EBU5504: Networks and Protocols
(Week 3)

Dr. Zhijin (Judy) Qin

Lecturer

z.qin@qmul.ac.uk

Queen Mary
University of London

Recognize me

- **Work experience**
  - Lecturer (2018-now), Queen Mary University of London, UK
  - Lecturer (2017-2018), Lancaster University, UK
  - Postdoc (2016-2017), Imperial College London, UK
- **Research interest**
  - Machine learning and compressive sensing for wireless signal processing
  - Internet of Things for smart cities

Course contents and schedule

The main topics covered by this course:

- Week 1&2:

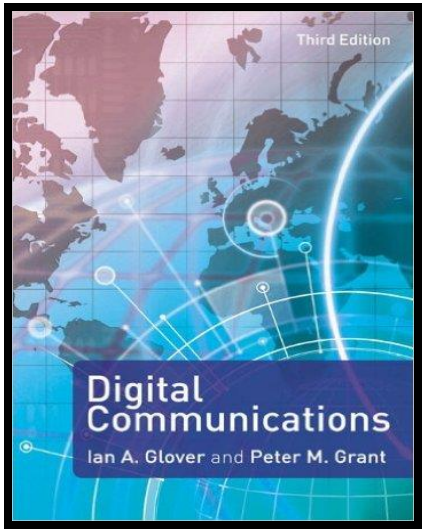
  Protocol layering, data link layer, MAC schemes for WSN and IoT.

- Week 3&4:

  System introduction, digitization, source coding and channel coding

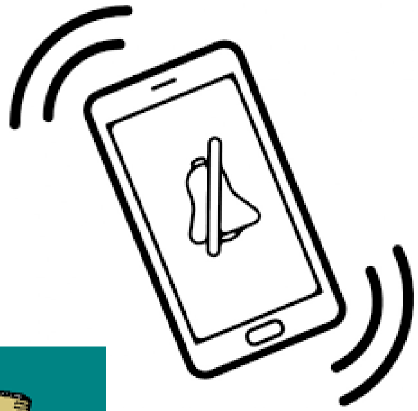
Recommended Text Book and References

- Majority of the content is available in ‘Glover and Grants’ book

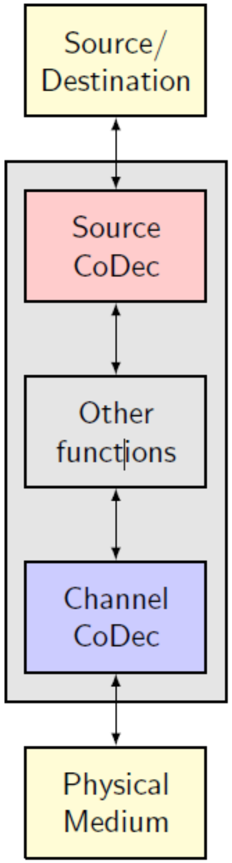


Manners

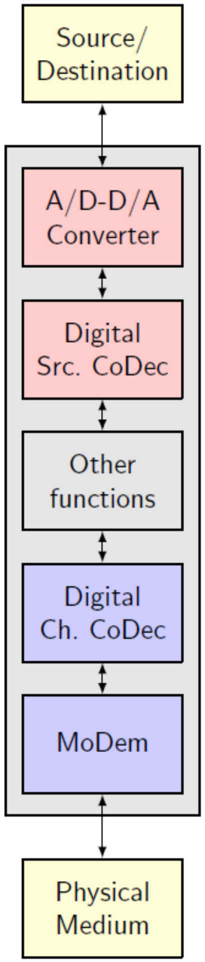
- No talking with each other
- But be **active** for interaction
- No smelly food
- Turn the phone to silent mode



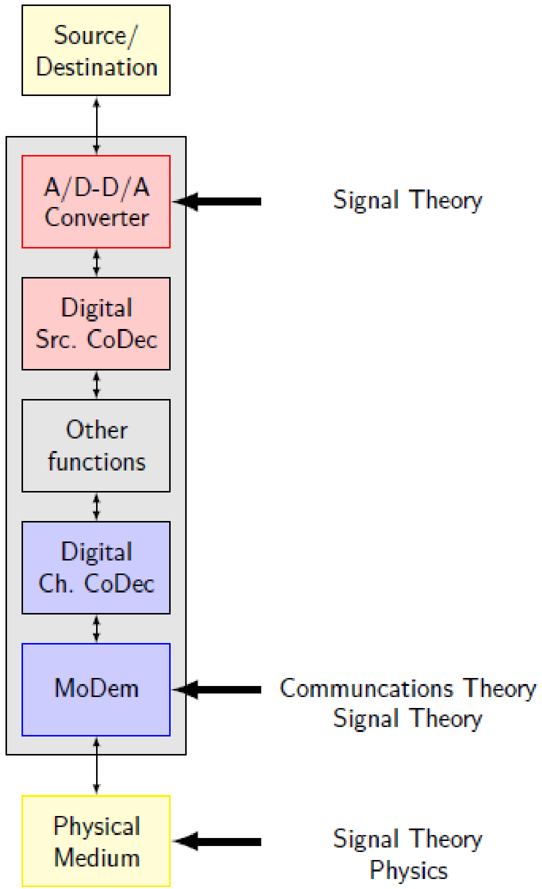
EBU5504 Networks and Protocols



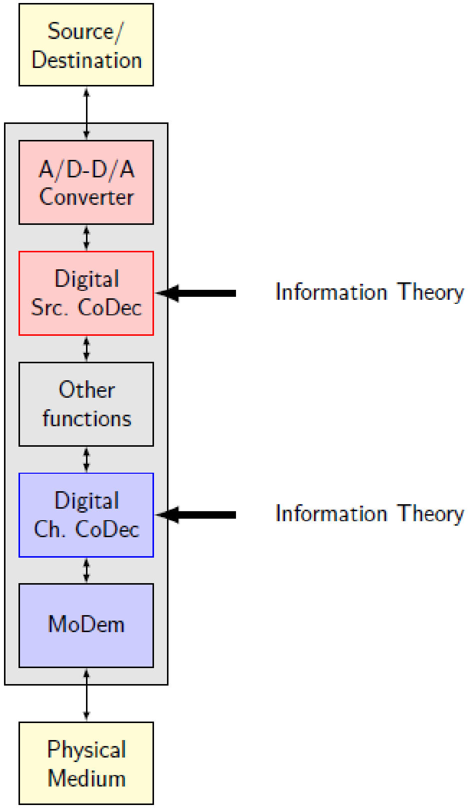
EBU5504 Networks and Protocols



Week 3: The Analog Interfaces



Week 4: Digital Source and Channel Coding



MODERN TELECOMMUNICATIONS

What is Communications?

- For us it is common to refer to various communication technologies, services and systems under the general term of Telecommunications.

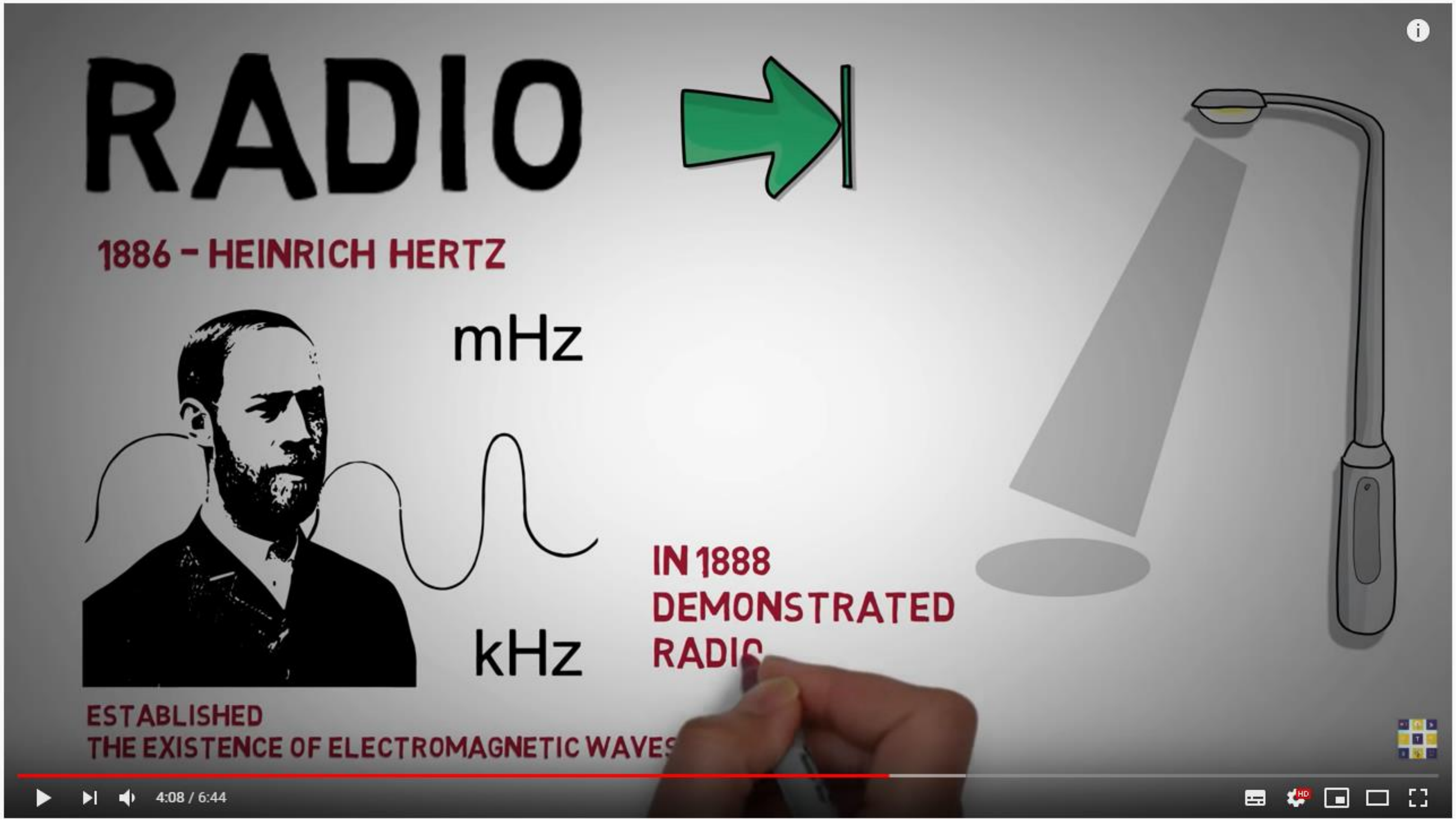
- The prefix "**Tele**" implies "at distance" and primarily allows us to distinguish between electrical communication and oral face-to-face communication.

