$\tilde{y}(t) = A(t) + n_{\mathrm{I}}(t)$$

- PDFs for PSK with equiprobable 0s and 1s in noise (use threshold 0 for detection):




Symbol 0 transmitted




Symbol 1 transmitted

# BER of PSK

$$P_{e0} = P_{e1} = P_{e,\mathsf{PSK}} = \int_A^\infty \frac{1}{\sigma\sqrt{2\pi}} e^{-y^2/2\sigma^2} dy$$

- Change variable of integration: $z \triangleq y/\sigma \Rightarrow dy = \sigma dz$

$$P_{e,\mathsf{PSK}} = \frac{1}{\sqrt{2\pi}} \int_{A/\sigma}^\infty e^{-z^2/2} dz = Q\left(\frac{A}{\sigma}\right)$$

- Average energy per bit: $E_b = A^2 T_b/2$
- Noise variance: $\sigma^2 = N_0/T_b$

$$\frac{A^2}{\sigma^2} = \frac{2E_b}{N_0}$$

- Probability of error in terms of energy per bit and noise PSD:

$$P_{e,\mathsf{PSK}} = Q\left(\sqrt{\frac{2E_b}{N_0}}\right)$$

## PSK Modulation System with Coherent Demodulation

- Carrier Amplitude $A = 0.7$ V
- Standard Deviation of White Gaussian Noise $\sigma = 0.125$ V
- Symbols "0" and "1" with equal probability

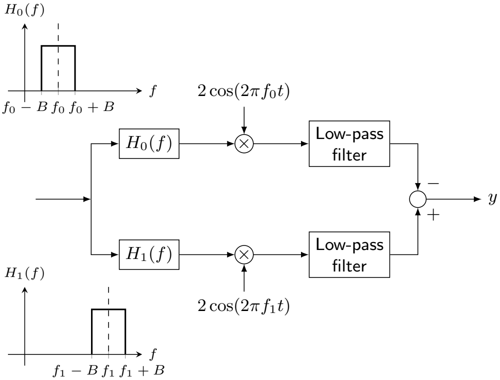
What is BER?

# FSK

$$s(t) = \begin{cases} s_0(t) = A\cos(2\pi f_0 t), & \text{if symbol 0 is transmitted} \\ s_1(t) = A\cos(2\pi f_1 t), & \text{if symbol 1 is transmitted} \end{cases}$$

Symbol recovery: two sets of coherent detectors at frequencies $f_0$ and $f_1$

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




Coherent FSK demodulation. The two BPF's are non-overlapping in frequency domain.

# FSK Output

- Each branch = an ASK detector

$$\text{LPF output on top branch} = \begin{cases} A + n_0(t), & \text{if symbol 0 is present} \\ n_0(t), & \text{if symbol 1 is present} \end{cases}$$

$$\text{LPF output on bottom branch} = \begin{cases} n_1(t), & \text{if symbol 0 is present} \\ A + n_1(t), & \text{if symbol 1 is present} \end{cases}$$

- $n_0(t), n_1(t)$: noise output of top and bottom branches, same statistics as $n_l(t)$!
- Output if transmitting "1"

$$y = y_1(t) = A + n_1(t) - n_0(t) \xlongequal{n(t)=n_1(t)-n_0(t)} A + n(t)$$

- Output if transmitting "0"

$$y = y_0(t) = -A + n(t)$$

- Average noise power

$$\mathbb{E}\{n^2(t)\} = \mathbb{E}\Big\{\big(n_1(t) - n_0(t)\big)^2\Big\} = \mathbb{E}\big\{n_0^2(t)\big\} + \mathbb{E}\big\{n_1^2(t)\big\} = 2\sigma^2$$

# BER of FSK

- Detection threshold $\lambda = 0$
- Noise term: $n(t) = n_1(t) - n_0(t)$
- Independent noise in the two channels: $\hat{\sigma}^2 = 2\sigma^2 = \frac{2N_0}{T_b}, \hat{\sigma} = \sqrt{2}\sigma$

$$P_{e,\text{FSK}} = Q\left(\frac{A}{\sqrt{2}\sigma}\right)$$

- Average energy per bit: $E_b = A^2 T_b / 2$
- Probability of error in terms of energy per bit and noise PSD:

$$P_{e,\text{FSK}} = Q\left(\sqrt{\frac{E_b}{N_0}}\right)$$

# Minimum-Shift Keying (MSK)

More on FSK:

- Symbol 0 $\rightarrow$ frequency $f_0$, symbol 1 $\rightarrow$ frequency $f_1$
- The unmodulated carrier frequency: $f_c = (f_0 + f_1)/2$
- Frequency separation: $\Delta f = |f_1 - f_0|$
- The symbol period: $T$
  - ✓ $f_c T \gg 1$ in practice, $\cos(2\pi f_1 t)$ and $\sin(2\pi f_0 t)$ orthogonal within the symbol period

$\Delta f$ (or $1/T$) should be large enough to make $\cos(2\pi f_1 t)$ and $\sin(2\pi f_0 t)$ orthogonal!

**Minimum-shift keying (MSK):** using the minimum separation $\Delta f = \frac{1}{2T}$

Why $\Delta f = \frac{1}{2T}$ is the minimum separation?

$$\frac{1}{T}\int_0^T \cos(2\pi f_1 t)\cos(2\pi f_0 t)dt = \frac{1}{2T}\int_0^T \cos\big(2\pi(f_1+f_0)t\big) + \cos\big(2\pi(f_1-f_0)t\big)dt$$

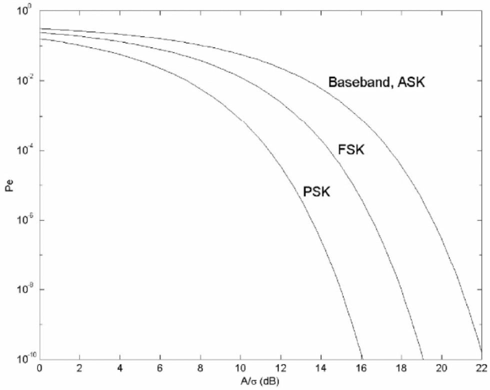
$$= \frac{1}{2T}\int_0^T \cos\big(4\pi f_c t\big) + \cos\big(2\pi \Delta f t\big)dt = \frac{1}{2}\left(\frac{\sin(4\pi f_c T)}{4\pi f_c T} + \frac{\sin(2\pi \Delta f T)}{2\pi \Delta f T}\right) \approx \frac{\sin(2\pi \Delta f T)}{4\pi \Delta f T}$$

$$2\pi \Delta f T = \pi \Rightarrow \Delta f = \frac{1}{2T} \text{ minimum separation!}$$

Whole bandwidth of MSK signal:

$$2B + \Delta f = 2 \times \frac{1}{2T} + \frac{1}{2T} = \frac{3}{2T}$$

# Comparison of Three Schemes




ASK:

- $\frac{E_b}{N_0} = \frac{A^2}{4\sigma^2}$
- $P_{e,\text{ASK}} = Q\left(\sqrt{\frac{E_b}{N_0}}\right) = Q\left(\frac{A}{2\sigma}\right)$

PSK:

- $\frac{E_b}{N_0} = \frac{A^2}{2\sigma^2}$
- $P_{e,\text{PSK}} = Q\left(\sqrt{\frac{2E_b}{N_0}}\right) = Q\left(\frac{A}{\sigma}\right)$

FSK:

- $\frac{E_b}{N_0} = \frac{A^2}{2\sigma^2}$
- $P_{e,\text{FSK}} = Q\left(\sqrt{\frac{E_b}{N_0}}\right) = Q\left(\frac{A}{\sqrt{2}\sigma}\right)$

# Application: GMSK

Gaussian minimum shift keying (GMSK), a special form of FSK preceded by Gaussian filtering, is used in GSM (Global Systems for Mobile Communications), a leading cellular phone standard in the world.

- Also known as digital FM, it was used in (Advanced Mobile Phone System) AMPS, the first-generation analog system (30 KHz bandwidth).
- Binary data are passed through a Gaussian filter to satisfy stringent requirements of out-of-band radiation.
- Minimum Shift Keying: its spacing between the two frequencies of FSK is minimum in a certain sense.
- GMSK is allocated bandwidth of 200 kHz, shared among 32 users. This provides a $(30/200) \times 32 = 4.8$ times improvement over AMPS.

## Note

# Lecture 14: Noncoherent Detection of Digital Modulation

Dr. Geoffrey Li

Department of Electrical & Electronic Engineering
Imperial College London

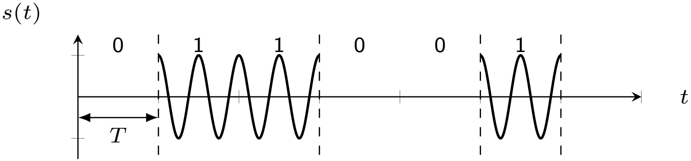
# Outline

- Noncoherent demodulation of ASK
- Noncoherent demodulation of FSK
- Differential demodulation of DPSK
- Reference
  - ✓ [Haykin] Chapter 9

# Coherent vs. Noncoherent Demodulation

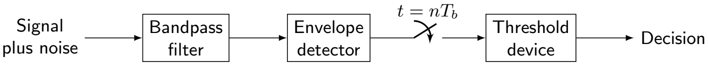
- Coherent demodulation assumes perfect synchronization
  - ✓ Needs a phase locked loop (very complicated)
- Accurate phase synchronization: difficult in a dynamic channel
  - ✓ Phase synchronization error is due to varying propagation delays, frequency drift, instability of the local oscillator, effects of strong noise ...
  - ✓ Performance of coherent detection will degrade severely
- For unknown phase, use non-coherent detection
  - ✓ No provision is made for carrier phase recovery
- Simpler circuitry/receiver but hard to analyze

# ASK Waveform




$$s(t) = \begin{cases} A\cos(2\pi f_c t), & \text{if transmitting "1"} \\ 0, & \text{otherwise} \end{cases}$$

# Noncoherent Demodulation of ASK




(Similar to noncoherent of AM)

- Output of the BPF

$$y(t) = \begin{cases} n(t), & \text{if 0 is sent,} \\ n(t) + A\cos(2\pi f_c t), & \text{if 1 is sent.} \end{cases}$$

- Recall

$$n(t) = n_{\mathsf{I}}(t)\cos(2\pi f_c t) - n_{\mathsf{Q}}(t)\sin(2\pi f_c t)$$

- Envelope

$$r(t) = \begin{cases} \sqrt{n_{\mathsf{I}}^2(t) + n_{\mathsf{Q}}^2(t)}, & \text{if 0 is sent,} \\ \sqrt{\left(A + n_{\mathsf{I}}(t)\right)^2 + n_{\mathsf{Q}}^2(t)}, & \text{if 1 is sent.} \end{cases}$$

# Distribution of the Envelope

$$r(t) = \begin{cases} \sqrt{n_{\mathsf{I}}^2(t) + n_{\mathsf{Q}}^2(t)}, & \text{if 0 is sent,} \\ \sqrt{\big(A + n_{\mathsf{I}}(t)\big)^2 + n_{\mathsf{Q}}^2(t)}, & \text{if 1 is sent.} \end{cases}$$

- Symbol 0 sent $\Rightarrow$ envelope $r(t) = \sqrt{n_{\mathsf{I}}^2(t) + n_{\mathsf{Q}}^2(t)}$: Rayleigh distribution

$$f(r) = \frac{r}{\sigma^2} e^{-r^2/2\sigma^2}, \quad r \geq 0 \quad \text{(Lect 3)}$$

- Symbol 1 sent $\Rightarrow$ envelope $r(t) = \sqrt{\big(A + n_{\mathsf{I}}(t)\big)^2 + n_{\mathsf{Q}}^2(t)}$: Rician distribution

$$f(r) = \frac{r}{\sigma^2} e^{-(r^2+A^2)/2\sigma^2} I_0\left(\frac{Ar}{\sigma^2}\right), \quad r \geq 0, \quad \text{(HW 1)}$$

