Digital Communications

Introduction

N. Lebedev

CPE Lyon
lebedev@cpe.fr

Plan I

1. Motivation
2. Blocks of digital communication system
3. Digital bandwidth measures

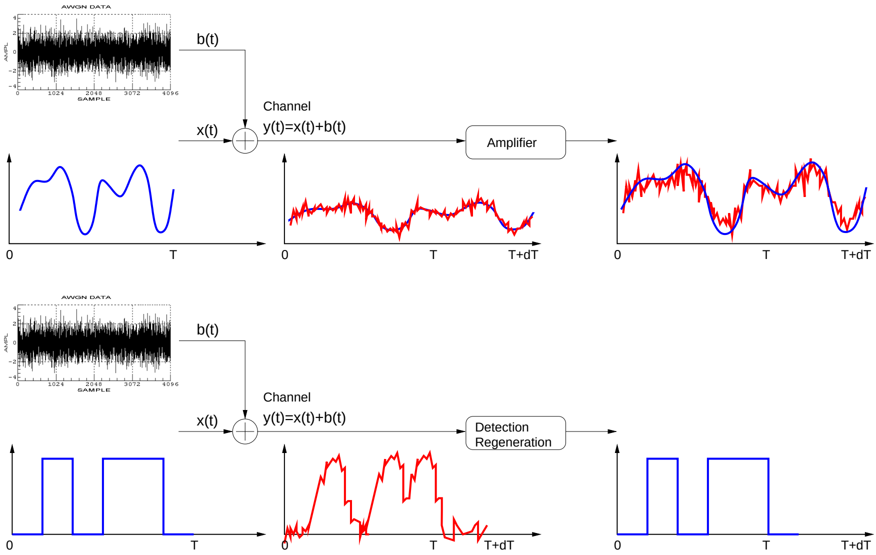
Motivation

Reasons for digital format and communication

- Regeneration: recover (detect) and repeat digital pulses.
- Digital: finite signal states $\Rightarrow$ less sensitive to distortions.
- Storage.
- Reproducibility.
- Coding—almost achieving channel capacity.
- Encryption—algebraic methods, in ASIC.
- Algorithms implementation—high-level language (e.g. Python) prototyping, then FPGA programmable circuits.

- DIGITAL is the UNIFIED FORMAT suitable for ANY SOURCE or CHANNEL or STORAGE.

Digital vs analog: a short comparison



Digital vs analog: a short comparison

Digital

- Finite set of M discrete msg's (waveforms)
- Easily identified (detected) if distorted $\Rightarrow$ recovery, amplification
- *Detect=identify the waveform* from the noisy channel output
- Processing regardless of appli "it's just bits"!
- Low $Prob(err)$ with error-correcting codes
- Simple electrical components:
  - Integration, VLSI circuits, μProc
  - Cheap
- Signal processing, algorithmic complexity
- Synchronization burden (VCO circuits)
- Performance metric: $Prob(detection\ error)$.

Analog

- Messages are continuous (infinite states)
- Distortion and signal are equally amplified
- *Detect=recover and restore* the transmitted signal with high precision.
- May require application-specific processing
- ECC unavailable for analog
- Complex and expensive circuits
- Easy synchro
- Performance metrics: measure of analog waveform distortion.

Applications

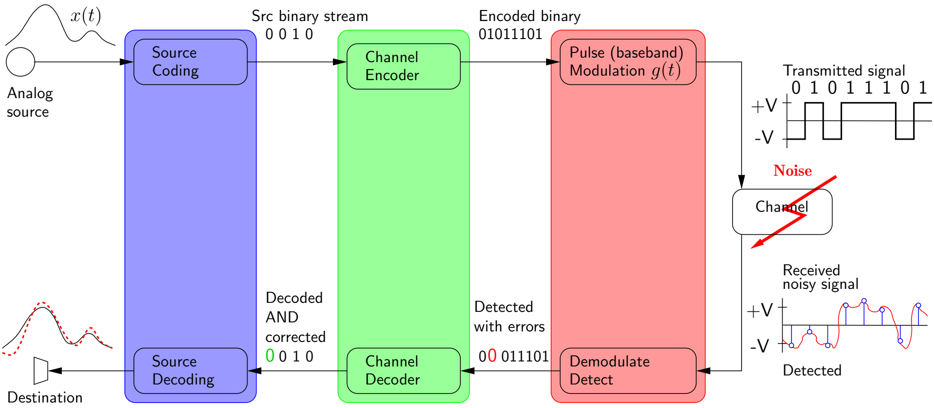
Communications:

- Digital telephony: switching appeared first (!), transmission came later.
- Computer and data networks.
- Multimedia: unified multiple digitized flows over the same channels (VoIP).
- Satellite communications: better detection of very weak signals.
- Industrial and embedded systems: automotive, aerospace, rail, production lines.

Storage:

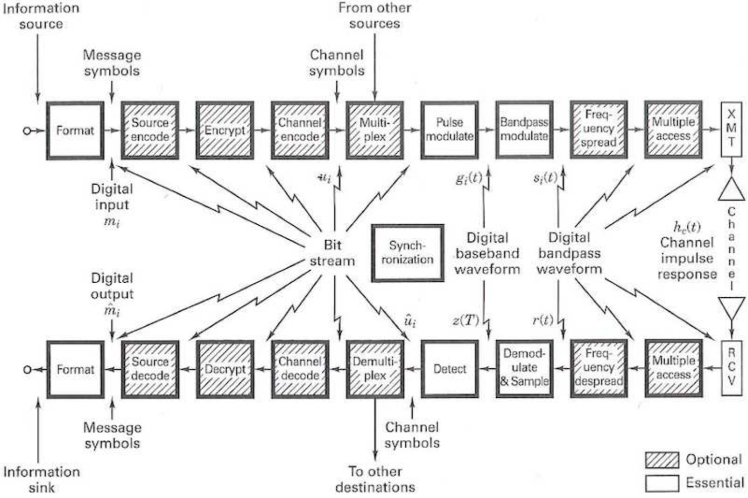
- Digital images, many formats: size vs precision.
- Ditigal videos: (size) vs (rate) vs (precision) vs (error rate); multiple criteria.
- Data: CD, hard drives, magnetic tapes, flash.

Minimal digital communication chain



Building blocks of digital communication system [Sklar]

- Digital transmission of discrete **data** over a **channel**.
- Source: analog or digital
- Channel: wired (electric), wireless (radio), optical, underwater, deep space;storage



Building blocks of digital communication system [Sklar]

- Formatting $\approx$ ADC, into binary $\{0, 1\}$, and possibly grouping.
  - Formatting = ADC, without compression
  - Grouping by k bits to form *digital messages or message symbols or words* $m_i, i = 1, \ldots, M = 2^k$.
  - ... (Optional: source coding, encryption, channel coding, framing and mux).
- Pulse baseband modulation: create physical signals $g_i(t), i = 1, \ldots, M$.
  - $M = 2$, binary messages $m_i \in \{0, 1\}$ $\rightarrow$ PCM binary waveform (*line codes*).
  - Non-binary messages m_i $\rightarrow$ PAM, PPM, PWM. Baseband waveform $g_i(t), i = 1, \ldots, M$.

  Objectives of the modulation:

  1. Mapping : M messages $\rightarrow$ baseband waveforms $g_i(t)$, binary or M-ary.
  2. Filtering (pulse shaping) $\rightarrow$ time waveforms longer than T_{symb} to limit the spectral content.
  3. ...Optional: if RF, frequency translation , or bandpass modulation.
     Appropriate waveform is to be used $s_i(t)$ ($g_i(t)$ translated to $f \gg$ freq$[g_i(t)]$)
- ($\longleftarrow$ Channel—not a part of the comm chain ! $\longrightarrow$).
  Received noisy signal: $r(t) = s_i(t) * h(t) + n(t)$.
- Demodulator: recover and convert the waveform into a baseband pulse
  - Frequency down conversion, and pulse shaping (MF—Match Filtering) to restore the waveform from $r(t)$ into a baseband $\tilde{z}(t)$.
  - May involve equalization if the channel delay spread $\tau_{max} \gtrapprox T_{symb}$ (echo).
  - *Sampling* to obtain $\tilde{z}(nT)$ (at the peak of MF output) to prepare the detection.
- Detection: decision making $\tilde{z}(nT) \rightarrow \hat{m}_i$.
- Synchronization is essential, drives all other blocks