- Define communications *"as the transmission of information between two distant points"*.

Transmission medium

- Copper
- Wireless
- Optical
- Increasing optical nearer the customer

History of Communications



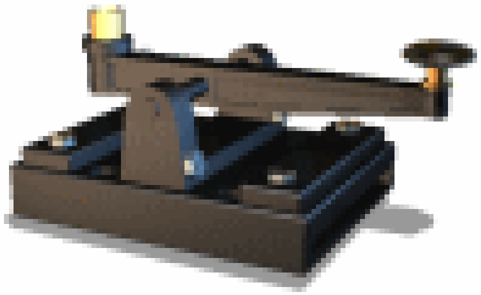
Modern Communications History

- Modern communications began in 1837 with the invention of the **Telegraph.**

- **Morse** thought if electricity could travel a short distance through wire, it could travel long distances through wire as well.

- Morse's idea was to string a wire

between two points, maybe miles apart.



Modern Communications History

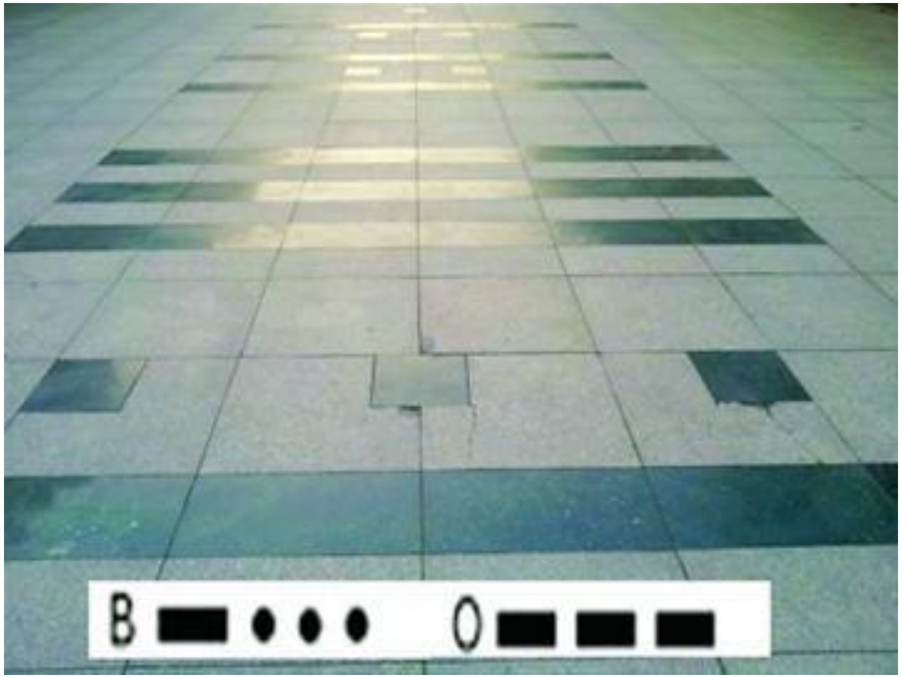
- **Morse** developed his telegraph idea and tested it by stringing 1,700 feet of wire around his room at New York University.

- It worked; his signals traveled from one end of the wire to the other.

- The famous Morse code was created.

Morse Code

- **A** • —
- **B** — • • •
- **E** •
- **F** • • — •
- **S** • • •
- Full Stop • — • — • —
- Comma — — • • — —



Modern Communications History

- By the early 1850s, overland telegraph lines spanned much of Europe, North America and the Middle East.
- In 1851 England was permanently connected to continental Europe by means of a cable laid between Dover and Calais.

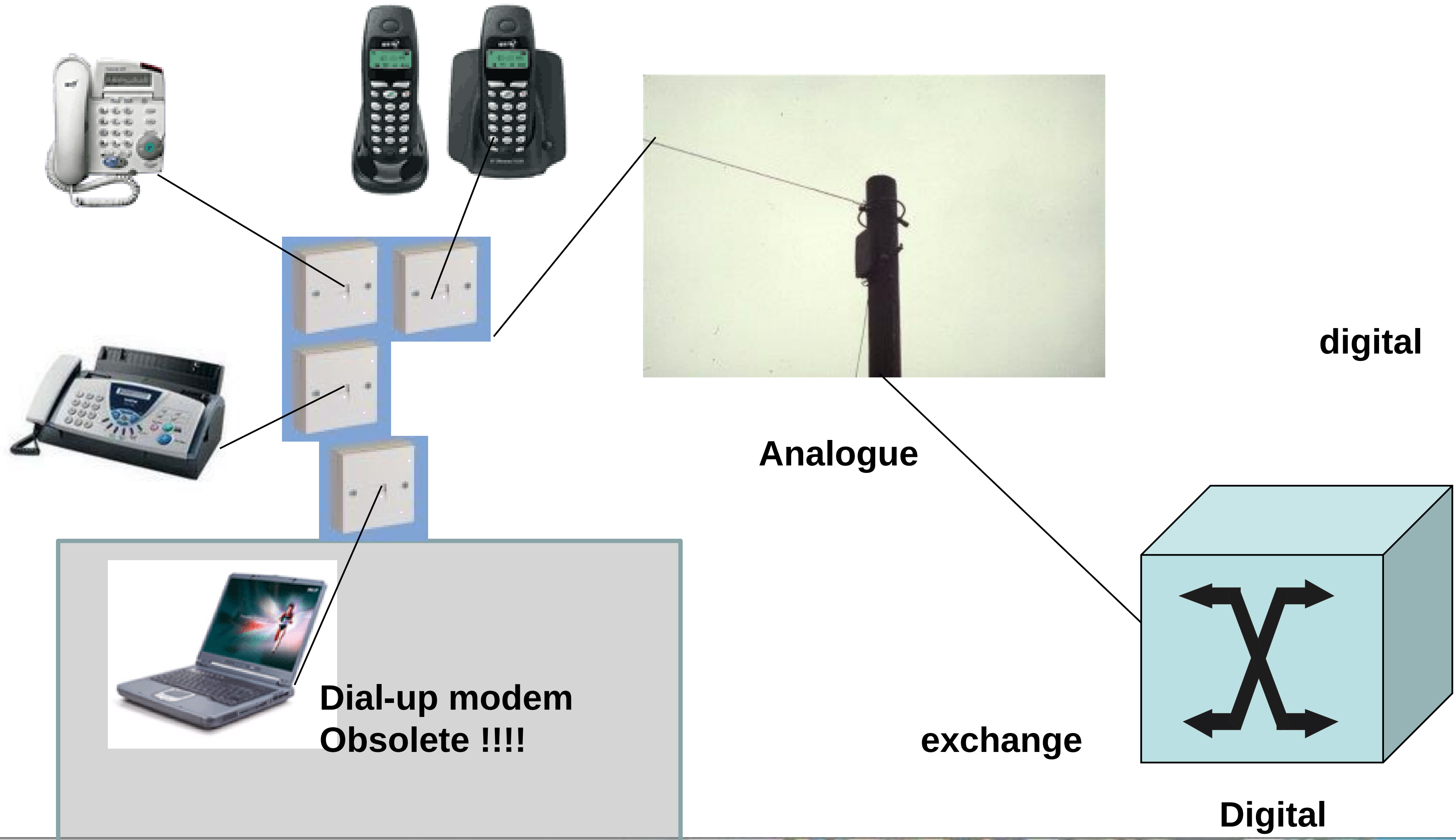
Modern Communications History

- Up to the late 19^{th} Century, all rapid long-distance depended upon the telegraph.

- In the 1870s, a rival technology was beginning its development that would again change the face of communication: the telephone

- The telegraph and telephone are both wire-based electrical systems.

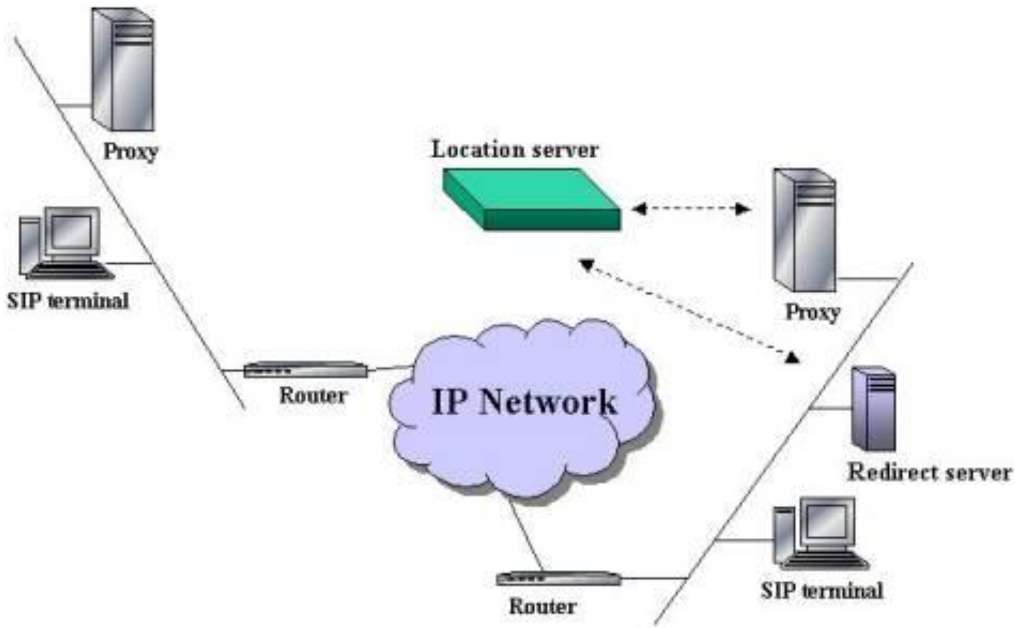
Legacy - wired communications using telephony