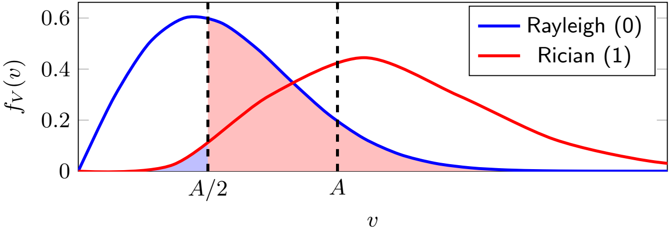
where $I_0(\cdot)$: the modified zero-order Bessel function of the first kind,

$$I_0(x) \triangleq \frac{1}{2\pi} \int_0^{2\pi} e^{x \cos\theta} d\theta$$

# Error Probability

- Let threshold $\lambda = A/2$ for simplicity
- Error probability: *dominated by symbol* $0$ and given by

$$P_{e,\text{ASK,noncoherent}} \approx \frac{1}{2}\int_{A/2}^{\infty} \frac{r}{\sigma^2} e^{-r^2/2\sigma^2} dr = \frac{1}{2} e^{-A^2/8\sigma^2}$$




- Coherent demodulation

$$P_{e,\text{ASK,coherent}} = Q\left(\frac{A}{2\sigma}\right) \le \frac{1}{2} e^{-A^2/8\sigma^2}$$

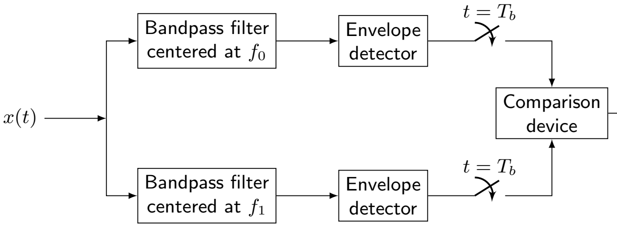
- Noncoherent demodulation results in some performance degradation

# ASK Example

## ASK Modulation System with Noncoherent Demodulation

- Carrier Amplitude $A = 0.7$ V
- Standard Deviation of White Gaussian Noise $\sigma = 0.125$ V
- Symbols "0" and "1" with equal probability

What is BER?




$$\begin{cases} \text{choose } 0, & \text{if } R_0 > R_1 \\ \text{choose } 1, & \text{if } R_0 < R_1 \end{cases}$$

# Distribution of Envelope

- If a symbol 1 is sent, output of the BPFs

$$y_0(t) = n_0(t)$$
$$y_1(t) = n_1(t) + A\cos(2\pi f_1 t)$$

- First branch: Rayleigh distribution

$$f_{R_0}(r_0) = \frac{r_0}{\sigma^2} e^{-r_0^2/2\sigma^2}, \quad r_0 \geq 0$$

- Second: Rician distribution

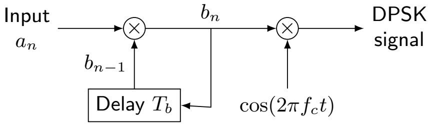
$$f_{R_1}(r_1) = \frac{r_1}{\sigma^2} e^{-(r_1^2+A^2)/2\sigma^2} I_0\left(\frac{Ar_1}{\sigma^2}\right), \quad r_1 \geq 0,$$

- Envelopes $R_0$ and $R_1$ are independent
- Error probability (derivation omitted)

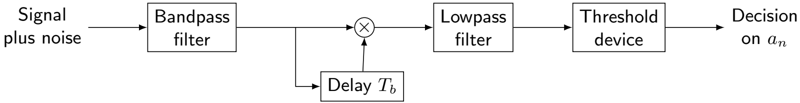
$$P_{e,\text{FSK,noncoherent}} = \text{Pr}\{R_1 < R_0\} = \frac{1}{2} e^{-A^2/4\sigma^2}$$

# DPSK: Differential PSK

- Impossible to demodulate PSK with an envelop detector since PSK signals have the same frequency and amplitude
- Demodulating PSK differentially, where phase reference is provided by a delayed version of the signal in the previous interval
- Essential to encode differentially: $b_n = b_{n-1} \times a_n$ where $a_n, b_n \in \pm 1$



# Differential Demodulation




- Probability of error ([Haykin] Chapter 9)

$$P_{e,\text{DPSK}} = \frac{1}{2} e^{-A^2/2\sigma^2}$$

- Coherent demodulation

$$P_{e,\text{PSK}} = Q\left(\frac{A}{\sigma}\right) \leq \frac{1}{2} e^{-A^2/2\sigma^2}$$

# Illustration of DPSK

| n | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|---|
| Information symbols $\{a_n\}$ | | 1 | -1 | -1 | 1 | -1 | -1 | 1 | 1 |
| $\{b_{n-1}\}$ | | 1 | 1 | -1 | 1 | 1 | -1 | 1 | 1 |
| DPSK sequence $\{b_n = a_n \times b_{n-1}\}$ | 1 | 1 | -1 | 1 | 1 | -1 | 1 | 1 | 1 |
| Transmitted phase (radians) | 0 | 0 | $\pi$ | 0 | 0 | $\pi$ | 0 | 0 | 0 |
| Output of lowpass filter (polarity) | | $+$ | $-$ | $-$ | $+$ | $-$ | $-$ | $+$ | $+$ |
| Decision | | 1 | -1 | -1 | 1 | -1 | -1 | 1 | 1 |

Symbol 1 ($b_0 = 1$) is inserted at the beginning of the differentially encoded sequence.

# PSK Example

## DPSK Modulation System with Differential Demodulation

- Carrier Amplitude $A = 0.7$ V
- Standard Deviation of White Gaussian Noise $\sigma = 0.125$ V
- Symbols "0" and "1" with equal probability

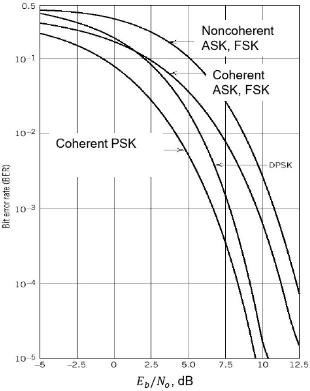
What is BER?

# Application: DPSK

- WLAN standard IEEE 802.11b
- Bluetooth2
- Digital audio broadcast (DAB): DPSK + OFDM (orthogonal frequency division multiplexing)
- Inmarsat (International Maritime Satellite Organization): now a London-based mobile satellite company

# Summary and Comparison

| Scheme | Bit-Error Rate (BER) |
|---|---|
| Coherent ASK | $Q\left(\frac{A}{2\sigma}\right) = Q\left(\sqrt{\frac{E_b}{N_0}}\right)$ |
| Coherent FSK | $Q\left(\frac{A}{\sqrt{2}\sigma}\right) = Q\left(\sqrt{\frac{E_b}{N_0}}\right)$ |
| Coherent PSK | $Q\left(\frac{A}{\sigma}\right) = Q\left(\sqrt{\frac{2E_b}{N_0}}\right)$ |
| Noncoherent ASK | $\frac{1}{2}\exp\left(-\frac{A^2}{8\sigma^2}\right) = \frac{1}{2}\exp\left(-\frac{E_b}{2N_0}\right)$ |
| Noncoherent FSK | $\frac{1}{2}\exp\left(-\frac{A^2}{4\sigma^2}\right) = \frac{1}{2}\exp\left(-\frac{E_b}{2N_0}\right)$ |
| Noncoherent PSK | $\frac{1}{2}\exp\left(-\frac{A^2}{2\sigma^2}\right) = \frac{1}{2}\exp\left(-\frac{E_b}{N_0}\right)$ |




- ASK: $E_b/N_0 = A^2/4\sigma^2$
- FSK: $E_b/N_0 = A^2/2\sigma^2$
- PSK: $E_b/N_0 = A^2/2\sigma^2$

# Conclusions

- Non-coherent demodulation retains the hierarchy of performance
- Non-coherent demodulation has error performance slightly worse than coherent demodulation, but approaches coherent performance at high SNR
- Non-coherent demodulators are considerably easier to build

# Note

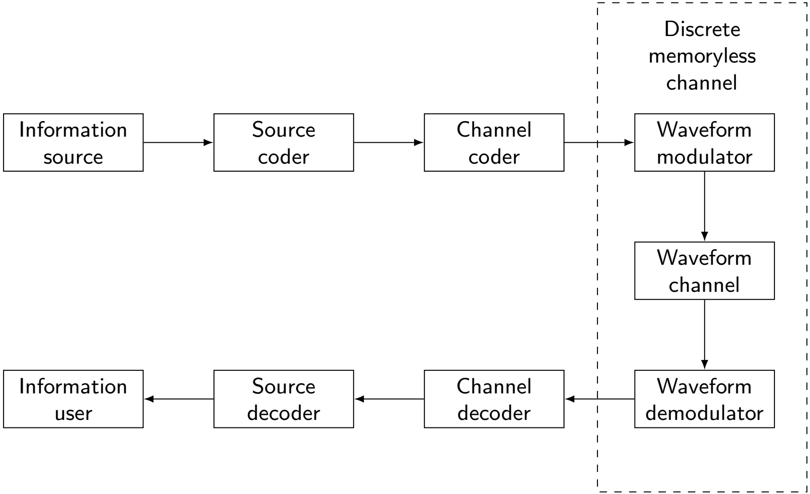
# Lecture 15: Information Theory

Dr. Geoffrey Li

Department of Electrical & Electronic Engineering
Imperial College London

# Outline

- Introduction to information theory
- Discrete memoryless source (DMS) and source entropy
- Discrete memoryless channel (DMC) and conditional entropy
- Mutual information and channel coding theorem
- Binary symmetric channel (BSC) and additive white Gaussian noise (AWGN) channel capacities
- Reference
  - ✓ [Haykin] Chapter 10



Discrete
memoryless
channel
Information
source
Source
coder
Channel
coder
Waveform
modulator
Waveform
channel
Information
user
Source
decoder
Channel
decoder
Waveform
demodulator

# What is Information?

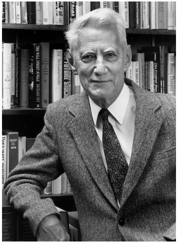
- Information: any new knowledge about something
  - ✓ How to store information efficiently?
  - ✓ How to transmit information over noisy channels?
- Information is everywhere
  - ✓ Collected by sensory system, transmitted via nervous system, processed in brain, ...
  - ✓ Stored in DNA, in hard-drives, in books, ...
  - ✓ Transmitted over the phone line, over the air, over generations, ...
- How to quantify information?

# What is Information Theory?

- C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, 1948

  *"The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point."*

- Two fundamental questions in information theory:
  - ✓ ultimate limit on data compression? (source coding)
  - ✓ ultimate transmission rate of reliable communication over noisy channels? (channel coding)



# What about non-uniform distribution?

**Example: Almost all students at Imperial are smart**

- Event that an imperial student is smart: *not so informative*
- Event that an imperial student is not smart: *very informative*

- Messages containing knowledge of a *high* probability of occurrence $\Rightarrow$ Not very informative
- Messages containing knowledge of *low* probability of occurrence $\Rightarrow$ More informative
- A small change in the probability of a certain output should not change the information delivered by that output by a large amount (it seems like a continuous function of the probability distribution)

# Definition

- Amount of information in a symbol $s$ with probability $p$:

$$I(s) = \log \frac{1}{p}$$

- Properties
  - ✓ $p = 1 \Rightarrow I(s) = 0$: a deterministic symbol contains no information
  - ✓ $0 \leq p \leq 1 \Rightarrow 0 \leq I(s) \leq \infty$: information measure is monotonic and non-negative
  - ✓ $p = p_1 \times p_2 \Rightarrow I(s) = I(s_1) + I(s_2)$: information from statistically independent events is additive

- Logarithm base 2 is commonly used, resulting in bits

## Example

# Discrete Memoryless Source (DMS)

- Suppose we have an information source emitting a sequence of symbols from a finite alphabet:

$$\mathcal{S} = \{s_1, s_2, \ldots, s_N\}$$

- Discrete memoryless source: The successive symbols are statistically independent and identically distributed (i.i.d.)
- Example: $\mathcal{S} = \{0, 1\}$, symbol sequence $= 001011000110\ldots$
- Assume that each symbol has probability $p_n$ for $n = 1, \ldots, N$, such that $\sum_{n=1}^{N} p_n = 1$

# Source Entropy

- We know that if symbol $s_n$ has occurred, this corresponds to amount of information,

$$I(s_n) = \log_2 \frac{1}{p_n} = -\log_2 p_n \text{ bits of information}$$

- For random variable $S \in \mathcal{S}$, expected value of $I(S)$ over the source alphabet

$$\mathbb{E}\{I(S)\} = \sum_{n=1}^{N} p_n I(s_n) = -\sum_{n=1}^{N} p_n \log_2 p_n$$