Formatting

A/D—Analog-to-Digital Conversion

Analog source:

- ADC: physical signal is formatted into $\{0, 1\}$ bit stream.
  - 3 steps: Sampling—Quantization—Encoding (ex: PCM)
  - Techniques: formatting (just ADC) or source coding (ADC+compression)

Discrete source:

- Symbols take values from a finite alphabet.
- Encoded or mapped one-by-one into a bit stream $\{0, 1\}$.
  - Ex: natural language [A..Z]. i) char-by-char encoded into ASCII, UTF-8, Morse; ii) then into binary $[A]_{ASCII} \rightarrow [01100001] = [0\text{x}61]$.

Digital (binary) source and transmission:

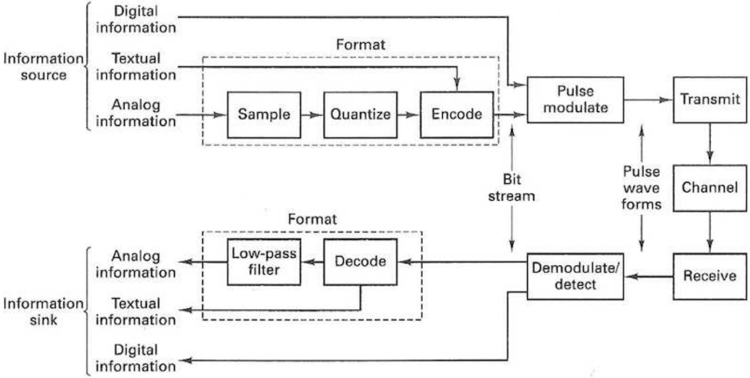
- Grouping by k bits to form digital messages (symbols) $m_i, i = 1, ..., M$ from alphabet of max $2^k = M$ such messages.
  - $k = 1, M = 2$, binary. Ex: $m_1 = [0],\ m_{M=2} = [1]$
  - $k = 2, M = 4$, quaternary. Ex: $m_1 = [00],\ m_2 = [01], m_3 = [10], m_{M=4} = [11]$
  - ...
  - $k = 8 = 1 \text{ octet}, M = 256$.

Digital communication system—1

Baseband [Sklar]

Stages

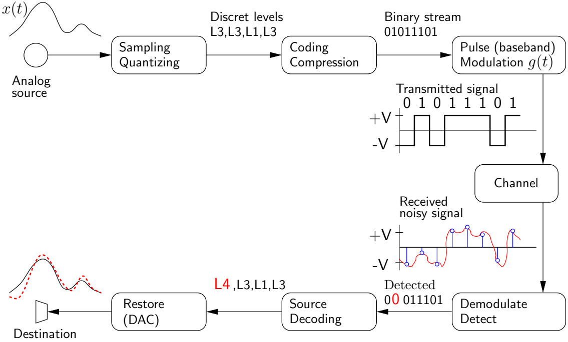
- Formatting—ADC (Analog-to-Digital Conversion)



Example: analog source coding

Source coding = Formatting + Compression

- Characterize the source.
  - Measure: minimal information content, entropy $H(X)$.
- Extract this information
  - Source coding (compression) technique: reduce the redundancy in the message.



- Ref : "Quadratic detection" in TSA module, 3ETI.

Compression

Consider an analog signal quantified to only 4 levels: L_1, L_2, L_3, L_4.
Question 1: How to encode ?