Voice is now going IP

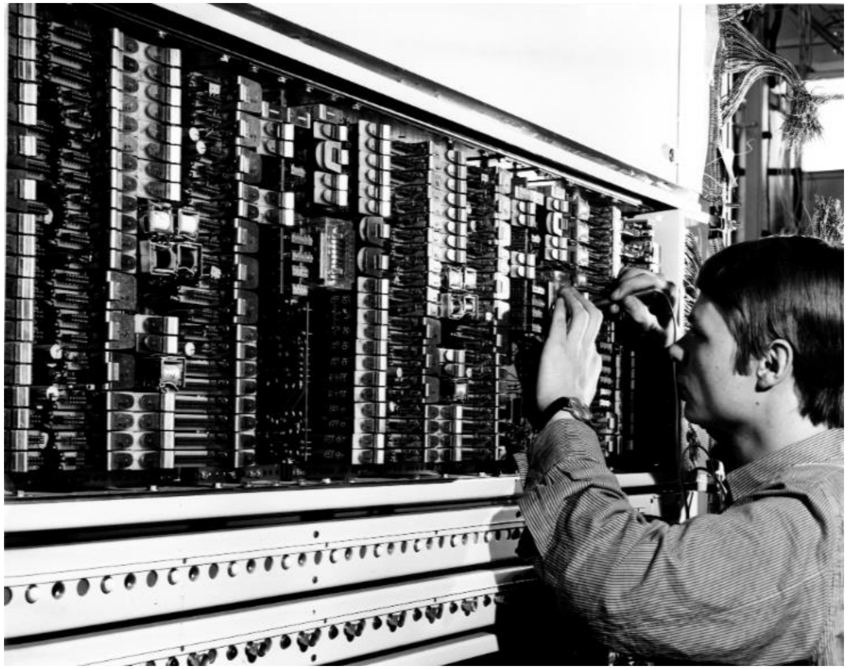
SIP based telephony

All telephony is going IP



Switching used to be manual

Then relays, then electronic – but specialised



Electromechanical exchange
picture courtesy of Nortel

1960s

Private electronic exchange
1983

Now just boxes of electronics – high volume

WLAN AP

IP router

Servers

All of these are just "boxes" of Electronics

IP switch

IP phone

The Public Switched Telephone Network

The PSTN has a dual analog-digital nature:

- Local loops are mainly **analog** (they are legacy increasingly structures), although digital (FTTH, Fiber To The Home).
- The core network (switching offices and trunks) is however **digital**. Local exchanges connected to the local loops digitize speech signals.

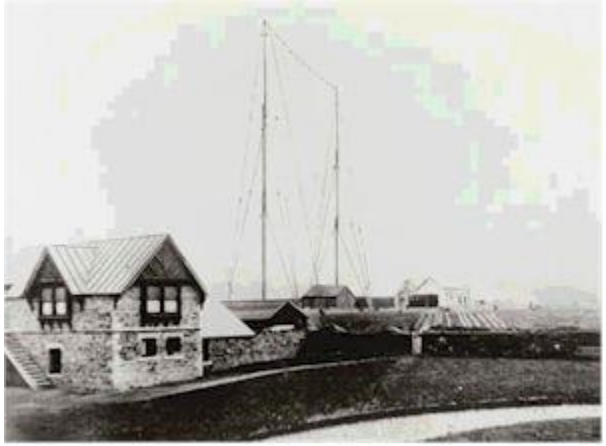
By contrast, due to the lack of legacy structures, mobile phone systems are **entirely digital.**

The Wireless Revolution

- Long-distance cable communications was expensive and caused signalling problems.

- In 1901, **Marconi** built a powerful wireless station at Cornwall (UK), in preparation for a transatlantic test.

- Many attempts were made to build suitable antennas some of which were destroyed by storms.



The Wireless Revolution

- On 12 December 1901, at Signal Hill in St John's (now part of Canada) Marconi received the first transatlantic radio signal from Cornwall, a distance of over 2,000 miles!

- The reception of transatlantic radio signals led to considerable advances in both science and technology.

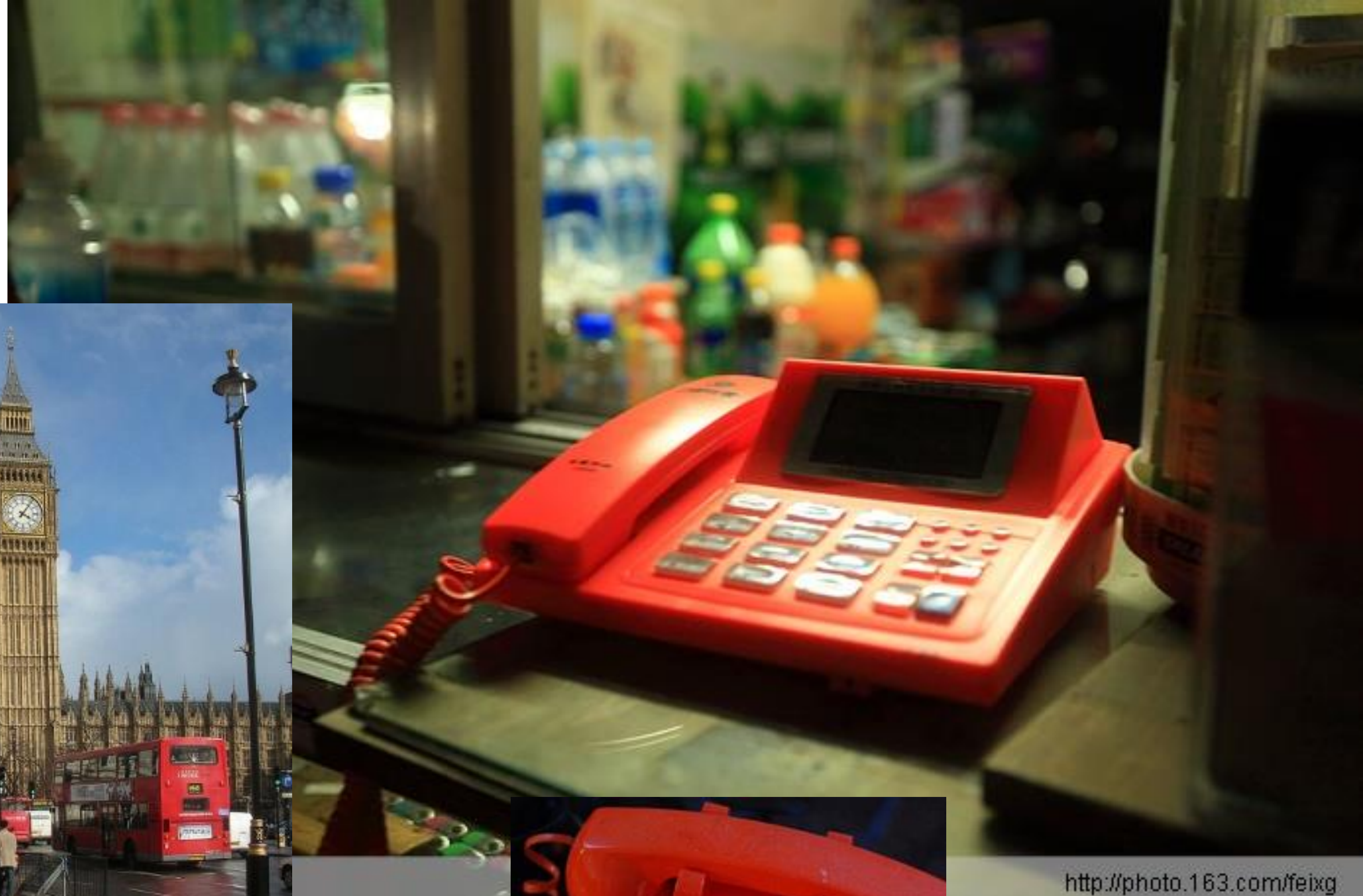
Types of Wireless Networks

- WPAN (Wireless Personal Area Network)
  - typically operates within about 30 feet
- WLAN (Wireless Local Area Network)
  - operates within 300 yards
- WMAN (Wireless Metropolitan Area Network)
  - operates within tens of miles
- WWAN (Wireless Wide Area Network)
  - operates over a large geographical area, mobile phone

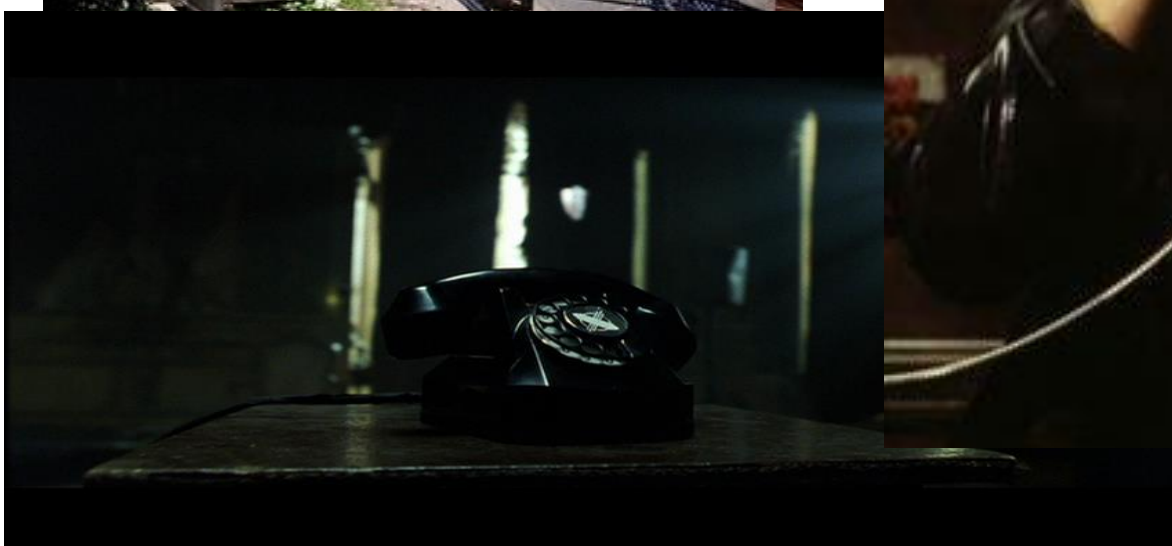
Why Wireless?