- Source entropy: average amount of information per source symbol

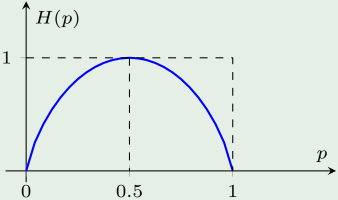
$$H(S) = -\sum_{n=1}^{N} p_n \log_2 p_n$$

- Units: bits/symbol

# Meaning of Entropy

- What does *entropy* tell us about the source?
- It is the amount of uncertainty before we receive it
- It tells us how many bits of information per symbol we get on the average by learning the source realization
- Relation with thermodynamic entropy
  - ✓ In thermodynamics: entropy measures disorder and randomness
  - ✓ In information theory: entropy measures uncertainty

## Entropy of a Binary Source



- Input alphabet: $\mathcal{X} = \{x_0, x_1, \ldots, x_{J-1}\}$
- Output alphabet: $\mathcal{Y} = \{y_0, y_1, \ldots, y_{K-1}\}$
- Transition probabilities (characterizing channel):

$$p(y_k|x_j) = P(Y = y_k|X = x_j), \quad \forall j, k$$

- Input probability distribution:

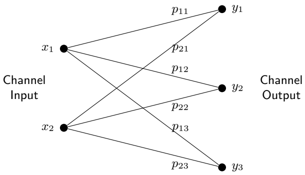
$$p(x_j) = P(X = x_j), \quad \forall j$$

- Joint probability distribution:

$$p(x_j, y_k) = p(y_k|x_j)p(x_j), \quad \forall j, k$$

- Marginal distribution of the channel output:

$$p(y_k) = P(Y = y_k) = \sum_{j=0}^{J-1} p(y_k|x_j)p(x_j), \quad \forall k$$




$$p_{jk} = p(y_k|x_j)$$

# Conditional Entropy

- Conditional entropy:

$$H(X|Y = y_k) = \sum_{j=0}^{J-1} p(x_j|y_k) \log_2 \frac{1}{p(x_j|y_k)}$$

- Probability of $H(X|Y = y_k)$:

$$\begin{array}{cccc} H(X|Y = y_0) & H(X|Y = y_1) & \dots & H(X|Y = y_{K-1}) \\ \downarrow & \downarrow & & \downarrow \\ p(y_0) & p(y_1) & \dots & p(y_{K-1}) \end{array}$$

- Average entropy:

$$H(X|Y) = \sum_{k=0}^{K-1} H(X|Y = y_k)p(y_k) = \sum_{k=0}^{K-1}\sum_{j=0}^{J-1} p(x_j|y_k)p(y_k) \log_2 \frac{1}{p(x_j|y_k)}$$
$$= \sum_{k=0}^{K-1}\sum_{j=0}^{J-1} p(x_j, y_k) \log_2 \frac{1}{p(x_j|y_k)}$$

- Interpretation: amount of uncertainty after observing the channel output

# Mutual Information

- Mutual information $I(X;Y)$: the uncertainty resolved by observing channel output

$$I(X;Y) = H(X) - H(X|Y) = H(Y) - H(Y|X)$$

- Also,

$$\begin{aligned} I(X;Y) &= \sum_{j=0}^{J-1} \sum_{k=0}^{K-1} p(x_j, y_k) \log_2 \frac{p(y_k|x_j)}{p(y_k)} \\ &= \sum_{k=0}^{K-1} \sum_{j=0}^{J-1} p(x_j, y_k) \log_2 \frac{p(x_j, y_k)}{p(x_j)p(y_k)} \end{aligned}$$

- Mutual information is
  - ✓ Non-negative: $I(X;Y) \geq 0$
  - ✓ Symmetric: $I(X;Y) = I(Y;X)$

- For a given channel $p(y_k|x_j)$ for $x_1, \ldots, x_J$ and $y_1, \ldots, y_K$, $I(X;Y)$ depends on $p(x_j)$ for $x_1, \ldots, x_J$

# Channel Capacity and Coding Theorem

- Capacity of a discrete memoryless channel is the maximum mutual information between the input and output, where the maximization is over all possible input probability distributions

$$C = \max_{p(x_0),\ldots,p(x_{J-1})} I(X;Y)$$

- How to calculate?
  - ✓ usually very complicated if analytically, except some symmetrical cases
  - ✓ easily to calculate numerically

## Channel Coding Theorem

- If the transmission rate $R \leq C$, then there exists a coding scheme such that R bits per channel use can be transmitted over the channel with an arbitrarily small probability of error.
- Conversely, if $R > C$, error probability is always bounded above zero when the transmission rate is above the capacity.

- How to code? We only know its existence but do not know how.
  - ✓ Polar code is the first code with an explicit construction to provably achieve the channel capacity for *symmetric binary-input*, discrete, memoryless channels (B-DMC) with polynomial dependence on the gap to capacity.

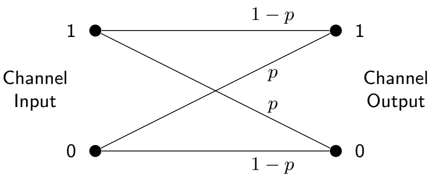
# Binary Symmetric Channel (BSC)

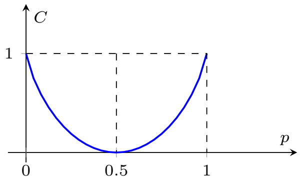
- Capacity of BSC

$$C = \max_{p(x_0),p(x_1)} I(X;Y) = 1 - h(p)$$

where

$$h(p) = -p\log_2 p - (1-p)\log_2(1-p)$$



# Additive White Gaussian Noise (AWGN) channel

- Capacity of an additive white Gaussian noise (AWGN) channel:

$$C = B\log_2(1 + \mathsf{SNR}) = B\log_2\left(1 + \frac{P}{N_0 B}\right) \text{ bps}$$

  - ✓ $B$: bandwidth of the channel
  - ✓ $P$: average signal power at the receiver
  - ✓ $N_0$: single-sided PSD of noise

- How can we achieve this rate?
  - ✓ Design powerful error correcting codes to correct as many errors as possible
  - ✓ Use good modulation schemes that do not lose information in the detection process
  - ✓ No simple way!

# Example

## Capacity of bandwidth-limited AWGN channel

# Note

# Lecture 16: Source Coding

Dr. Geoffrey Li

Department of Electrical & Electronic Engineering
Imperial College London

# Outline

- Average codeword length
- Fixed and variable length coding
- Source coding theorem
- Huffman coding
- Reference
  - ✓ [Haykin] Chapter 10

# Source Entropy

- If symbol $s_n$ has occurred, this corresponds to

$$I(s_n) = \log_2 \frac{1}{p_n} = -\log_2 p_n \text{ bits of information}$$

- For random variable $S$, expected value of $I(S)$ over the source alphabet

$$\mathbb{E}\{I(S)\} = \sum_{n=1}^{N} p_n I(s_n) = -\sum_{n=1}^{N} p_n \log_2 p_n$$

- Source entropy: average amount of information per source symbol

$$H(S) = -\sum_{n=1}^{N} p_n \log_2 p_n$$

- Units: bits/symbol

# Example

## Three-Symbol Alphabet

- $A$: occurs with probability 0.7
- $B$: occurs with probability 0.2
- $C$: occurs with probability 0.1

- Source entropy:

$$H(S) = -0.7\log_2(0.7) - 0.2\log_2(0.2) - 0.1\log_2(0.1) = 1.157 \text{ bits/symbol}$$

- How can we encode these symbols in order to transmit them?

- We need 2 bits/symbol if encoded as

$$A = 00, B = 01, C = 10 \quad \text{(fix-length coding)}$$

- Entropy prediction: the average amount of information is only 1.157 bits/symbol

- We are wasting bits!

- Source encoding: concerned with minimizing the actual number of source bits that are transmitted to the user
- Channel encoding: concerned with introducing redundant bits to enable the receiver to detect and possibly correct errors that are introduced by the channel.
- What is the minimum number of bits required to transmit a particular symbol?
- How can we encode symbols so that we achieve (or at least come arbitrarily close to) this limit?

# Average Codeword Length

- Definitions
  - ✓ $l_n$: number of bits used to code the $n$-th symbol
  - ✓ $N$: total number of symbols
  - ✓ $p_n$: probability of occurrence of symbol $n$
- Average codeword length

$$\bar{L} = \sum_{n=1}^{N} p_n l_n$$

- An idea to reduce average codeword length:
  - ✓ symbols that occur often should be encoded with short codewords
  - ✓ symbols that occur rarely may be encoded using long codewords
- Make sure that the codewords are uniquely decodable!

# Codeword Length

- In a system with 2 symbols that are equally likely:
  - ✓ Probability of each symbol to occur: $p = 1/2$, $H(p) = 1$ bit
  - ✓ Best one can do: encode each with 1 bit only (0 or 1), $\bar{L} = 1 = H(p)$ bit

- In a system with 2 symbols that are unequally likely:
  - ✓ $H(p) < 1$ bit
  - ✓ Encode each with 1 bit only (0 or 1), $\bar{L} = 1 > H(p)$ bit

- A system with $N$ ($N = 2^k$ for some integer $k$) symbols that are equally likely:
  - ✓ Probability of each symbol to occur: $p = 1/N$
  - ✓ One needs $\bar{L} = \log_2 N = k$ $(= -\log_2 p)$ bits to represent the symbols
  - ✓ For example, $N = 4$ requires $L = 2$ bits

- What is the minimum average codeword length for a particular source?

# Fixed Length Coding

Fixed length code: the same codeword length of different codewords.

**Example: 4-symbol source $p(a_1) = 1/2$, $p(a_2) = 1/4$, $p(a_3) = p(a_4) = 1/8$**

| Symbol (codeword) | Prob | Code I | Code II | Code III |
|---|---|---|---|---|
| $a_1$ | $1/2$ | 00 | 0 | 0 |
| $a_2$ | $1/4$ | 01 | 10 | 11 |
| $a_3$ | $1/8$ | 10 | 110 | 110 |
| $a_4$ | $1/8$ | 11 | 111 | 111 |

For example, using Code I:

$$a_1 a_3 a_4 a_3 \to 00, 10, 11, 10 \to 00101110$$
$$10110100 \to 10, 11, 01, 00 \to a_3 a_4 a_2 a_1$$

- Source entropy: $H(S) = \frac{1}{2}\log_2 2 + \frac{1}{4}\log_2 4 + 2 \times \frac{1}{8}\log_2 8 = 1.75$ bits
- Average length of codewords: $\bar{L} = 2 > H(S) = 1.75$

Fixed length coding is always **uniquely decodable** as long as you assign different symbols to different codewords!

# Variable Length Coding

Variable length code: codewords may have different lengths

**Example: 4-symbol source $p(a_1) = 1/2$, $p(a_2) = 1/4$, $p(a_3) = p(a_4) = 1/8$**

| Symbol (codeword) | Prob | Code I | Code II | Code III |
|---|---|---|---|---|
| $a_1$ | $1/2$ | 00 | 0 | 0 |
| $a_2$ | $1/4$ | 01 | 10 | 11 |
| $a_3$ | $1/8$ | 10 | 110 | 110 |
| $a_4$ | $1/8$ | 11 | 111 | 111 |

Using Code II:

- Encoding: $a_1a_3a_4a_1 \to 0, 110, 111, 0 \to 01101110$
- Decoding: $01101110 \to 0, 110, 111, 0 \to a_1a_3a_4a_1$

Using code III:

- Encoding: $a_1a_3a_4a_1 \to 0, 110, 111, 0 \to 01101110$
- Decoding:
  - ✓ $01101110 \to 0, 110, 111, 0 \to a_1a_3a_4a_1$
  - ✓ $01101110 \to 0, 11, 0, 111, 0 \to a_1a_2a_1a_4a_1$

Not uniquely decodable!

Some variable length codes are not uniquely decodable, and we only consider **uniquely decodable** coding subsequently.

# Source Coding Theorem

## Source Coding Theorem

Given a discrete memoryless source of entropy $H(S)$, average codeword length $\bar{L}$ for any **uniquely decodable source coding scheme** is (lower) bounded by $H(S)$, that is,

$$\bar{L} \geq H(S)$$

# Huffman Coding

- Basic Idea: choosing codeword lengths so that more-probable sequences have shorter codewords
- Code Construction:
  - ✓ Sort source symbols in order of decreasing probability
  - ✓ Take two smallest $p(x_i)$ and assign each a different bit (i.e., 0 or 1), then merge into a single symbol
  - ✓ Repeat until only one symbol remains
- Properties:
  - ✓ Huffman Coding (among other algorithms): uniquely decodable with average coding length satisfying $H(S) \leq \bar{L} < H(S) + 1$
  - ✓ The shortest average codeword length
  - ✓ Easy to implement this algorithm: used in JPEG, MP3, ...