- Proposal 1 (Code 1): simple binary PCM

Symbols	L_1	L_2	L_3	L_4
PCM code	11	10	01	00

  - Quantizer output symbols and binary PCM encoded stream: $L_2, L_3, L_2, L_3, L_3, L_2, L_4, L_3, L_2, L_2, L_1, L_2, ...$
  - 10, 01, 10, 01, 01, 10, 00, 01, 10, 10, 11, 10, = 24 bits

- Proposal 2 (Code 2):

Symbols	L_1	L_2	L_3	L_4
Proba	0.1	0.4	0.4	0.1
Code	101	0	11	100

  - Same quantizer output symbols and Code 2 binary stream $L_2, L_3, L_2, L_3, L_3, L_2, L_4, L_3, L_2, L_2, L_1, L_2, ...$
  - 0, 11, 0, 11, 11, 0, 100, 11, 0, 0, 101, 0, = 20 bits

Question 2: What is the limit of source coding ?
Question 3: Is there the best code = shortest average codewords ?

Types of compression techniques

Many source codes exist, for different source types. Often:

- Lossy compression used for speech: based on psycho-acoustic models.
- Lossless coding (entropy coding) for data: Huffman.

Example : CD quality audio, uncompressed

- Analog signal is sampled at 44.1 [kHz] $= 44100$ [samples/s]
- **Stereo 2 channels**, each sample is encoded onto 16 [bits].
- Source binary rate is: $R = 44100 \times 16 \times 2 \approx 1.41$ [Mbit/s]

Example : mp3 storage format. Compression standard MPEG1 Layer 3

- Many rates, 128 [kbit/s] is common. Compression ratio CD/mp3 $\approx$ 11:1 !

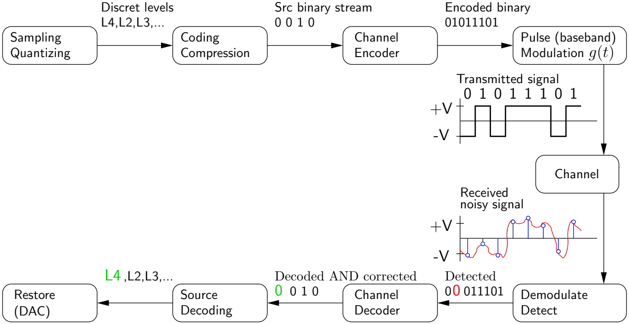
Channel coding

Error Correcting Coding (ECC)

- Protect against channel noise, fading, jamming
- Detect and (possibly) correct errors
- 2 categories:
  - Waveform coding: use signals that improve detection
  - Structured: ARQ, FEC: block, convolutional, turbo-codes

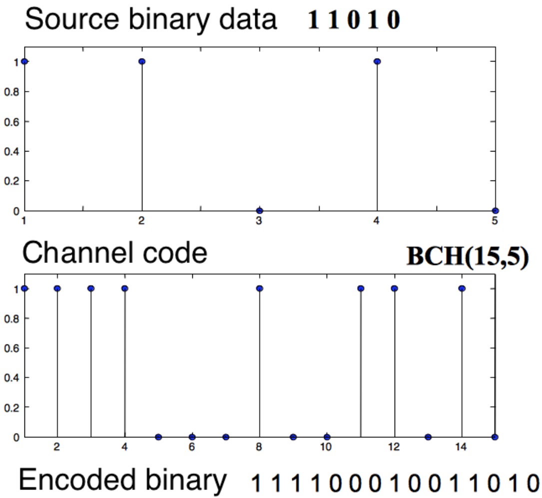
Multiple tradeoffs

$Prob(err)$ vs (SNR) vs (bandwidth) vs (complexity)