- No more cables
  - No cost for installing wires or rewiring
  - Wiring is infeasible or costly in some areas, e.g.. rural areas, old buildings…
- Mobility and convenience
  - Allows users to access services while moving: walking, in vehicles…
- Flexibility
  - Roaming allows connection any where and any time
- Scalability
  - Easier to expand network coverage compared to wired networks.

Touch-Tone dialling

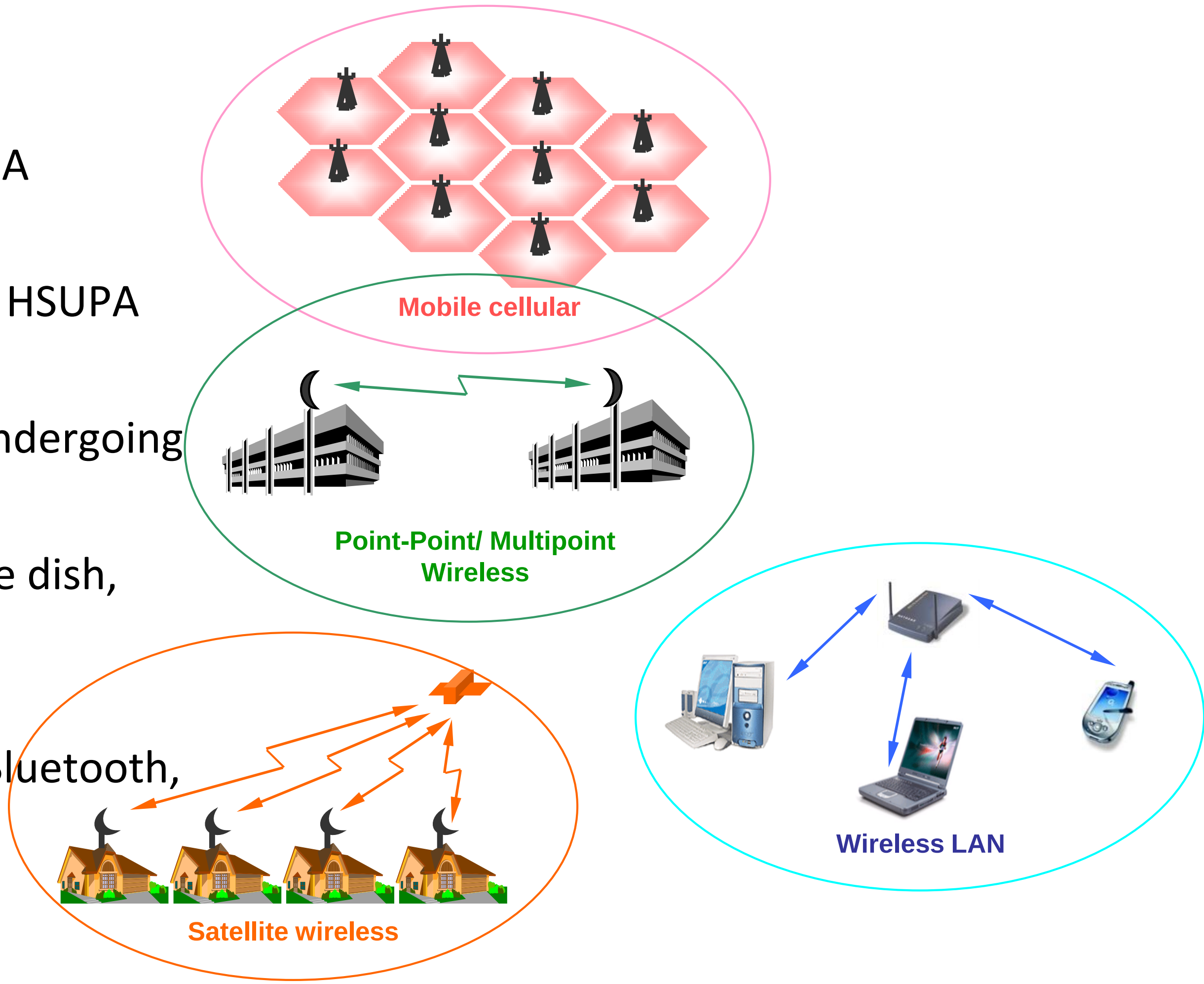


Pulse Phone

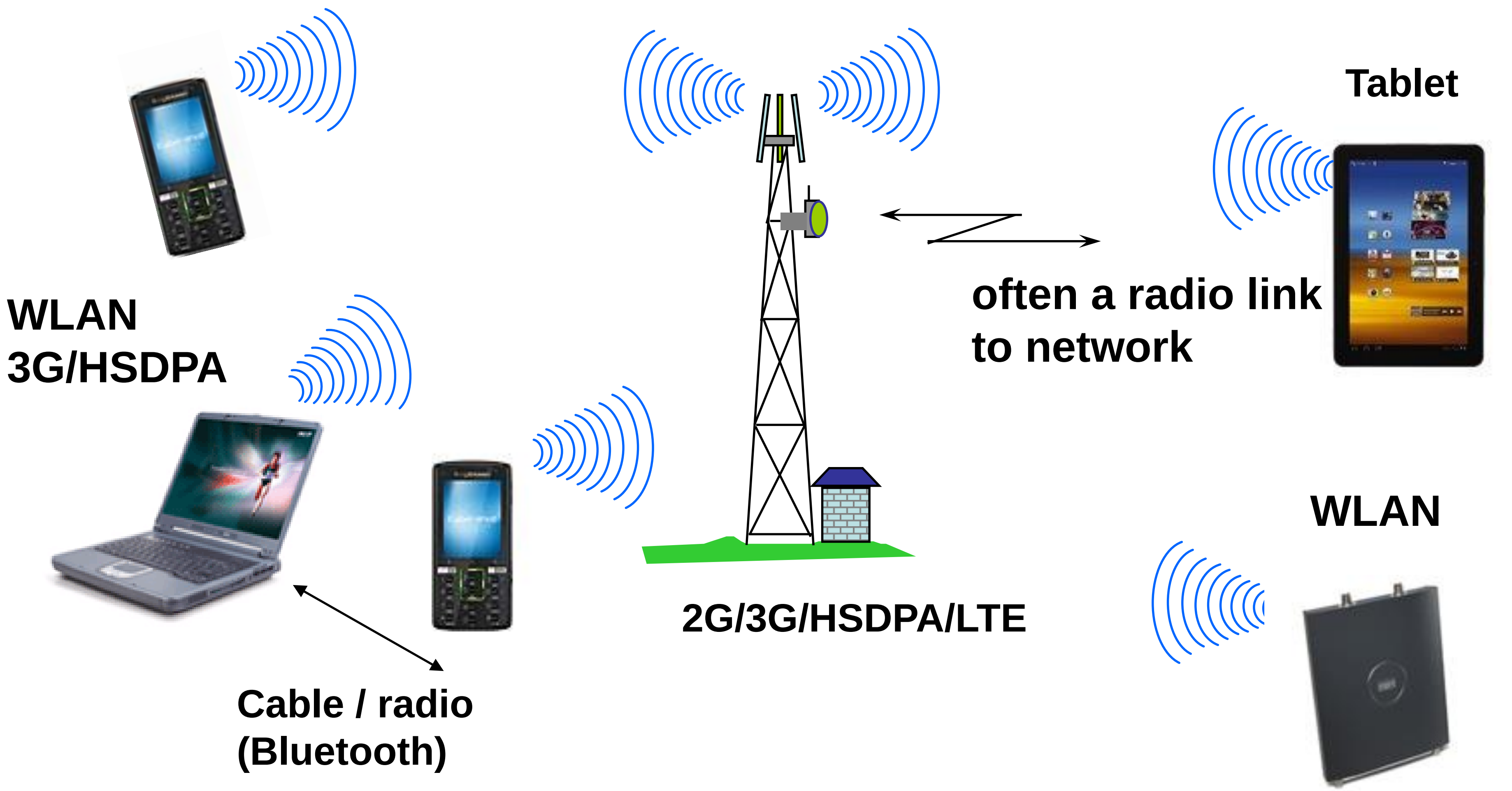


Emerging and existing wireless technology

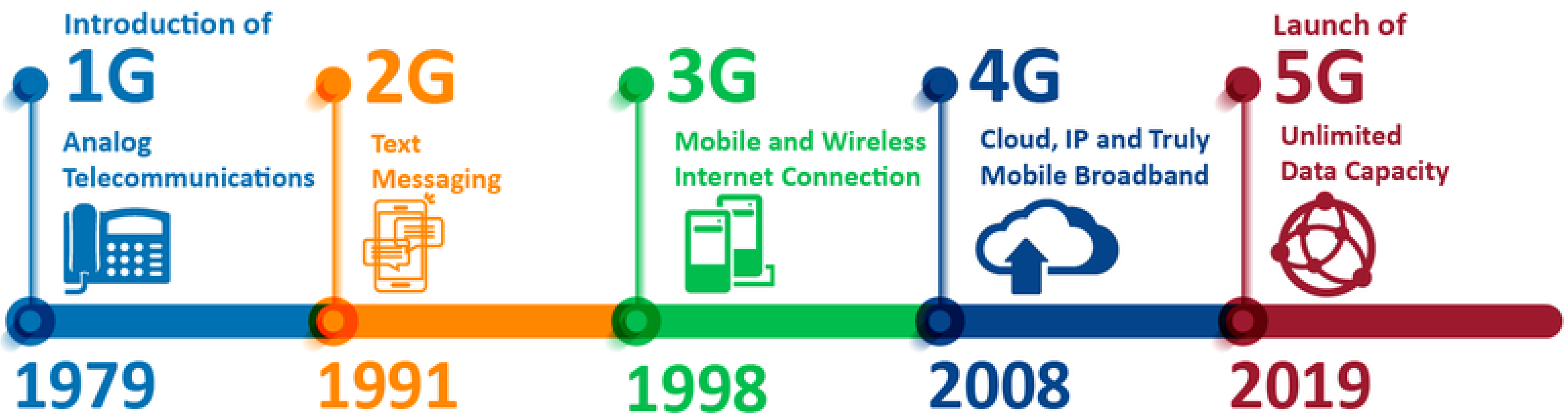
- Mobile Wireless:
  - 1G: Analog
  - 2G: GSM, TDMA, CDMA
  - 2.5G: EDGE, GPRS
  - 3G: W-CDMA, HSDPA, HSUPA
  - 4G: LTE
  - 5G: Standardization undergoing
- Fixed Wireless:
  - MMDS, LMDS, Satellite dish, Microwave
- Wireless LAN:
  - IEEE 802.11, Ad-hoc, Bluetooth,
- WiMax