# Example

## Huffman coding

# Compound Symbol using Huffman Coding

- Two symbol source: two symbols $s_1$, $s_2$
  - ✓ probabilities $\Pr(s_1) = p_1$, $\Pr(s_2) = p_2$
  - ✓ $H_1(S) = -p_1 \log_2 p_1 - p_2 \log_2 p_2$
  - ✓ Average length of Huffman code: $H_1(S) \leq \bar{L}_1 < H_1(S) + 1$

- Compound-symbol source by combining every two symbols:
  - ✓ Four compound symbols $s_1 s_1$, $s_1 s_2$, $s_2 s_1$, $s_2 s_2$
  - ✓ Probabilities

$$\Pr(s_1 s_1) = p_1^2, \Pr(s_1 s_2) = \Pr(s_2 s_1) = p_1 p_2, \Pr(s_2 s_2) = p_2^2$$

  - ✓ Compound-symbol source entropy

$$\begin{aligned} H_2(S) &= -\sum_{i,j} p_i p_j \log(p_i p_j) = -\sum_{i,j} p_i p_j \log(p_i) - \sum_{i,j} p_i p_j \log(p_j) \\ &= -\sum_{i} p_i \log(p_i) - \sum_{j} p_j \log(p_j) \\ &= 2H_1(S) \end{aligned}$$

  - ✓ Average length of Huffman code per the compound-symbol: $\bar{L}_2$

$$H_2(S) \leq \bar{L}_2 < H_2(S) + 1$$

  - ✓ Average length per symbol: $\bar{L}_2/2$

$$2H_1(S) \leq \bar{L}_2 < 2H_1(S) + 1 \Rightarrow H_1(S) \leq \frac{\bar{L}_2}{2} < H_1(S) + \frac{1}{2}$$

# Compound Symbol using Huffman Coding

- Compound-symbol source by combining K symbols:
  - ✓ Probability Pr$(s_{n_1} s_{n_2} \ldots s_{n_K}) = p_{n_1} p_{n_2} \ldots p_{n_K} = \prod_{k=1}^{K} p_{n_k}$
  - ✓ Compound-symbol source entropy

$$H_K(S) = K H_1(S)$$

  - ✓ Average length of Huffman code per the compound-symbol: $\bar{L}_K$

$$H_K(S) \leq \bar{L}_K < H_K(S) + 1$$

  - ✓ Average length per symbol: $\bar{L}_K / K$

$$H_1(S) \leq \frac{\bar{L}_K}{K} < H_1(S) + \frac{1}{K}$$

- When $K \to \infty$,

$$H_1(S) \leq \lim_{K \to \infty} \frac{\bar{L}_K}{K} \leq H_1(S) + \lim_{K \to \infty} \frac{1}{K}$$

Average length per symbol:

$$\lim_{K \to \infty} \frac{\bar{L}_K}{K} = H_1(S)$$

## Note

# Lecture 17: Channel Coding

Dr. Geoffrey Li

Department of Electrical & Electronic Engineering
Imperial College London

# Outline

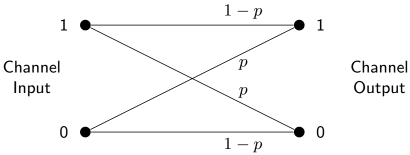
- Channel coding
- Linear block codes and generator matrix
- Hamming weight and distance
- Error detection and correction
- Reference
  - ✓ [Haykin] Chapter 10

# Channel Coding: Noise and Errors

- Noise can corrupt the information during transmission
- Corruption of a signal should be avoided if possible
- Different systems will generally require different levels of protection against errors
- Consequently, a number of different channel coding techniques have been developed to detect and correct different types and number of errors

# Channel Model

- Binary Symmetric Channel (BSC):




- Error probabilities are symmetric, errors are stationary and statistically independent
- $p$ is presumed to be less than $1/2$, or $p < 1 - p$

# Boolean Algebra I

- Two numbers: 0, 1
- Addition $+$: $0+0=1+1=0$; $0+1=1+0=1$
- Multiplication $\times$: $0\times 0=0\times 1=1\times 0=0$; $1\times 1=1$
- Calculation order: same as regular number calculation (multiplication first, from left to right), for example,

$$1\times 1+1\times 0+0=(1\times 1)+(1\times 0)+0=1+0+0=1$$

- Algebraic in Modulo 2:

$$\text{vector } \boldsymbol{a} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \text{matrix } \boldsymbol{B} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

$$\boldsymbol{B}\boldsymbol{a} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \times 1 + 1 \times 0 + 0 \times 1 \\ 1 \times 1 + 0 \times 0 + 1 \times 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

# Simple Error Checks: Repetition

If the error probability is small and information is fairly fault-tolerant, it is possible to use simple methods to detect errors (e.g., repetition, parity bits)

- Repetition – Repeating each bit in the message
  - ✓ If two symbols in an adjacent pair are different, it is likely that an error has occurred
  - ✓ However, this is not very efficient (bit rate is halved)
  - ✓ One repetition provides a means for error detection, but not for error correction
  - ✓ More repetitions are needed for error correction

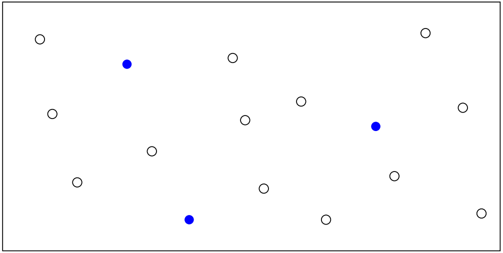
# Simple Error Checks: Adding a "Parity Bit"

- Parity bit – Use of a "parity bit" at the end of the message
  - ✓ A parity bit is a single bit that corresponds to the sum of the other message bits (modulo 2)
  - ✓ $\boldsymbol{u} = [u_1 \quad \ldots \quad u_k] \Rightarrow \boldsymbol{c} = [c_1 \quad \ldots \quad c_k \quad p]$, $p = c_1 + c_2 + \ldots + c_k$
  - ✓ For example, $011 \rightarrow 0110$; $010 \rightarrow 0101$
  - ✓ This allows any odd number of errors to be detected, but not even numbers
  - ✓ A single parity bit only allows error detection, not error correction
  - ✓ More efficient than simple repetition

# Block Codes

An important class of codes that can detect and correct some errors are block codes

- Encode a series of symbols from the source, a "block", into a longer string: codeword or code block
- Error detection: if the received coded block is not a valid codeword
- Error correction: "decode" and associate a corrupted block to a valid coded block by its proximity (as measured by the "Hamming distance")



● Codeword

○ Non-codeword

An $(n, k)$ binary linear block code takes a block of $k$ bits of source data and encodes them using $n$ bits.

$$\text{code rate} = \frac{\text{information bits}}{\text{codeword bits}} = \frac{k}{n}$$

- Linearity: the *Boolean sum* of any two codewords *must* be another codeword, e.g., if $\boldsymbol{a} = [1 \quad 0 \quad 0]$ and $\boldsymbol{b} = [1 \quad 0 \quad 1]$ are codewords, then $\boldsymbol{c} = \boldsymbol{a} + \boldsymbol{b} = [1+1 \quad 0+0 \quad 0+1] = [0 \quad 0 \quad 1]$ is, too
- The set of codewords forms a vector space, within which mathematical operations can be defined and performed

# Generator Matrix

- To construct a linear block code we define a matrix, the generator matrix $\boldsymbol{G}$, which converts blocks of source symbols into longer blocks corresponding to codewords
- $\boldsymbol{G}$ is a $k \times n$ matrix ($k$ rows, $n$ columns) that takes a source block $\boldsymbol{u}$ (a binary vector of length $k$), to a codeword $\boldsymbol{x}$ (a binary vector of length $n$)

$$\boldsymbol{x} = \boldsymbol{u} \cdot \boldsymbol{G}$$

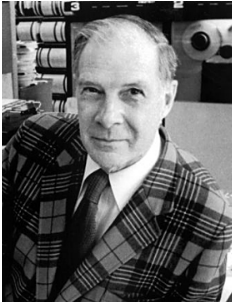
Message vector $\boldsymbol{u}$ → $\boxed{\text{Generator matrix } \boldsymbol{G}}$ → Code vector $\boldsymbol{x}$

$$\boldsymbol{u} = \begin{bmatrix} 1 & 0 \end{bmatrix}, \quad \boldsymbol{G} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}, \quad \boldsymbol{x} = \boldsymbol{u}\boldsymbol{G} = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \end{bmatrix}$$

- Linearity: Summation of two codewords is another codeword.
  - ✓ If $\boldsymbol{x}_1 = \boldsymbol{u}_1\boldsymbol{G}$, $\boldsymbol{x}_2 = \boldsymbol{u}_2\boldsymbol{G}$ are two codewords, then $\boldsymbol{x}_1 + \boldsymbol{x}_2 = \boldsymbol{u}_1\boldsymbol{G} + \boldsymbol{u}_2\boldsymbol{G} = (\boldsymbol{u}_1 + \boldsymbol{u}_2)\boldsymbol{G}$ is another codeword!

# Hamming Weight

Richard Hamming (1915 – 1998) established code theory and method when he worked at AT&T Bell Labs, New Jersey, USA.



- Hamming weight of a binary vector $\boldsymbol{a}$ (written as $w_{\mathrm{H}}(\boldsymbol{a})$), is the number of non-zero elements it contains. For example:
  - ✓ 001110011 has a Hamming weight of 5
  - ✓ 000000000 has a Hamming weight of 0

# Hamming Distance

- Hamming Distance between two binary vectors, $\boldsymbol{a}$ and $\boldsymbol{b}$, is written as $d_{\mathsf{H}}(\boldsymbol{a}, \boldsymbol{b})$, and is equal to the Hamming weight of their (Boolean) sum

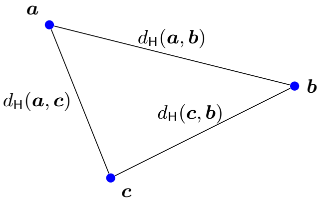
$$d_{\mathsf{H}}(\boldsymbol{a}, \boldsymbol{b}) = w_{\mathsf{H}}(\boldsymbol{a} + \boldsymbol{b})$$

For example, 01110011 and 10001011 have a Hamming distance of

$$d_{\mathsf{H}}(01110011, 10001011) = w_{\mathsf{H}}(01110011 + 10001011) = w_{\mathsf{H}}(11111000) = 5$$

- Triangle inequality

$$d_{\mathsf{H}}(\boldsymbol{a}, \boldsymbol{b}) \leq d_{\mathsf{H}}(\boldsymbol{a}, \boldsymbol{c}) + d_{\mathsf{H}}(\boldsymbol{c}, \boldsymbol{b})$$



# Example

## (7,4) Hamming Code

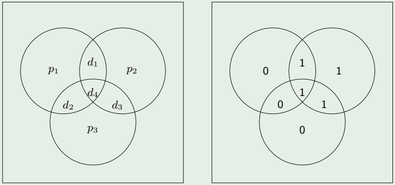
- Data bits: $\boldsymbol{u} = \begin{bmatrix} d_1 & d_2 & d_3 & d_4 \end{bmatrix}$
- Generator matrix:

$$\boldsymbol{G} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{P} \end{bmatrix}$$

- Codeword:

$$\boldsymbol{x} = \boldsymbol{uG} = \begin{bmatrix} d_1 & d_2 & d_3 & d_4 & d_2{+}d_3{+}d_4 & d_1{+}d_3{+}d_4 & d_1{+}d_2{+}d_4 \end{bmatrix}$$

- Parity bits: $p_1 = d_1{+}d_2{+}d_4, \quad p_2 = d_1{+}d_3{+}d_4, \quad p_3 = d_2{+}d_3{+}d_4$