Channel coding

Example : Bose, Ray-Chaudhuri et Hocquenghem



Correction capability vs Redundancy vs Bandwidth vs Rate

Modulation

Adaptation of signal parameters to the channel

Baseband: DC (0 Hz) to several MHz

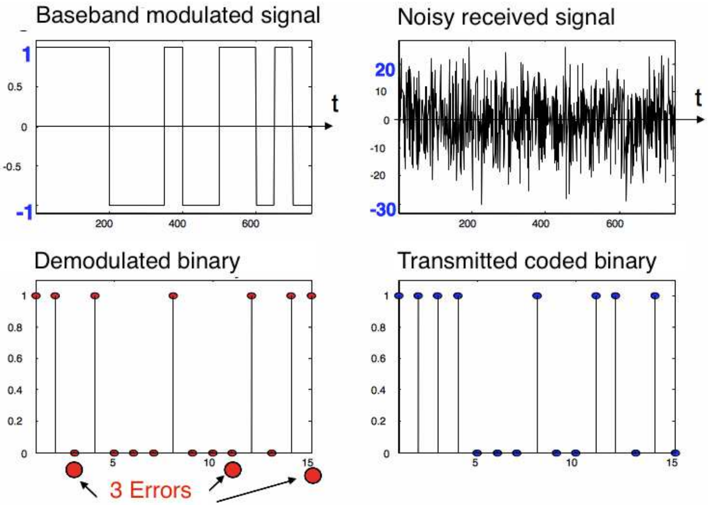
- Mapping **one-to-one** to transform each digital message into a **waveform**: $m_i \rightarrow g_i(t)$ of symbol duration T_s (out of M-ary waveform alphabet).
- Filtering or pulse shaping to limit the signal spectrum.
- Change voltage levels from device internal V_{bus} to transmission line V_{tx}.
- Many waveforms with different parameters. 2 types:
  - Binary = line codes: NRZ, RZ, phase encoded (Manchester in Ethernet), multilevel (HDB3 in ISDN).
  - M-ary: PAM (Pulse Amplitude Modulation), PPM (Position), PWM (Width).

RF (passband): carrier frequency $f_c \gg$ MHz

- Specific waveform $s_i(t)$ is to be used:
  - $s_i(t)$ is obtained by translating $g_i(t)$ to high-freq carrier $f_c \gg$ freq$\{g_i(t)\}$
  - The resulting BW is doubled ! $2f_{max}$ centered at f_c.

Optimal receiver

- Steps: synchronization, demodulation, detection (decision).
- NB: after demodulation, errors are still present !



Demodulation

Demodulation: recover the waveform (or baseband pulse) from $r(t)$ to $\tilde{z}(t)$:

- Frequency down conversion: revert from RF f_c to baseband.
- Filtering:
  - LPF (lowpass) to remove the high freq harmonics and keep the baseband signal
  - Matched Filter (MF) for waveform restoration, pulse shaping
  - Optional: equalization if the channel delay spread $\tau_{max} \gtrapprox T_s$ (time dispersive channel).
- Sampling to obtain $\tilde{z}(nT)$ (estimate of $g(t)$) to prepare the detection.

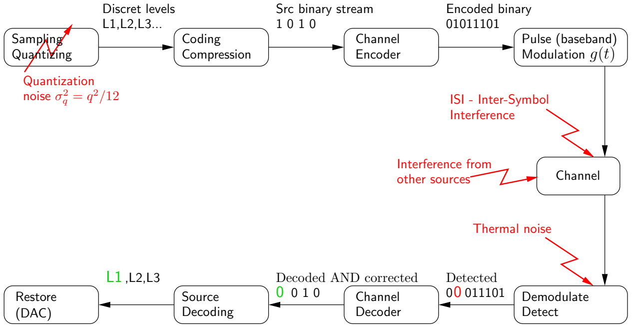
Detection: decision making $\tilde{z}(nT) \rightarrow \hat{u}_i$.

- Performance metric: $Prob(detection\ error) = \mathbb{E}\left\{Prob(\hat{u}_i \neq u_i)\right\}$

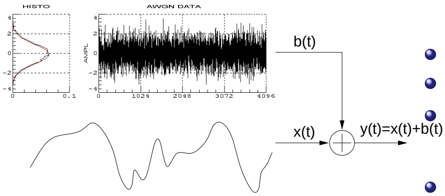
Noise and interference

The main effects are (but there are many others !)

- Irreducible quantization noise during formatting. Its variance is $\sigma_q^2 = q^2/12$.
- Interference due to co-channel or adjacent frequency channel transmissions.
- At the receiver, the thermal noise $n(t)$ is a random additive perturbation.