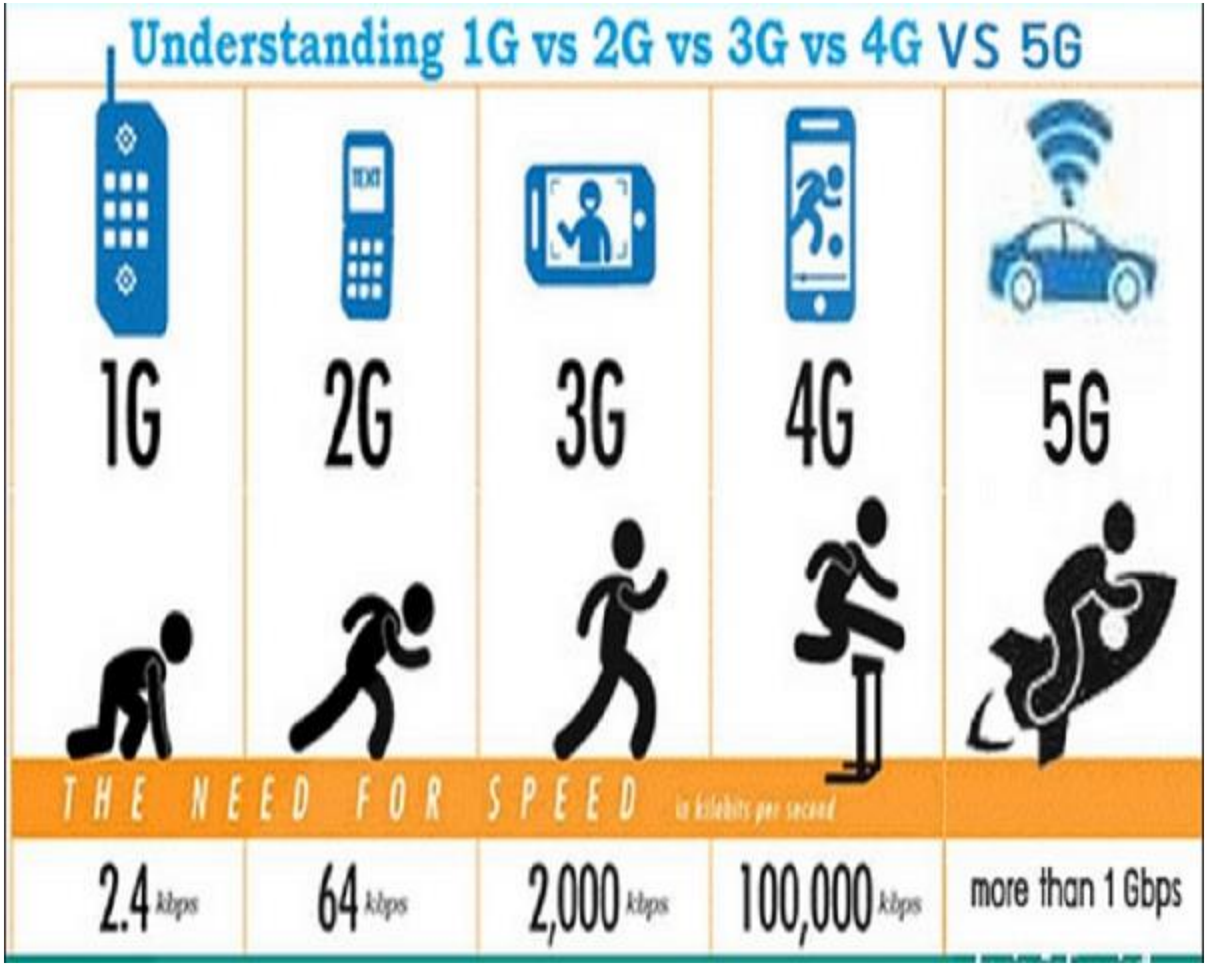
Mobile communications



Wireless Revolution

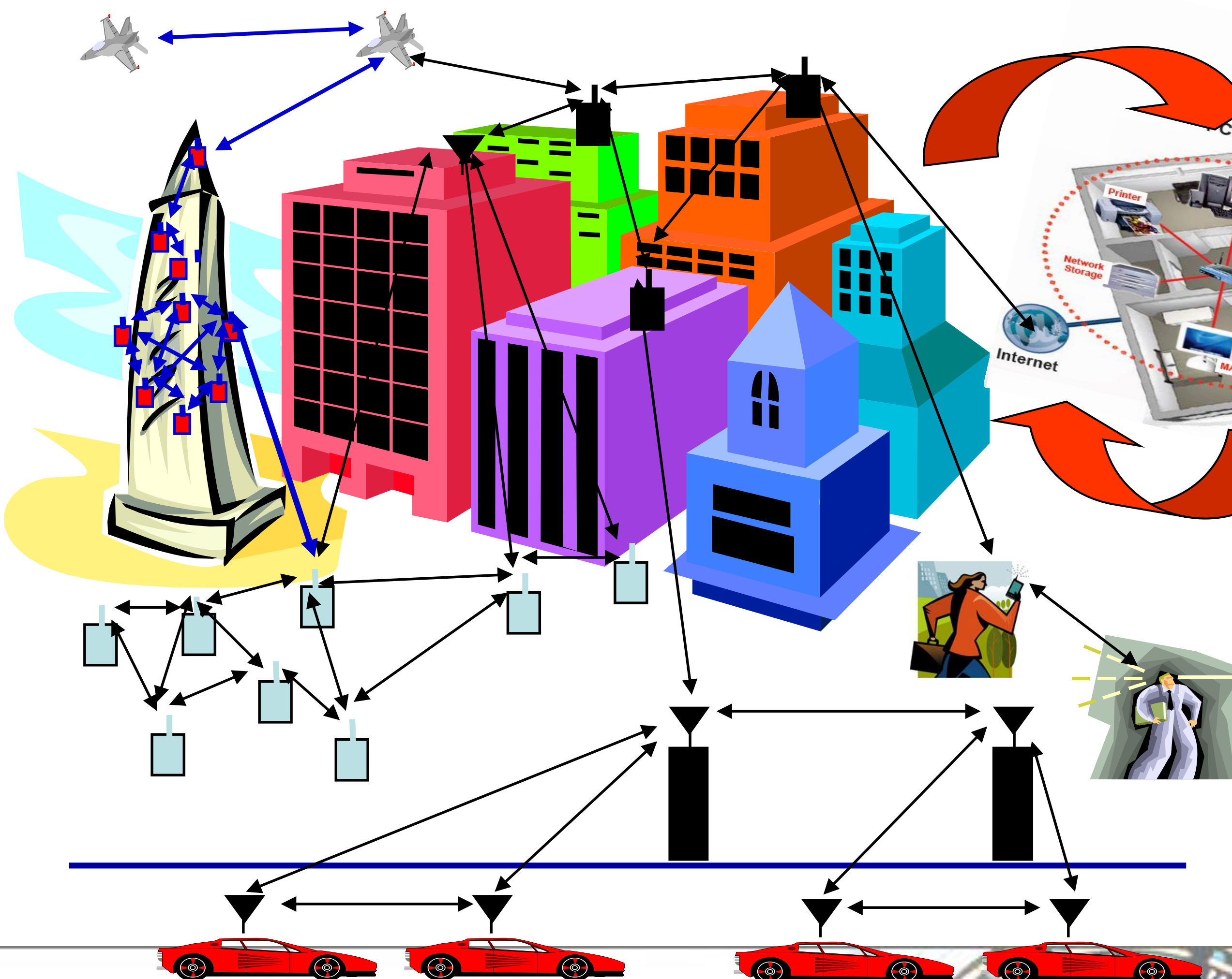


Wireless Revolution



Future Wireless Networks

Ubiquitous Communication Among People and Devices

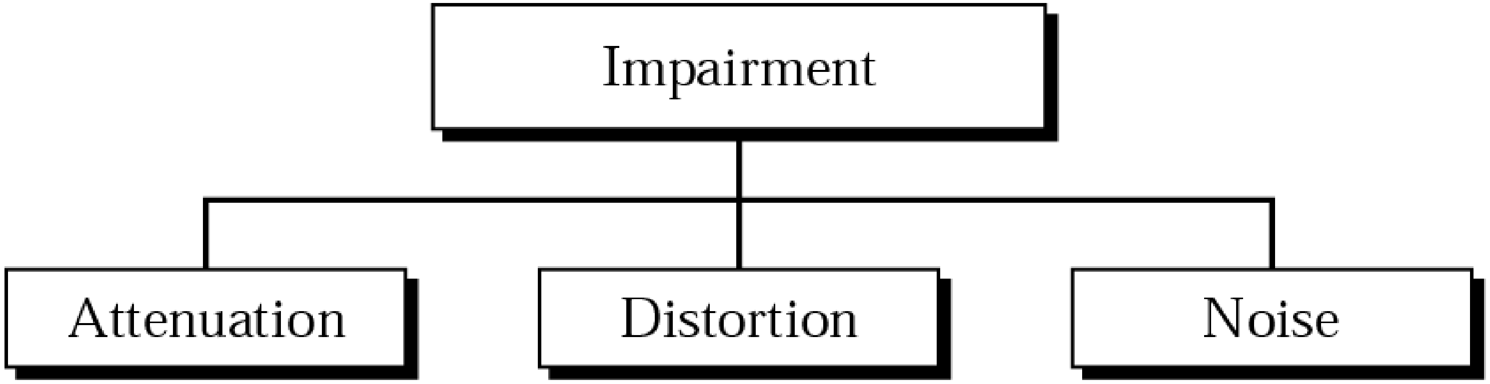


Next-generation Cellular
Wireless Internet Access
Wireless Multimedia
Sensor Networks
Smart Homes/Spaces
Automated Highways
In-Body Networks
All this and more …

Queen Mary
University of London

Transmission Impairment

- Signals that travel through transmission medium will always be *corrupted* by attenuation, distortion and noise.