# Example

## (7,4) Hamming Code

$$G = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}$$

- Data bits: $\boldsymbol{u}_1 = [1\ 0\ 1\ 0] \Rightarrow$ codeword $\boldsymbol{x}_1 = \boldsymbol{u}_1\boldsymbol{G} = [1\ 0\ 1\ 0\ 1\ 0\ 1]$
- Data bits: $\boldsymbol{u}_2 = [1\ 1\ 0\ 1] \Rightarrow$ codeword $\boldsymbol{x}_2 = \boldsymbol{u}_2\boldsymbol{G} = [1\ 1\ 0\ 1\ 0\ 0\ 1]$
- Data bits: $\boldsymbol{u}_3 = [0\ 0\ 1\ 0] \Rightarrow$ codeword $\boldsymbol{x}_3 = \boldsymbol{u}_3\boldsymbol{G} = [0\ 0\ 1\ 0\ 1\ 1\ 0]$
- Calculate the Hamming distances between each pair of codewords $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3$ and compare them to the Hamming distances between each pair of data bits $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3$

$$d_{\mathsf{H}}(\boldsymbol{x}_1, \boldsymbol{x}_2) = 4 > d_{\mathsf{H}}(\boldsymbol{u}_1, \boldsymbol{u}_2) = 3$$
$$d_{\mathsf{H}}(\boldsymbol{x}_1, \boldsymbol{x}_3) = 3 > d_{\mathsf{H}}(\boldsymbol{u}_1, \boldsymbol{u}_3) = 1$$
$$d_{\mathsf{H}}(\boldsymbol{x}_2, \boldsymbol{x}_3) = 7 > d_{\mathsf{H}}(\boldsymbol{u}_2, \boldsymbol{u}_3) = 4$$
$$\boldsymbol{x} = \boldsymbol{u}\boldsymbol{G} = \boldsymbol{u}\begin{bmatrix}\boldsymbol{I} & \boldsymbol{P}\end{bmatrix} = \begin{bmatrix}\boldsymbol{u} & \boldsymbol{u}\boldsymbol{P}\end{bmatrix}$$
$$d_{\mathsf{H}}(\boldsymbol{x}_i, \boldsymbol{x}_j) = d_{\mathsf{H}}(\boldsymbol{u}_i, \boldsymbol{u}_j) + d_{\mathsf{H}}(\boldsymbol{u}_i\boldsymbol{P}, \boldsymbol{u}_j\boldsymbol{P})$$

- To determine the number of errors a particular code can *detect* and *correct*, we look at the minimum Hamming distance between any two codewords.
- From linearity, the zero vector must be a codeword. The minimum Hamming distance of a code is the same as minimum weight of non-zero codewords.
- We define the minimum distance between any two codewords to be

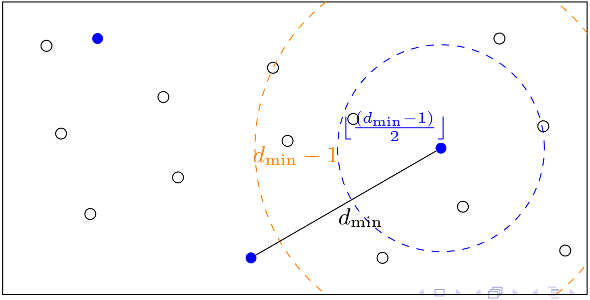
$$d_{\min} = \min_{\substack{\boldsymbol{a},\boldsymbol{b}\in\mathcal{C} \\ a\neq b}} d_{\mathsf{H}}(\boldsymbol{a},\boldsymbol{b}) = \min_{\substack{\boldsymbol{a},\boldsymbol{b}\in\mathcal{C} \\ a\neq b}} d_{\mathsf{H}}(0,\boldsymbol{a}+\boldsymbol{b}) = \min_{\boldsymbol{c}\in\mathcal{C},\boldsymbol{c}\neq\boldsymbol{0}} w_{\mathsf{H}}(\boldsymbol{c})$$

where $\mathcal{C}$ is the set of codewords.

- The number of errors that can be detected is then $d_{\min} - 1$ since $d_{\min}$ errors may turn an input codeword into a different valid codeword. Less than $d_{\min}$ errors will turn an input codeword into a vector that is not a valid codeword.

- Number $t$ of errors that can be corrected is $t = \lfloor (d_{\min} - 1)/2 \rfloor$, simply the number of errors that can be detected divided by two and rounded down to the nearest integer since any output vector with less than this number of errors will be "nearer" to the input codeword.

- (7,4) Hamming code has $d_{\min} = 3$. It can detect one or two bit errors, and correct any single bit error.



## Applications of Coding

- The first success was the application of convolutional codes in deep space probes 1960's-70's.
  - ✓ Mariner Mars, Viking, Pioneer missions by NASA
- Voyager, Galileo missions were further enhanced by concatenated codes (RS + convolutional).
- The next chapter was trellis coded modulation (TCM) for voice-band modems in 1980's.
- 1990's saw turbo codes approached capacity limit (now used in 3G).
- Followed by another breakthrough – space-time codes in 2000's (used in WiMax, 4G)
- The current frontier: LDPC, fountain codes, network coding, polar codes in 5G

# Note

# Recaps

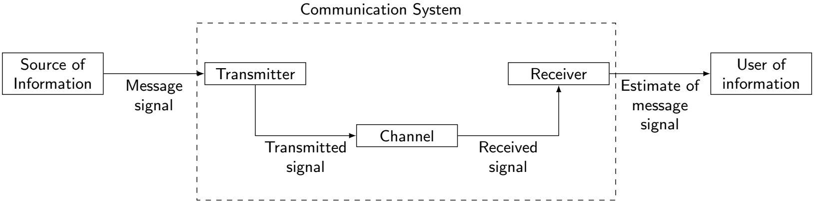
Dr. Geoffrey Li

Department of Electrical & Electronic Engineering
Imperial College London

# Table of Contents

# Communication Systems: Four Basic Elements

- Information source: voice, music, picture, video, ...
- Transmitter: converting information in the source into a form suitable for transmission over the channel
- Channel: the physical medium, introducing distortion, noise, and interference
- Receiver: reconstructing a recognizable form of the source signal




Block diagram of a communication system

# Table of Contents

# Fourier Transform (FT) and Inverse Fourier Transform (IFT)

| | In $f$-Domain | In $\omega$-Domain |
|---|---|---|
| FT | $X(f) = \int_{-\infty}^{\infty} x(t)e^{-j2\pi ft}dt$ | $X(\omega) = \int_{-\infty}^{\infty} x(t)e^{-j\omega t}dt$ |
| IFT | $x(t) = \int_{-\infty}^{\infty} X(f)e^{j2\pi ft}df$ | $x(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} X(\omega)e^{j\omega t}d\omega$ |

# Properties of the Fourier Transform

| Function | Fourier Transform |
|---|---|
| $a_1 x_1(t) + a_2 x_2(t)$ | $a_1 X_1(f) + a_2 X_2(f)$ |
| $x(t - t_0)$ | $X(f)e^{-j2\pi f t_0}$ |
| $x(at) \quad (a > 0)$ | $\frac{1}{a} X\left(\frac{f}{a}\right)$ |
| $x(-t)$ | $X(-f) = X^*(f)$ |
| $X(t)$ | $x(-f)$ |
| $x(t)e^{j\omega_0 t}$ | $X(f - f_0)$ |
| $x(t)\cos(\omega_0 t)$ | $\frac{1}{2} X(f - f_0) + \frac{1}{2} X(f + f_0)$ |
| $x(t)\sin(\omega_0 t)$ | $\frac{1}{2j} X(f - f_0) - \frac{1}{2j} X(f + f_0)$ |
| $\frac{d^n x(t)}{dt^n}$ | $(j2\pi f)^n X(f)$ |
| $\int_{-\infty}^{t} x(\tau)d\tau$ | $(j2\pi f)^{-1} X(f) + \frac{1}{2} X(0)\delta(f)$ |
| $\int_{-\infty}^{\infty} x_1(t-\tau)x_2(\tau)d\tau = \int_{-\infty}^{\infty} x_1(\tau)x_2(t-\tau)d\tau$ | $X_1(f)X_2(f)$ |
| $x_1(t)x_2(t)$ | $\int_{-\infty}^{\infty} X_1(f-\xi)X_2(\xi)d\xi = \int_{-\infty}^{\infty} X_1(\xi)x_2(f-\xi)d\xi$ |

# Mean and Variance

- Mean or expectation (corresponding to DC level of signals):

$$\mathbb{E}\{X\} = \mu_X = \int_{-\infty}^{\infty} x f_X(x) dx, \quad \mathbb{E}\{\cdot\} \text{ is the expectation operator}$$

- Variance (corresponding to power of zero-mean signals):

$$\sigma_X^2 = \mathbb{E}\{(X-\mu_X)^2\} = \int_{-\infty}^{\infty} (x-\mu_X)^2 f_X(x) dx = \mathbb{E}\{X^2\} - \mu_X^2$$

## Evaluation of mean

If $Y = G(X)$, then

$$\mathbb{E}\{Y\} = \int_{-\infty}^{\infty} G(x) f_X(x) dx$$

# Table of Contents

# Joint Distribution

- Joint cdf for two random variables $X$ and $Y$

$$F_{XY}(x,y) = P(X \leq x, Y \leq y)$$

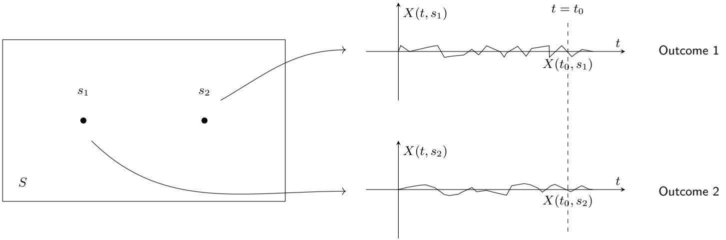
- Joint pdf

$$f_{XY}(x,y) = \frac{\partial^2 F_{XY}(x,y)}{\partial x \partial y}$$

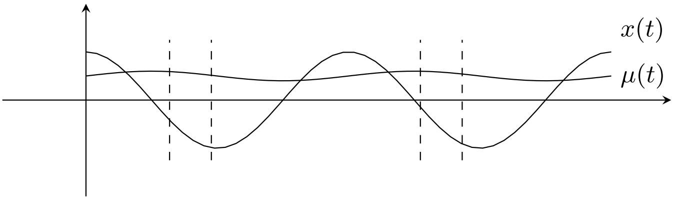
- Properties
  1. $F_{XY}(\infty,\infty) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(u,v)dudv = 1$
  2. $f_X(x) = \int_{y=-\infty}^{\infty} f_{XY}(x,y)dy,\ f_Y(y) = \int_{x=-\infty}^{\infty} f_{XY}(x,y)dx$
  3. Independent: $f_{XY}(x,y) = f_X(x)f_Y(y)$
  4. Uncorrelated: $\mathbb{E}\{XY\} = \mathbb{E}\{X\}\mathbb{E}\{Y\}$ or $\mathbb{E}\big\{(X-\mu_X)(Y-\mu_Y)\big\}$

# Stochastic Process

- Stochastic process $X(t, s)$: a collection of random variables over time. It represents the *evolution* of a random system.
- At a given time $t_0$, $X(t_0, s)$ is a random variable.
- At a sample outcome $s_j$, $X(t, s_j)$ is a deterministic function over time.
- Stochastic process $X(t, s)$ is often denoted by $X(t)$ for simplicity.
- Noise is often modelled as a Gaussian stochastic process.



- A stochastic process is wide-sense stationary (WSS) if and only if:
  1. The mean is not a function of time:
  $$\mu_X(t) = \mu_X, \quad \forall t$$
  2. The autocorrelation function only depends on time difference:
  $$R_X(t+\tau, t) = R_X(\tau), \quad \forall t, \tau$$




- Noise and message signals are often modelled as WSS processes.