Thermal noise



- AWGN—Additive White Gaussian Noise: $\mathcal{N}(0, \sigma^2)$.
- Average PSD in bandpass: $\sigma^2 = N_0/2$
- PSD: $N_0 = 4 \cdot 10^{-21}$ Watt/Hz $= -174$ dBm/Hz, thermal noise threshold for 1 Hz at room temperature $T = 20^o$C $= 293$ K
- N, noise power within the band B

$$N = N_0 \cdot B = k \cdot T_k \cdot B, \text{ Watt},$$

where $k = 1,38 \cdot 10^{-23}$ $[\frac{W}{Hz \cdot K}]$, Boltzmann constant)

Example: GSM

useful bandwidth $B = 271$ KHz, at $T = 293$ K we have:

$$N \approx 1 \text{ femtoW} = 10^{-15} \text{ Watt} = -120 \text{ dBm}$$

Example: LoraWAN

useful bandwidth 500 kHz or 250 kHz, or 125 KHz at $T = 293$ K we have:

$$N_{500khz} =? \quad N_{250khz} =? \quad N_{125khz} =?$$

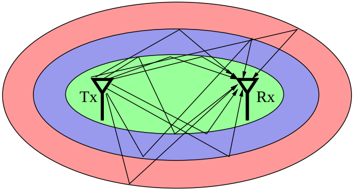
Fading and time dispersive channels

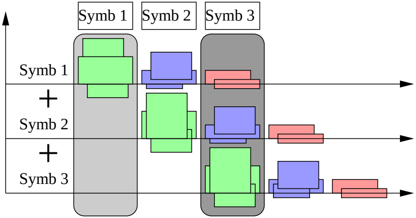
Distortion

Channel effects:

- Often, the channel is modeled as linear time variant system with impulse response $h_c(t)$
- At the receiver, the thermal noise $n(t)$ is a random additive perturbation
- The received signal $r(t)$ is a distorted version of the original waveform:

$$r_i(t) = s_i(t) * h_c(t) + n(t), \quad i = 1, 2, \ldots, M$$



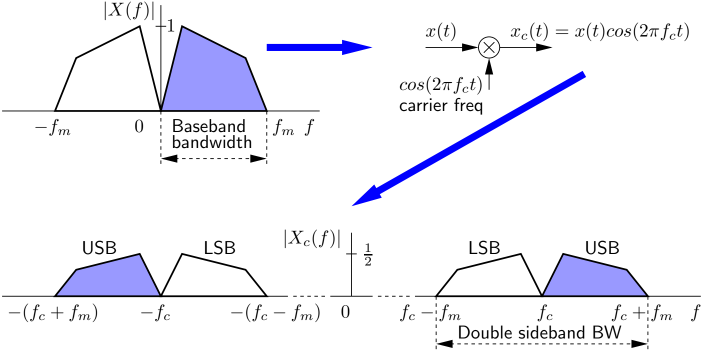


Reminder on data units

- Analog waveform $x(t)$: any physical signal. Ex: sensor output.
- Discret source: finite alphabet. Ex: $A, ..., Z$.
- Message: character, text: "A", "HELLO, WORLD !". Can be encoded into a binary word. Ex: ASCII, $[01100001] = [A]_{ASCII}$.
- Bit—BInary digiT, $\{0, 1\}$.
- Bit stream. $[...0010110010000101...]$. Information rate of a bit stream is $R_b = 1/T_b$ [bit/s]. Result of source formatting, either analog or digital source.
- Baseband digital message (symbol)—group of k bits: $m_i = [0110]_{k=4}$. Grouping is arbitrary, not related to the source, but to the transmission method.
- Baseband message is mapped to the baseband waveform $m_i \Rightarrow g_i(t)$, line coding.
- Symbol in RF or carrier modulation $s_i(t)$.
  - Channel (or symbol, or pulse) rate $R_s = R_b/k = 1/T_s$ [Baud] = [symbols/s].
- Digital waveform $g_i(t)$ or $s_i(t)$. Parameters:
  - Baseband: amplitude, width, duration of pulses.
  - Carrier: amplitude A, frequency f_c, phase φ for the carrier $A\sin(2\pi f_c t + \varphi)$.