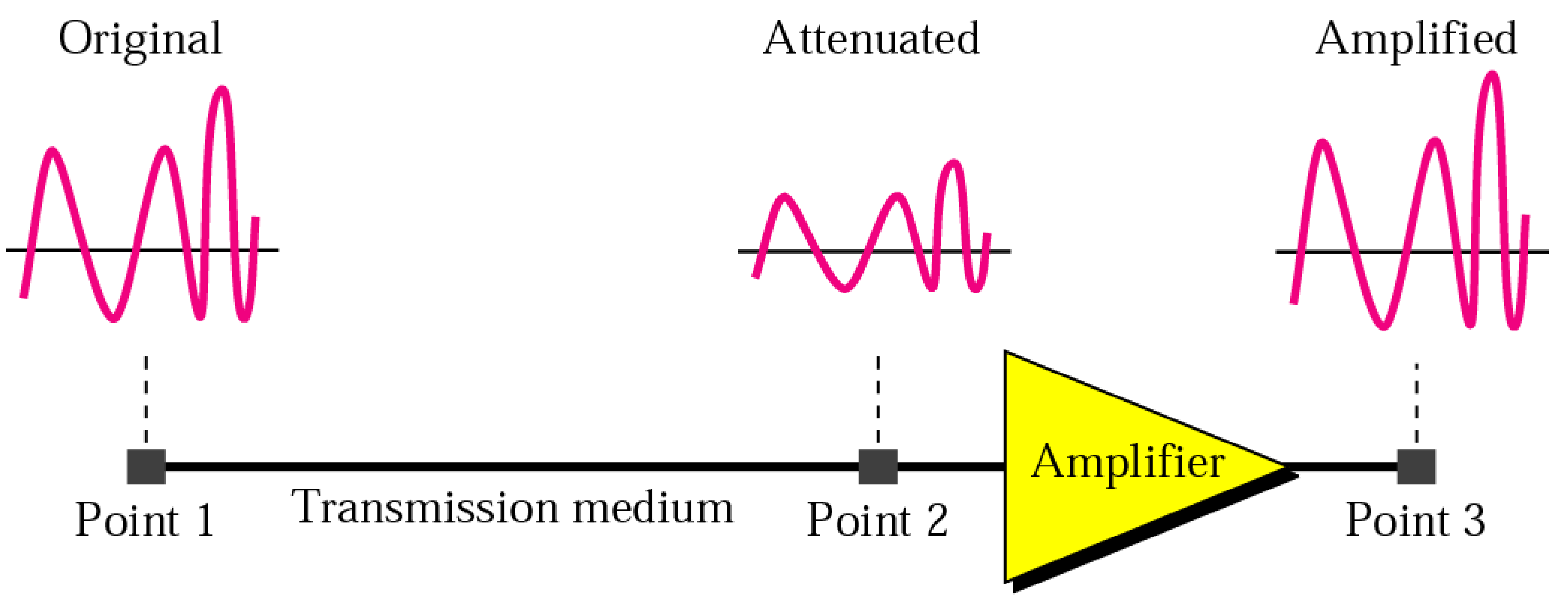


Attenuation

- Attenuation means the loss of energy.
- When signals travel through a medium they lose some energy so that they can **overcome** the ***resistance*** of this medium.

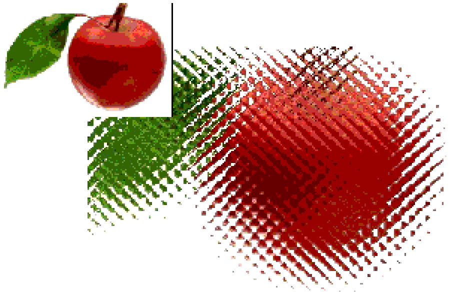
Attenuation

- To compensate for energy loss, **amplifiers** are used to boost the signal back up to its original level.



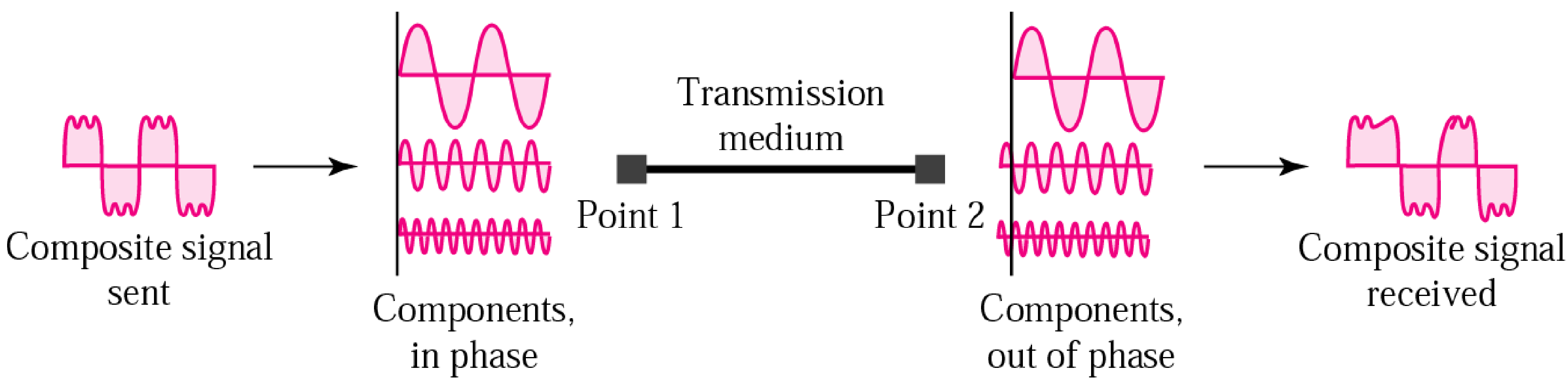
Distortion

- **Distortion**: signal *changes* in its **form** or **shape**.
- Typically effects complex or composite signals.
- Distortion takes place when a composite signal carrying different frequencies suffers from the delay of some of these frequencies.



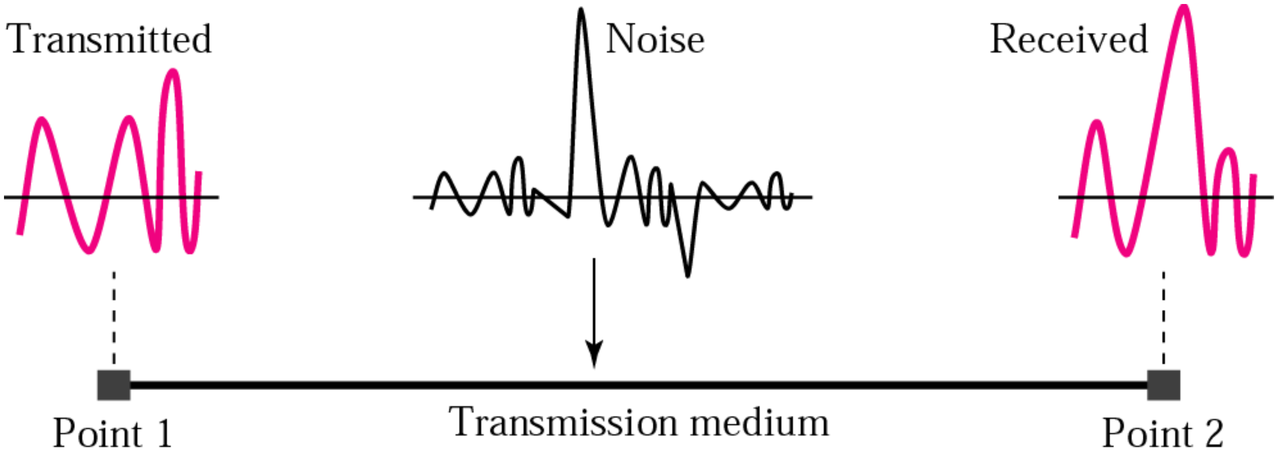
Distortion

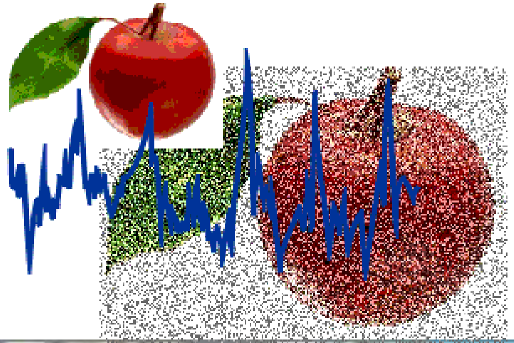
- Each frequency component has its own propagation **attenuation** through a medium.



Noise

- Noise is the **main source** of a signal being <u>corrupted</u>.