# Table of Contents

# Stochastic Process Through Linear Time-Invariant (LTI) Systems

$$X(t) \longrightarrow \boxed{\text{Impulse Response } h(t)} \longrightarrow Y(t) = X(t) * h(t) = \int h(\tau)X(t-\tau)d\tau$$

- If the system is bounded-input bounded-output (BIBO) stable, we have

$$\mu_Y(t) = \mathbb{E}\{Y(t)\} = \mu_X(t) * h(t)$$
$$R_Y(t,u) = h(t) * h(u) * R_X(t,u)$$

- If $X(t)$ is real WSS, then

$$\mu_Y(t) = \mu_X \cdot \text{DC response} = \text{constant}$$
$$R_Y(\tau) = h(\tau) * h(-\tau) * R_X(\tau)$$
$$S_Y(f) = |H(f)|^2 S_X(f)$$

- If $X(t)$ is WSS / Gaussian, then $Y(t)$ is also WSS / Gaussian

# Power Spectral Density (PSD)

- PSD measures the distribution of power of a random process over its spectrum
- PSD is defined only for WSS processes, whose average power is

$$P = \mathbb{E}\{X^2(t)\} = R_X(0) = \int_{-\infty}^{\infty} S_X(f)df$$

**Einstein-Wiener-Khintchine relation:**

The PSD of a wide sense stationary process is equal to the Fourier transform of its autocorrelation function:

$$S_X(f) = \int_{-\infty}^{\infty} R_X(\tau)e^{-j2\pi f\tau}d\tau \geq 0$$

$$R_X(\tau) = \int_{-\infty}^{\infty} S_X(f)e^{j2\pi f\tau}df$$

# Table of Contents

**Energy and Power**

- Energy: $E = \int_{-\infty}^{\infty} |s(t)|^2 dt = \int_{-\infty}^{\infty} |S(f)|^2 df$
- Power for deterministic signal: $P = \frac{1}{T} \int_{-T/2}^{T/2} |s(t)|^2 dt$
- Power for random signal: $P = \mathbb{E}\{|s(t)|^2\} = R_S(0) = \int_{-\infty}^{\infty} S(f) df$

**Signal Bandwidth**

- Real-valued signal: conjugate-symmetric in the frequency domain; bandwidth measures one sideband
- Complex-valued signal: the complex envelope of a real-valued passband signal; not necessarily conjugate-symmetric; two-sided bandwidth across DC

**Baseband and Passband Signals**

- Baseband signal: usually complex-valued in time domain, real in special case; energy concentrated around DC
- Passband signal: always real in time domain and conjugate-symmetric in frequency domain; energy concentrated far away from DC

**Modulation and Demodulation**

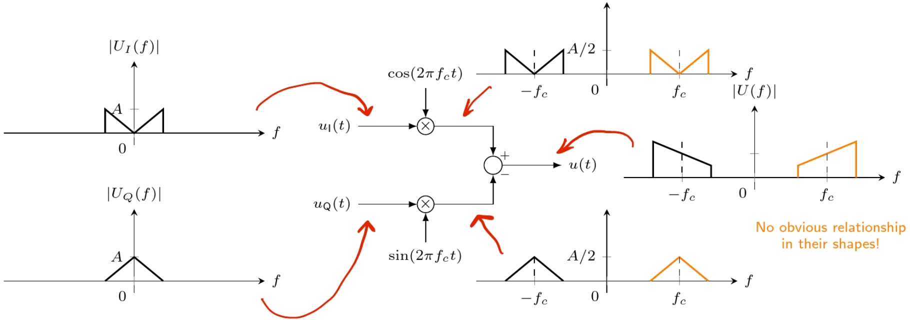
- Modulation (upconversion): converting baseband to passband; one passband signal maps to two baseband signals (passband modulation is two-dimensional)
- Demodulation (downconversion): converting passband to baseband

**Hilbert Transform**

- Hilbert transform: $\hat{g}(t) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{g(\tau)}{t-\tau} d\tau = g(t) * \frac{1}{\pi t}$,

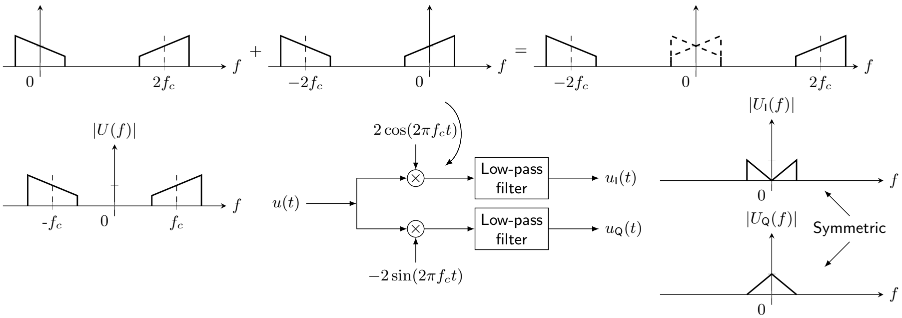
$$\hat{G}(f) = \mathcal{F}\left(\frac{1}{\pi t}\right) G(f) = -j\mathrm{sgn}(f) G(f), \quad \mathrm{sgn}(f) = \begin{cases} +1, & f > 0, \\ 0, & f = 0, \\ -1, & f < 0 \end{cases}$$

- Pre-envelope: $u_+(t) = u(t) + j\hat{u}(t)$ removes the negative frequency components




Since $u_\mathsf{I}(t)$ and $u_\mathsf{Q}(t)$ are real,
their spectra are conjugate symmetric

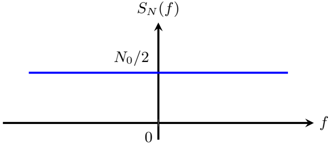
# Downconversion: Passband to Baseband

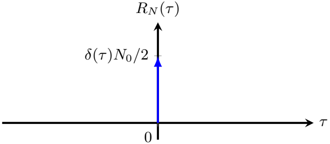


# Table of Contents

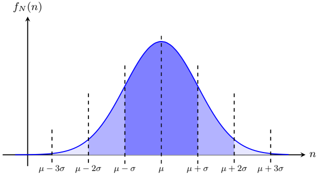
# White and Gaussian Noise

- White noise: shape of PSD is flat!
  - Autocorrelation function of $n(t)$: $R_N(\tau) = \mathcal{F}^{-1}\{S_N(f)\} = \frac{N_0}{2}\delta(\tau)$
  - Uncorrelated samples at different time instants
  - Also colored noise
  - White noise: either Gaussian or non-Gaussian




- Gaussian noise:
  - Gaussian distribution for a sample at any *time instant*
  - Colored or white Gaussian noise




- White noise (shape of PSD) and Gaussian noise (distribution, CDF) are different concepts
- In communications, typically additive white Gaussian noise (AWGN)

# Properties of Baseband Noise

- If noise $n(t)$ is Gaussian, then so are $n_{\mathsf{I}}(t)$ and $n_{\mathsf{Q}}(t)$
- $n_{\mathsf{I}}(t)$ and $n_{\mathsf{Q}}(t)$ have the same variance (i.e., same power) as $n(t)$
- Both in-phase and quadrature components have the same PSD:

$$S_{N_I}(f) = S_{N_Q}(f) = \begin{cases} S_N(f - f_c) + S_N(f + f_c), & |f| \leq B \\ 0, & \text{otherwise} \end{cases}$$

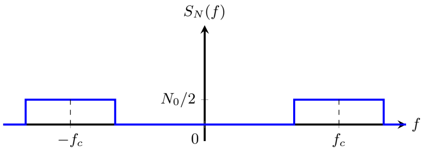
# Noise Power

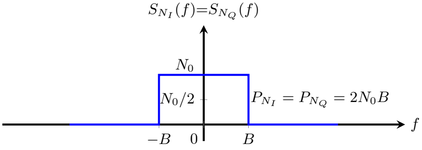
- For ideally filtered narrowband white noise, the PSDs of $n_{\mathsf{I}}(t)$ and $n_{\mathsf{Q}}(t)$:

$$S_{N_I}(f) = S_{N_Q}(f) = \begin{cases} N_0, & |f| \leq B, \\ 0, & \text{otherwise} \end{cases}$$

- The average power in *each* of the baseband waveforms $n_{\mathsf{I}}(t)$ and $n_{\mathsf{Q}}(t)$ is identical to the average power in the bandpass noise waveform $n(t)$.
- For ideally filtered narrowband noise, the variance of $n_{\mathsf{I}}(t)$ and $n_{\mathsf{Q}}(t)$:

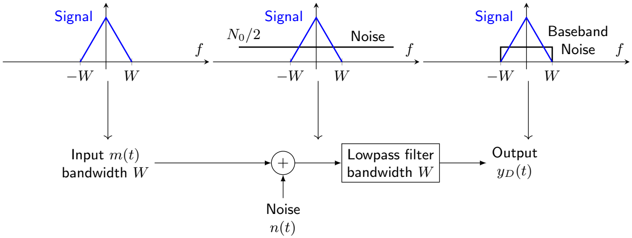
$$P_{N_I} = P_{N_Q} = 2N_0 B$$



# Table of Contents

# Baseband Communication

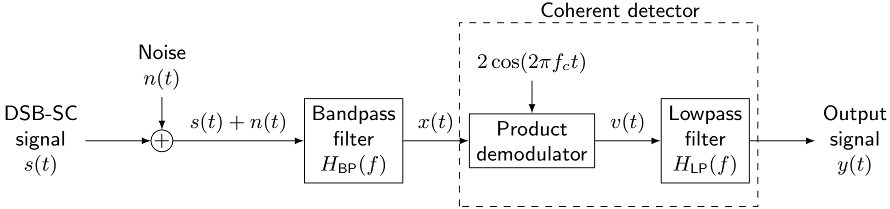



- Transmit power $P_T$ = receive power $P_R$ = message power $P$ = area under the triangular
- Average noise power $P_N$ = area under the straight line $= 2WN_0/2 = WN_0$
- SNR at receiver output:

$$\text{SNR}_{\text{baseband}} = \frac{P_T}{N_0 W}$$

- Improve SNR by:
  - Increasing the transmitted power $P_T \uparrow$
  - Making the channel/receiver less noisy $N_0 \downarrow$

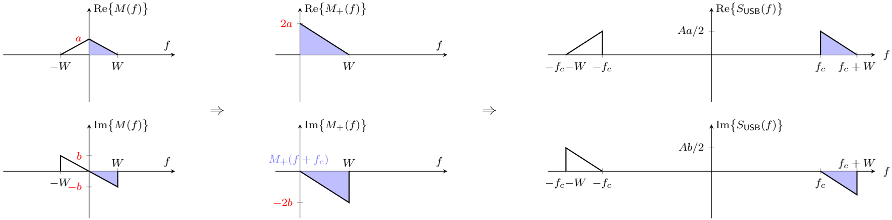
# Double Sideband-Suppressed Carrier (DSB-SC) Modulation




- DSB-SC modulated signal $s(t) = m(t)A\cos(2\pi f_c t)$
- Coherent detection: multiply received signal by $2\cos(2\pi f_c t)$ and apply LP filter
- Transmit power $P_T = \mathbb{E}\{A^2 m^2(t)\cos^2(2\pi f_c t)\} = A^2 P/2$
- Receiver output power $P_R = \mathbb{E}\{A^2 m^2(t)\} = A^2 P$
- Noise power $P_N = 2N_0 W$
- SNR at receiver output:

$$\text{SNR}_{\text{DSB-SC}} = \frac{A^2 P}{2N_0 W} = \frac{P_T}{N_0 W} = \text{SNR}_{\text{baseband}}$$

# Table of Contents




- SSB modulated signal $s(t) = \frac{A}{2} \underbrace{m(t)}_{\text{I component}} \cos(2\pi f_c t) - \frac{A}{2} \underbrace{\hat{m}(t)}_{\text{Q component}} \sin(2\pi f_c t)$
- Transmit power $P_T = A^2 P/4 =$ receiver output power $P_R$
- Noise power $P_N = N_0 W$ (halved compared to DSB)
- SNR at receiver output:

$$\text{SNR}_{\text{SSB}} = \frac{A^2 P}{4 N_0 W} = \frac{P_T}{N_0 W} = \text{SNR}_{\text{baseband}}$$

# Amplitude Modulation

- Amplitude modulated signal $s(t) = \big(A + m(t)\big)\cos(2\pi f_c t)$
- Envelope detection requires modulation index $\mu = m_p/A \leq 1$; coherent detection does not
- Transmit power $P_T = (A^2 + P)/2$
- For coherent detection:
  - Receiver output power $P_S = P$
  - Noise power $P_N = 2N_0 W$
  - SNR at the receiver output:

$$\mathsf{SNR}_{\mathsf{AM}} = \frac{P}{2N_0 W} = \frac{P}{A^2 + P}\mathsf{SNR}_{\mathsf{baseband}}$$

- For envelope detection with small noise, SNR performance is close to coherent detection