Modulated signal bandwidth (BW)

- Simplest modulator is heterodyne (multiplier) by $f_c \gg f_{max}$ of the signal.
- Modulation: frequency translation AND mirroring AND scaling.
- **BW on carrier $= 2\times$ BW baseband ! $W_{DSB} \approx 2f_{max}$**; DSB-Dual Side Band.



The bandwidth dilemma [Sklar]—1

General form single-sided PSD $G_x(f)$ for a single modulated pulse $x_c(t)$, centred at f_c.

- The infinite duration ($\overset{FT}{\rightleftharpoons}$ bandlimited) signals are not physically realizable.
- **The finite duration signals can be realized, but $\overset{FT}{\leftrightharpoons}$ have infinite bandwidth.**

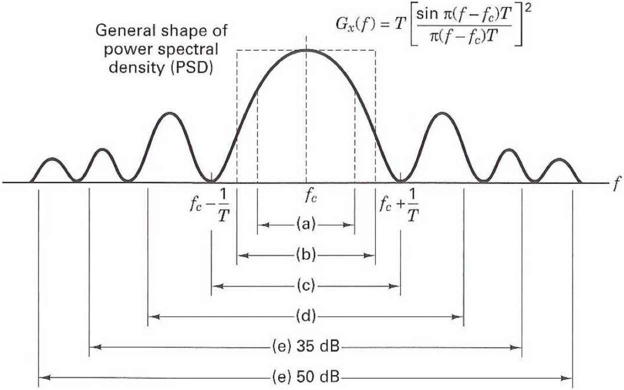


Figure 1.20 Bandwidth of digital data. (a) Half-power. (b) Noise equivalent. (c) Null to null. (d) 99% of power. (e) Bounded PSD (defines attenuation outside bandwidth) at 35 and 50 dB.

The bandwidth dilemma—2

Comments:

a) Half-power (-3 dB) wrt the maximum of PSD at f_c, that is $G_x(f_c)$.

b) Rectangular bandwidth (or *noise equivalent BW* of the spectral window).

$$W_N = \frac{P_x}{G_x(f_c)} = \frac{\int_{-\infty}^{\infty} G_x(f)df}{G_x(f_c)},$$

ratio of P_x, the total signal power of modulated pulse and $G_x(f_c)$, the maximum of the PSD at f_c. In other terms, W_N is the width of the rectangle with area P_x, and height $G_x(f_c)$.

c) **Most used, width of the main spectral lobe, containing most (≈ 90 %) power**, also named nul-to-nul bandwidth.

d) Fractional power BW: 99% of power 0.5 % left/right). Ex: for $\Pi_T(t)$...20 lobes !

e) Bounded PSD: BW beyond which the attenuation is > some level: 35 dB, 50 dB.

Summary on BW:

- Choice of BW definition depends on the application !

Fundamental limits—Nyquist sampling rate

Nyquist criterion for analog source considered as random process $X(t)$

- Stationary with autocorrelation $R_x(\tau)$ and PSD $G_x(f)$ (power signal).
- If $X(t)$ is bandlimited, $G_x(f) = 0$ for $|f| \geq f_m$, it can be represented as infinite sum of samples taken at $f_{samp} = 2f_m$. **Sampling theorem (Nyquist)**

$$X(t) = \sum_n X\left(\frac{n}{2f_m}\right) \times \text{sinc}\left(2f_m\left(t - \frac{n}{2f_m}\right)\right)$$