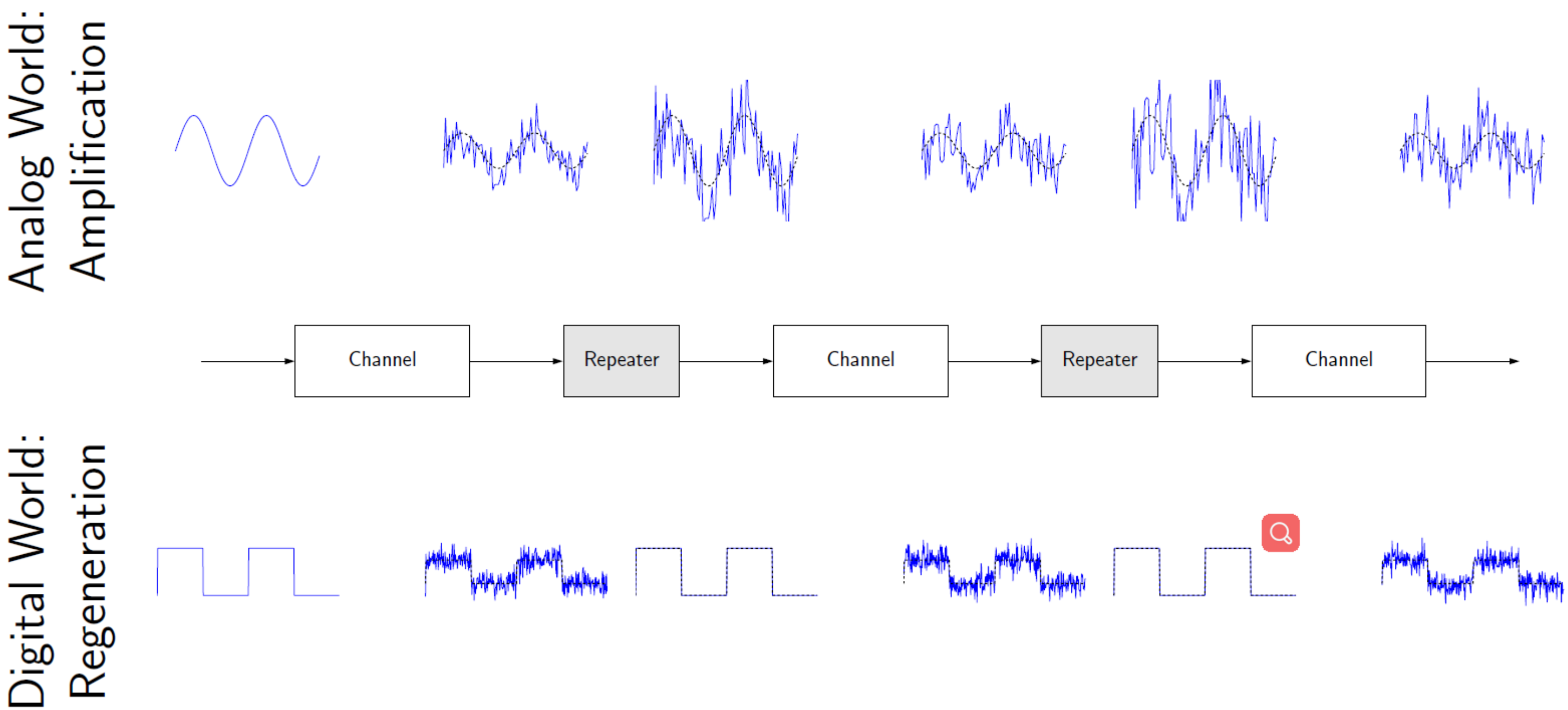


Amplification vs Regeneration

When a signal travels through a channel, it suffers **channel impairment**. Since their negative effects increase with the distance, special equipment called **repeaters** are inserted along the way.

- In analog systems continuously-varying waveforms are transmitted. In order to preserve the transmitted waveforms, repeaters essentially **filter**, **equalise** and **amplify** the signal.
- In digital systems sequences predefined waveforms (symbols) are transmitted. Repeaters **regenerate** such waveforms.

Amplification vs Regeneration



Efficiency: Circuit switching vs Packet switching

In communication networks, such as the PSTN information, sources and destinations are usually connected through intermediate systems, such as switches and routers.

- In **circuit switching** a dedicated path connecting source and destination is established. Resources are guaranteed for the whole connection time but they might be wasted.
- In **packet switching** no resources are reserved and the information to be sent is split into packets that can follow different paths to reach destination. Since resources are not reserved, they can be shared.

Hence, packet switching makes a more efficient use of the existing resources.

Resilience: Circuit switching vs Packet switching

Resilience: system can recover quickly when parts of it fail. Originally, packet switching was proposed as an alternative to circuit switching to guarantee the resilience.

- Traditionally, circuits were established in hierarchical networks by finding a common switching office. Network could be split into different isolated networks if a few switching offices failed.
- By contrast, in mesh networks many different paths can be established. Combining mesh networks topologies with packet switching results in fault tolerant systems, that is, systems that are resilient.

Incidentally, digital data are more amenable to be implemented in packet switching than analog data.

Summary

- Modern communications history
- Transmission medium
- Channel impairment

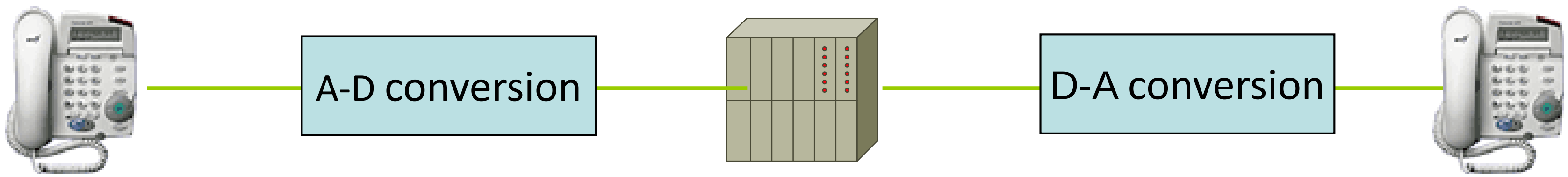
INFORMATION CONVERSION

Types of Information

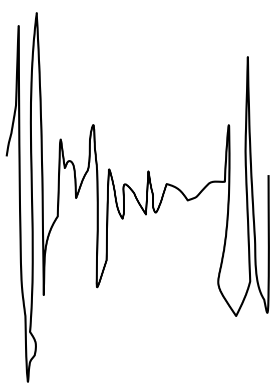
In communication systems, information to be transmitted can be categorized into three broad groupings:

- Text.
- Audio/Sound.
- Image/Video.

Transmission of information



Information:
'Hello! How are you?' You and I understand but not the telephone!



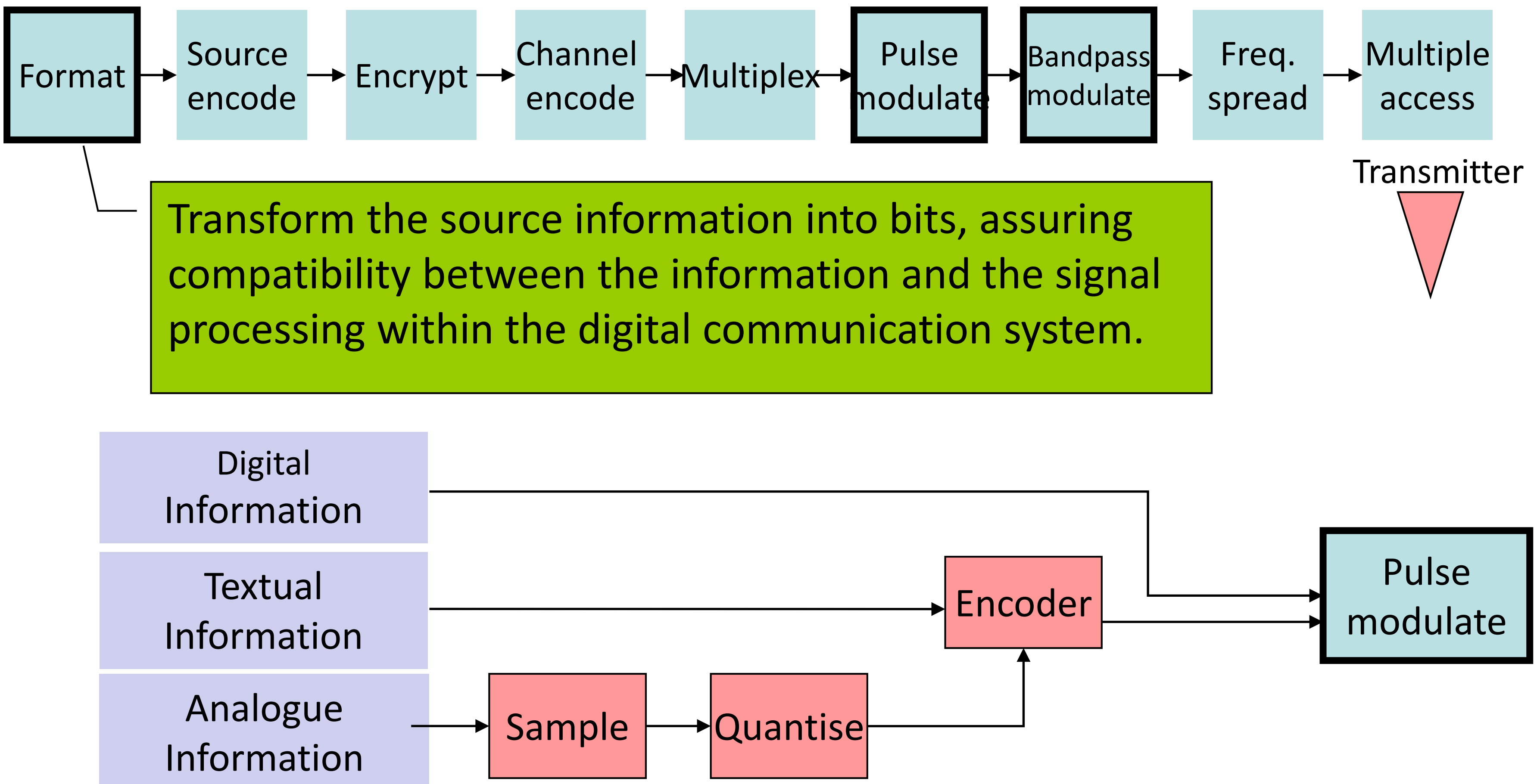
Analogue signal can be understood by electrical systems but problematic!