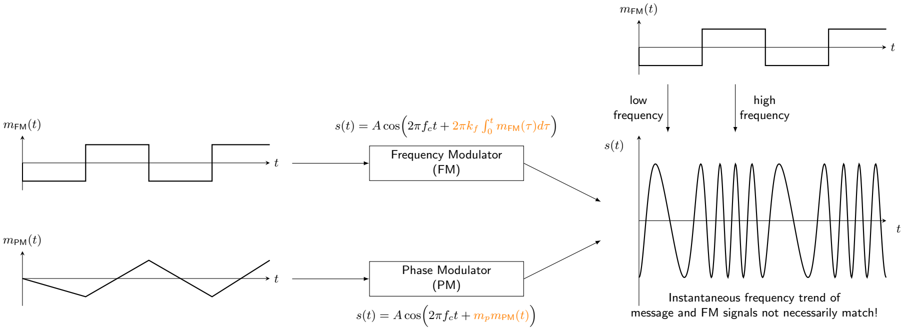
# Summary

| (De-) modulation format | Output SNR | Transmitted power | Baseband reference SNR | Output SNR / reference SNR |
|---|---|---|---|---|
| AM coherent detection | $\frac{P}{2N_0W}$ | $\frac{A^2+P}{2}$ | $\frac{A^2+P}{2N_0W}$ | $\frac{P}{A^2+P} < 1$ |
| DSB-SC coherent detection | $\frac{A^2P}{2N_0W}$ | $\frac{A^2P}{2}$ | $\frac{A^2P}{2N_0W}$ | 1 |
| SSB coherent detection | $\frac{A^2P}{4N_0W}$ | $\frac{A^2P}{4}$ | $\frac{A^2P}{4N_0W}$ | 1 |
| AM envelope detection (small noise) | $\frac{P}{2N_0W}$ | $\frac{A^2+P}{2}$ | $\frac{A^2+P}{2N_0W}$ | $\frac{P}{A^2+P} < 1$ |
| AM envelope detection (large noise) | Poor | $\frac{A^2+P}{2}$ | $\frac{A^2+P}{2N_0W}$ | Poor |

- $A$: carrier amplitude
- $P$: power of message signal
- $N_0$: single-sided PSD of noise
- $W$: message bandwidth

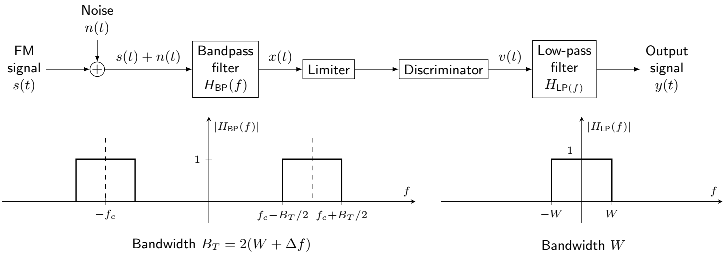
# Table of Contents

# PM-FM Equivalence




- Transmit power $P_T = A^2/2$
- Bandwidth $B_T = 2W(\beta + 1) = 2(\Delta f + W)$, where $\Delta f = k_f m_p, \beta = \Delta f/W$

# FM Receiver




$$\text{SNR}_{\text{FM}} = \frac{3A^2 k_f^2 P}{2N_0 W^3} = 3\beta^2 \frac{P}{m_p^2} \text{SNR}_{\text{baseband}}$$

# Comparison of Analog Systems

- Assumptions:
  - Single-tone modulation $m(t) = A_m \cos(2\pi f_m t)$
  - Message bandwidth $W = f_m$
  - For AM system, modulation index $\mu = m_p/A = A_m/A = 1$, $m_p = \max|m(t)| = A_m$
  - For FM system, modulation index $\beta = \Delta f/W = 5$, $\Delta f = k_f m_p = k_f A_m$ (used in commercial FM transmission with $\Delta f = 75$ kHz and $W = 15$ kHz)

- SNR expressions for various modulation schemes

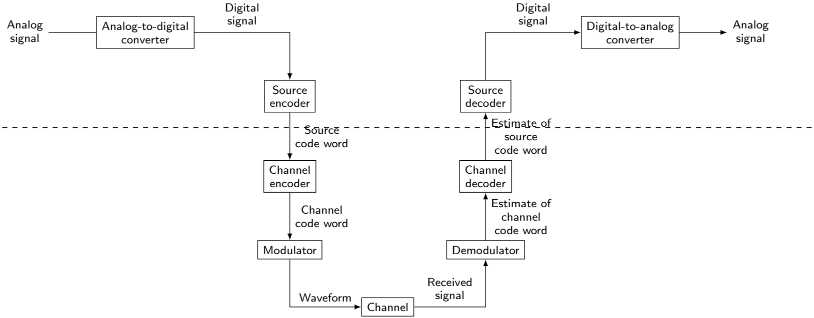
$$\text{SNR}_{\text{DSB-SC}} = \text{SNR}_{\text{baseband}} = \text{SNR}_{\text{SSB}}$$

$$\text{SNR}_{\text{AM}} = \frac{P}{A^2 + P}\text{SNR}_{\text{baseband}} = \frac{\mu^2}{2+\mu^2}\text{SNR}_{\text{baseband}} \leq \frac{1}{3}\text{SNR}_{\text{baseband}}$$

$$\text{SNR}_{\text{FM}} = \frac{3\beta^2}{2}\text{SNR}_{\text{baseband}} = \underbrace{\frac{75}{2}}_{15.7\text{dB}}\text{SNR}_{\text{baseband}} \quad \text{(without pre/de-emphasis)}$$

# Table of Contents

# Block Diagram of Digital Communication



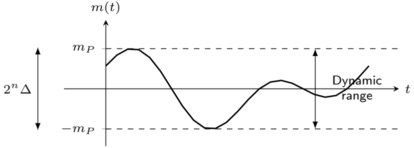
### Sampling theorem

For distortionless recovery, sampling rate $f_s \geq 2W$ for a (real) signal with bandwidth $W$. The Nyquist frequency is

$$f_{\mathrm{N}} = 2W$$

# Quantization Gap $\Delta$

- Assume that encoded symbol has $n$ bits
  - maximum number of quantizing levels is $L = 2^n$
  - maximum peak-to-peak dynamic range of the quantizer $2^n\Delta$




- $P$ : power of the message signal

$$P = \mathbb{E}\{m^2(t)\} \xlongequal{\text{periodic}} \frac{1}{T}\int_{-T/2}^{T/2} |m(t)|^2 dt$$

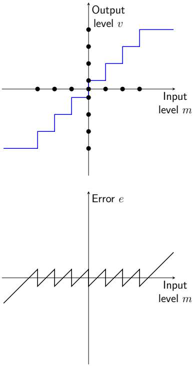
- SNR at quantizer output:

$$\text{SNR}_\text{o} = \frac{3P}{m_p^2} 2^{2n}$$

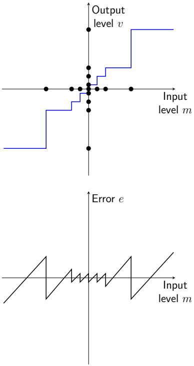
$$\text{SNR}_\text{o}(\text{dB}) = 6n + 10\log_{10}\left(\frac{3P}{m_p^2}\right)$$

- An extra bit in the encoder improves the output SNR by 6 dB

# Uniform and Nonuniform Quantization




Uniform

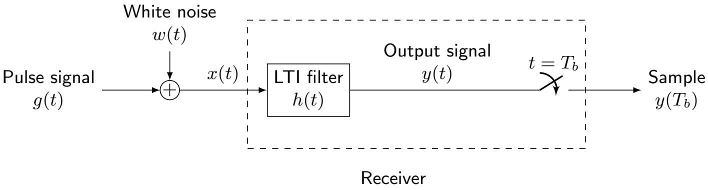



Nonuniform

- Compressing and expanding

# Table of Contents

# Matched Filter

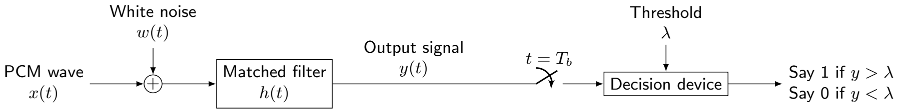



$$x(t) = g(t) + w(t), \quad 0 \le t \le T_b$$
$$y(t) = x(t) * h(t)$$

## Matched filter

- optimal $h(t)$ to maximize the SNR at $y(T_b)$:

$$h_{\text{opt}}(t) = kg(T_b - t) \quad (k \neq 0)$$




- If a symbol 0 was transmitted, $Y = N$
  - $Y \sim \mathcal{N}(0, \sigma^2)$, $f_0(y) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{y^2}{2\sigma^2}\right)$
- If a symbol 1 was transmitted, $Y = A + N$
  - $Y \sim \mathcal{N}(A, \sigma^2)$, $f_1(y) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(y-A)^2}{2\sigma^2}\right)$
- Use $\lambda$ as the decision threshold:
  - choose symbol 0 if $y < \lambda$
  - choose symbol 1 if $y > \lambda$

# Error Cases

Case I:

$$\text{Prob}(\text{error}|\text{symbol 0 was transmitted}) \times \text{Prob}(\text{symbol 0 was transmitted})$$

$$P_{\mathrm{I}} = P_{e0} \times p_0$$

- $p_0$ : a priori probability of transmitting a symbol 0
- $P_{e0}$ : conditional probability of error if symbol 0 was transmitted

$$P_{e0} = \int_{\lambda}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{y^2}{2\sigma^2}\right) dy$$

Case II:

$$\text{Prob}(\text{error}|\text{symbol 1 was transmitted}) \times \text{Prob}(\text{symbol 1 was transmitted})$$

$$P_{\mathrm{II}} = P_{e1} \times p_1$$

- $p_1$ : a priori probability of transmitting a symbol 1
- $P_{e1}$ : conditional probability of error if symbol 1 was transmitted

$$P_{e1} = \int_{-\infty}^{\lambda} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(y-A)^2}{2\sigma^2}\right) dy$$

- Total error probability:

$$P_e(\lambda) = P_{\mathsf{I}} + P_{\mathsf{II}} = p_0 P_{e0} + p_1 P_{e1}$$

- Optimal threshold:

$$\lambda_{\text{opt}} = -\frac{\sigma^2}{A} \log \frac{p_1}{1 - p_1} + \frac{A}{2}$$

- If $p_0 = p_1 = 0.5$:

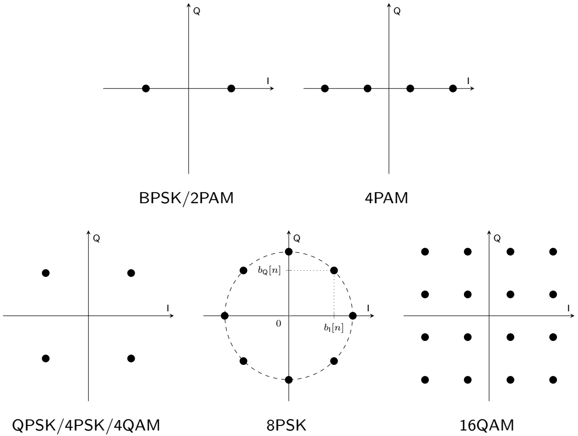
$$\lambda_{\text{opt}} = \frac{A}{2}$$

$$P_e = P_{e_0} = P_{e_1} = Q\left(\frac{A}{2\sigma}\right) = Q\left(\sqrt{\frac{E_b}{N_0}}\right)$$

# Table of Contents

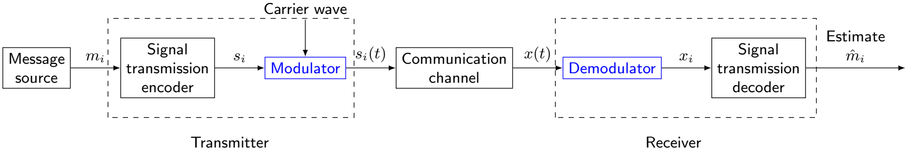
# Quadrature Amplitude Modulation (QAM)

$$s(t) = u_{\mathsf{I}}(t)\cos(2\pi f_c t) - u_{\mathsf{Q}}(t)\sin(2\pi f_c t),$$

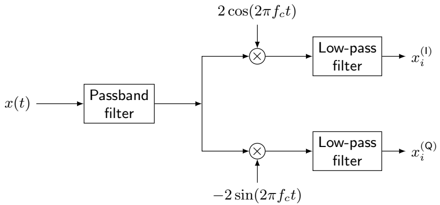



- ASK and QSK: special cases of QAM
- FSK: not a special case

# QAM-Demodulation




- Coherent (synchronous) demodulation/detection
  - Use a band-pass filter (BPF) to reject out-of-band noise
  - Multiply the incoming waveform with a cosine and a sine of the carrier frequency
  - Use a low-pass filter (LPF)
  - Require carrier regeneration (both frequency and phase synchronization using a phase-locked loop)

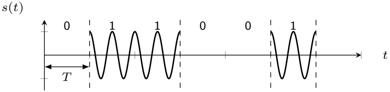


# Table of Contents

# Basic Forms: ASK, PSK, and FSK

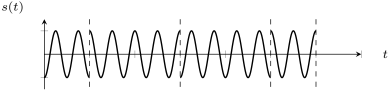
- Amplitude-shift keying (ASK)

$$s(t) = \begin{cases} A\cos(2\pi f_c t), & \text{if transmitting “1”} \\ 0, & \text{otherwise} \end{cases}$$