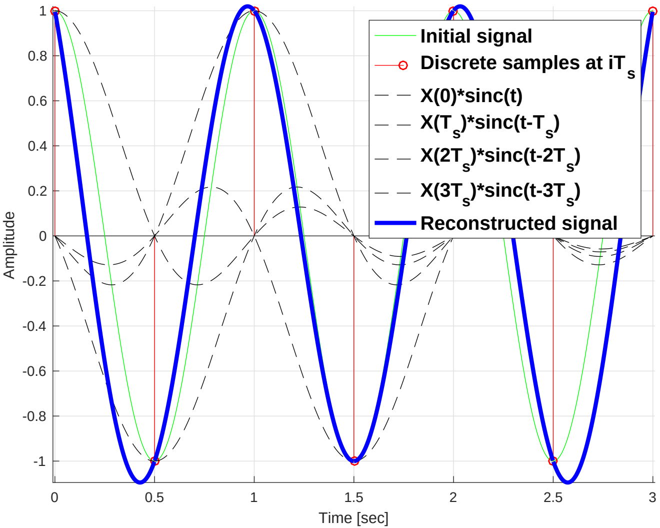
- The source output is described statistically by the joint pdf $p(X_i, ..., X_{i+n-1})$, of n unquantized samples $X(\frac{n}{2f_m}) \in \mathbb{R}_{[-V_{min}, V_{max}]}$.

DEMO

Formatting (ADC) = sampling + quantizing (+ PCM) $\Rightarrow$ equivalent discrete source.

- Remind, **formatting is lossy** due to quantization noise.
- Delivers symbols X out of finite alphabet of size L, nb of quantization levels.
- Followed by binary coding or compression

Nyquist rate: example of signal reconstruction



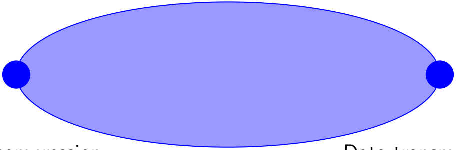
Fundamental limits—entropy H(X)

Minimal average source code length

The 2 fundamental questions of communication theory are:

1. What is the low limit of data compression ?
   Answer: Entropy $H(X)$, minimal information content of the source.
   $H(X) \leq R$ [bits/symbol], average number of bits to encode the source symbols.
2. What is the maximum transmission rate over a channel ?
   Answer: Channel Capacity C, so that $R_b \leq C$ [bits/s].

The communication theory methods and techniques lie in between.
Information and communication theory limits [Cover,Thomas]



Source—channel rates relationship

Source:

- S [samples/s] or [symbols/s] from the source (see CD example)
- $\overline{R}$ [bits/symbol] is the average source output rate after the formatting - compression. These bits can further be grouped, to form messages m_i.
- R [bits/s], bitrate
  is the source raw binary rate obtained by multiplying $\overline{R}$ by the sample (source symbol) rate S [samples/s] to obtain the $\overline{R} \times S = R$ [bits/s].
- Ex: PCM 8 bits/sample $\times$ 8000 samples/s $=$ 64000 bits/s

Channel:

- R_b is the channel transmission bit rate, the bits were eventually redundant channel coded. In this latter case, $R = R_c R_b$, where $0 < R_c \leq 1$ is the coding rate.
- Obviously, to make the communication possible, we should have

$$\underbrace{H(X) \leq R}_{source} \leq \underbrace{R_b \leq C}_{channel}$$

Course content

- Source coding and compression
- Channel error correcting codes (ECC)
- Carrier frequency modulation (digital bandpass signaling)
- Reception in Gaussian noise channel. Optimal detectors

Bibliography

1. FR:
   Didier Le Ruyet et Mylène Pischella, *Bases de communications numériques. Tomes 1,2*, ISTE editions, Sept. 2015.
   ISBN : 978-1-78406-093-0 (ebook) : 9,90 €
2. http://www.inference.org.uk/mackay/itila/
   David MacKay, *ITILA - Information Theory, Inference and Learning Algorithms*, University of Cambridge. Free book in .pdf.
   - http://www.inference.org.uk/mackay/itprnn/Videos.shtml, Companion videos (English) to view or to download
   - Source code for some programs available.
   - e-Campus : follow the announcements, and reading list.

NB :

One of the best and widely recognised references—didactic, light but still rigorous maths, full of examples and applications, funny in some places...
Don't waste your time on the web looking for alternatives !