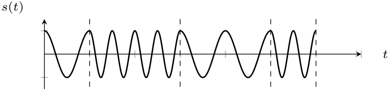

- Phase-shift keying (PSK)

$$s(t) = \begin{cases} A\cos(2\pi f_c t), & \text{if transmitting “1”} \\ A\cos(2\pi f_c t + \pi), & \text{otherwise} \end{cases}$$

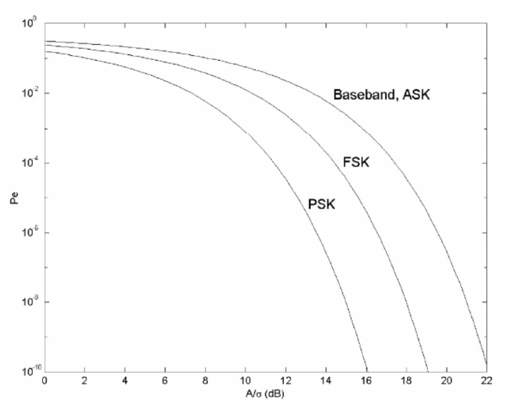



- Frequency-shift keying (FSK)

$$s(t) = \begin{cases} A\cos(2\pi f_0 t), & \text{if transmitting “0”} \\ A\cos(2\pi f_1 t), & \text{if transmitting “1”} \end{cases}$$



# Comparison of Three Schemes




ASK:

- $\frac{E_b}{N_0} = \frac{A^2}{4\sigma^2}$
- $P_{e,\text{ASK}} = Q\left(\sqrt{\frac{E_b}{N_0}}\right) = Q\left(\frac{A}{2\sigma}\right)$

PSK:

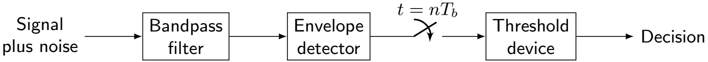
- $\frac{E_b}{N_0} = \frac{A^2}{2\sigma^2}$
- $P_{e,\text{PSK}} = Q\left(\sqrt{\frac{2E_b}{N_0}}\right) = Q\left(\frac{A}{\sigma}\right)$

FSK:

- $\frac{E_b}{N_0} = \frac{A^2}{2\sigma^2}$
- $P_{e,\text{FSK}} = Q\left(\sqrt{\frac{E_b}{N_0}}\right) = Q\left(\frac{A}{\sqrt{2}\sigma}\right)$

# Table of Contents

# Noncoherent Demodulation of ASK




(Similar to noncoherent of AM)

- Output of the BPF

$$y(t) = \begin{cases} n(t), & \text{if 0 is sent,} \\ n(t) + A\cos(2\pi f_c t), & \text{if 1 is sent.} \end{cases}$$

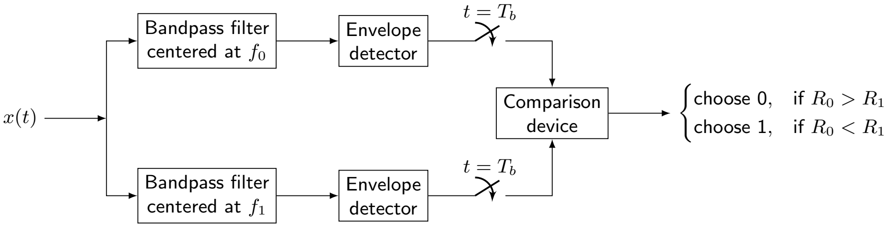
- Recall

$$n(t) = n_{\mathsf{I}}(t)\cos(2\pi f_c t) - n_{\mathsf{Q}}(t)\sin(2\pi f_c t)$$

- Envelope

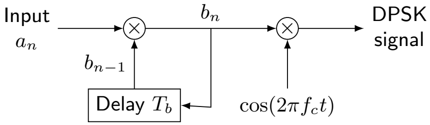
$$r(t) = \begin{cases} \sqrt{n_{\mathsf{I}}^2(t) + n_{\mathsf{Q}}^2(t)}, & \text{if 0 is sent,} \\ \sqrt{\left(A + n_{\mathsf{I}}(t)\right)^2 + n_{\mathsf{Q}}^2(t)}, & \text{if 1 is sent.} \end{cases}$$

# Noncoherent Demodulation of FSK



# DPSK: Differential PSK

- Impossible to demodulate PSK with an envelop detector since PSK signals have the same frequency and amplitude
- Demodulating PSK differentially, where phase reference is provided by a delayed version of the signal in the previous interval
- Essential to encode differentially: $b_n = b_{n-1} \times a_n$ where $a_n, b_n \in \pm 1$



# Table of Contents

- We know that if symbol $s_n$ has occurred, this corresponds to amount of information,

$$I(s_n) = \log_2 \frac{1}{p_n} = -\log_2 p_n \text{ bits of information}$$

- For random variable $S \in \mathcal{S}$, expected value of $I(S)$ over the source alphabet

$$\mathbb{E}\{I(S)\} = \sum_{n=1}^{N} p_n I(s_n) = -\sum_{n=1}^{N} p_n \log_2 p_n$$

- Source entropy: average amount of information per source symbol

$$H(S) = -\sum_{n=1}^{N} p_n \log_2 p_n$$

- Units: bits/symbol

- Capacity of a discrete memoryless channel is the maximum mutual information between the input and output, where the maximization is over all possible input probability distributions

$$C = \max_{p(x_0),\ldots,p(x_{J-1})} I(X;Y)$$

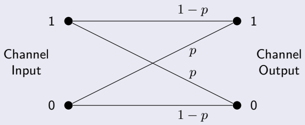
## Channel Coding Theorem

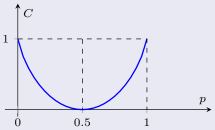
- If the transmission rate $R \leq C$, then there exists a coding scheme such that R bits per channel use can be transmitted over the channel with an arbitrarily small probability of error.
- Conversely, if $R > C$, error probability is always bounded above zero when the transmission rate is above the capacity.

# Binary Symmetric Channel and Additive White Gaussian Noise Channel

## Capacity of BSC

$$C = \max_{p(x_0),p(x_1)} I(X;Y) = 1 - h(p), \quad \text{where } h(p) = -p\log_2 p - (1-p)\log_2(1-p)$$




## Capacity of AWGN channel

$$C = B\log_2(1+\mathsf{SNR}) = B\log_2\left(1+\frac{P}{N_0 B}\right) \text{ bps}$$

- $B$: bandwidth of the channel
- $P$: average signal power at the receiver
- $N_0$: single-sided PSD of noise

# Table of Contents

## Source Coding Theorem

Given a discrete memoryless source of entropy $H(S)$, average codeword length $\bar{L}$ for any **uniquely decodable source coding scheme** is (lower) bounded by $H(S)$, that is,

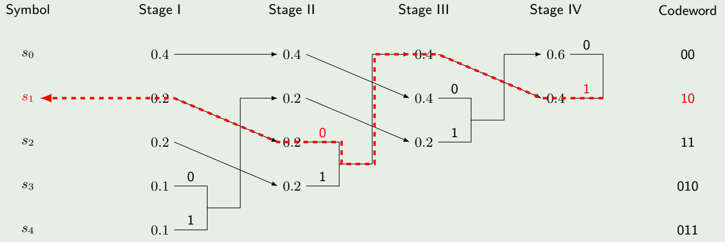
$$\bar{L} \geq H(S)$$

# Huffman Coding

- Basic Idea: choosing codeword lengths so that more-probable sequences have shorter codewords
- Code Construction:
  - Sort source symbols in order of decreasing probability
  - Take two smallest $p(x_i)$ and assign each a different bit (i.e., 0 or 1), then merge into a single symbol
  - Repeat until only one symbol remains
- Properties:
  - Huffman Coding (among other algorithms): uniquely decodable with average coding length satisfying $H(S) \leq \bar{L} < H(S) + 1$
  - The shortest average codeword length
  - Easy to implement this algorithm: used in JPEG, MP3, . . .

# Example

## Huffman coding




Read diagram backwards for codewords

- Average codeword length:

$$\bar{L} = (2 \times 0.4) + (2 \times 0.2) + (2 \times 0.2) + (3 \times 0.1) + (3 \times 0.1) = 2.2$$

- Huffman code is not unique (you can reorder equal probabilities)
- Huffman code is the most efficient prefix code
- More than the entropy $H(S) = 2.12$ bits/symbol

# Table of Contents

# Generator Matrix

- To construct a linear block code we define a matrix, the generator matrix $\boldsymbol{G}$, that converts blocks of source symbols into longer blocks corresponding to codewords
- $\boldsymbol{G}$ is a $k \times n$ matrix ($k$ rows, $n$ columns) that takes a source block $\boldsymbol{u}$ (a binary vector of length $k$), to a codeword $\boldsymbol{x}$ (a binary vector of length $n$)

$$\boldsymbol{x} = \boldsymbol{u} \cdot \boldsymbol{G}$$

$$\begin{array}{c}\text{Message}\\ \text{vector } \boldsymbol{u}\end{array} \longrightarrow \boxed{\begin{array}{c}\text{Generator}\\ \text{matrix } \boldsymbol{G}\end{array}} \longrightarrow \begin{array}{c}\text{Code}\\ \text{vector } \boldsymbol{x}\end{array}$$

$$\boldsymbol{u} = \begin{bmatrix} 1 & 0 \end{bmatrix}, \quad \boldsymbol{G} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}, \quad \boldsymbol{x} = \boldsymbol{u}\boldsymbol{G} = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \end{bmatrix}$$

- Linearity: Summation of two codewords is another codeword.
  - If $\boldsymbol{x}_1 = \boldsymbol{u}_1\boldsymbol{G}$, $\boldsymbol{x}_2 = \boldsymbol{u}_2\boldsymbol{G}$ are two codewords, then $\boldsymbol{x}_1 + \boldsymbol{x}_2 = \boldsymbol{u}_1\boldsymbol{G} + \boldsymbol{u}_2\boldsymbol{G} = (\boldsymbol{u}_1 + \boldsymbol{u}_2)\boldsymbol{G}$ is another codeword!

- Hamming Distance between two binary vectors, $\boldsymbol{a}$ and $\boldsymbol{b}$, is written as $d_{\mathsf{H}}(\boldsymbol{a}, \boldsymbol{b})$, and is equal to the Hamming weight of their (Boolean) sum

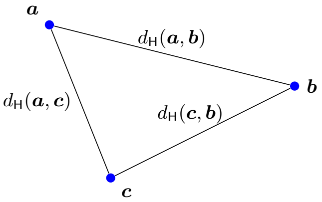
  $$d_{\mathsf{H}}(\boldsymbol{a}, \boldsymbol{b}) = w_{\mathsf{H}}(\boldsymbol{a} + \boldsymbol{b})$$

  For example, 01110011 and 10001011 have a Hamming distance of

  $$d_{\mathsf{H}}(01110011, 10001011) = w_{\mathsf{H}}(01110011 + 10001011) = w_{\mathsf{H}}(11111000) = 5$$

- Triangle inequality

  $$d_{\mathsf{H}}(\boldsymbol{a}, \boldsymbol{b}) \leq d_{\mathsf{H}}(\boldsymbol{a}, \boldsymbol{c}) + d_{\mathsf{H}}(\boldsymbol{c}, \boldsymbol{b})$$



# Error Detection and Correction

- To determine the number of errors a particular code can *detect* and *correct*, we look at the minimum Hamming distance between any two codewords.
- From linearity, the zero vector must be a codeword. The minimum Hamming distance of a code is the same as minimum weight of non-codewords.
- We define the minimum distance between any two codewords to be

$$d_{\min} = \min_{\substack{\boldsymbol{a},\boldsymbol{b}\in\mathcal{C}\\ a\neq b}} d_{\mathsf{H}}(\boldsymbol{a},\boldsymbol{b}) = \min_{\substack{\boldsymbol{a},\boldsymbol{b}\in\mathcal{C}\\ a\neq b}} d_{\mathsf{H}}(0,\boldsymbol{a}+\boldsymbol{b}) = \min_{\boldsymbol{c}\in\mathcal{C},\boldsymbol{c}\neq\boldsymbol{0}} w_{\mathsf{H}}(\boldsymbol{c})$$

where $\mathcal{C}$ is the set of codewords.