So all new systems are digital.

Information Conversion

- Different sources of information need different methods to transform the source information to a digital format
  - **Text** – ASCII, UTF
  - **Voice** (PSTN) – Pulse Code Modulation (G711a/u) 64kps
  - Voice (GSM) – GSM codec (13kbps) EFR (improved quality)
  - 3G WCDMA – AMR (adaptive Multi Rate)
  - Picture – JPEG …..
  - **Video** – MPEG2, MPEG4, H264
- Aim is to **minimise bitrate** but **maintain quality**.

Transmission side



Formatting Textual Data

- Textual information compromises a sequence of alphanumeric characters.

- Each alphanumeric character is transformed into **binary** by character coding.
  - Most popular character coding method is ASCII.

- Encoded into sequence of *k* bits called [symbols]

SAMPLING

Introduction to A/D conversion

Analog information has the following properties:

- Continuous in time.
- Continuous in amplitude.

Digital information consists of sequences of discrete values.

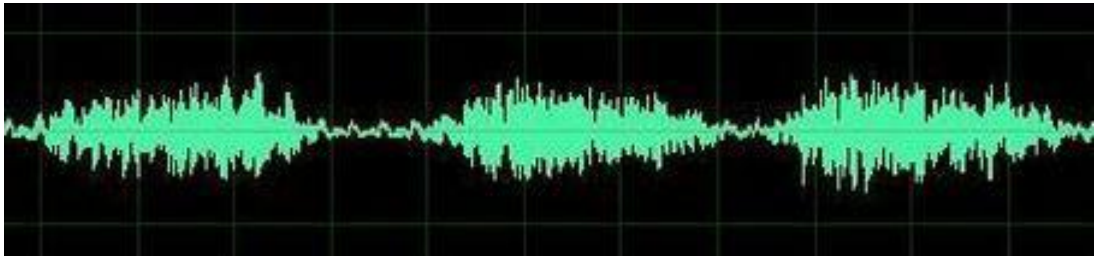
- Discrete in time.
- Discrete in amplitude.

The process of converting analog information into digital form is called **digitisation**.

Digitisation consists of discretising analog information both in time (**sampling**) and in amplitude (**quantisation**).

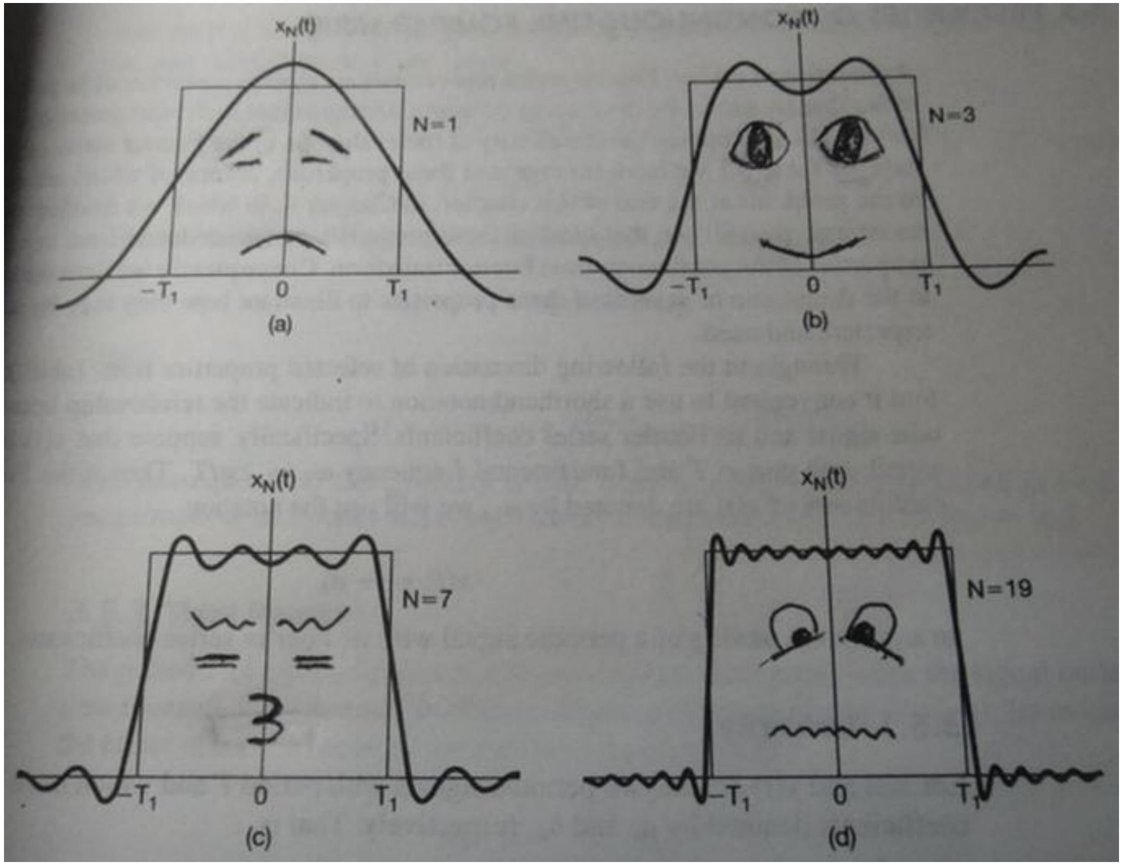
Review: Time Domain VS Frequency Domain

Time Domain

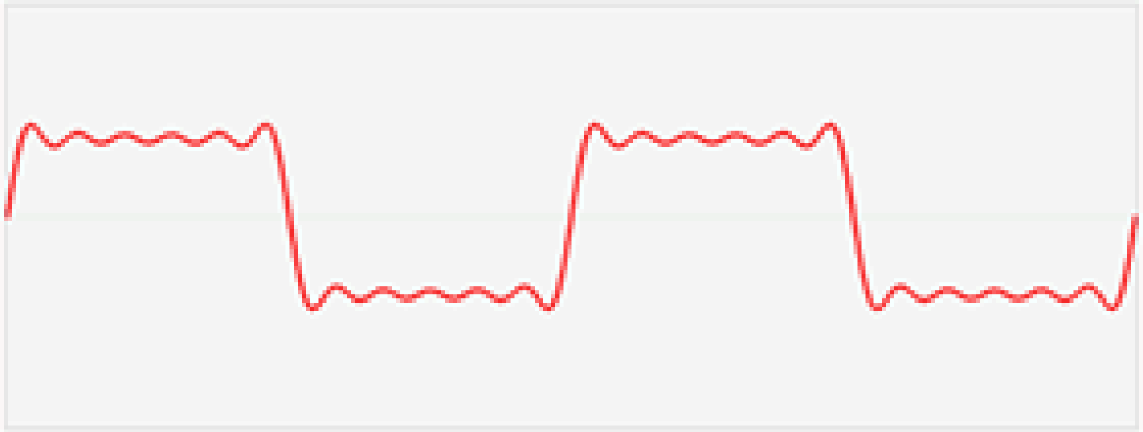


Frequency Domain

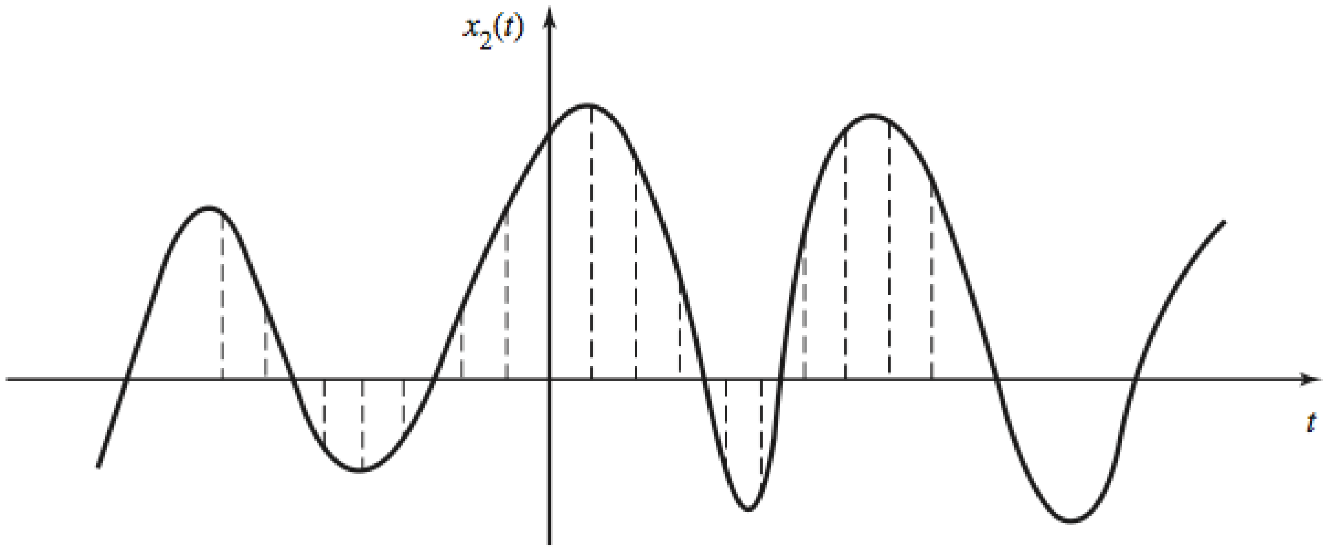
Fourier Series of Square Wave



Time Domain VS Frequency Domain： Dynamic



Sampling in the Time Domain

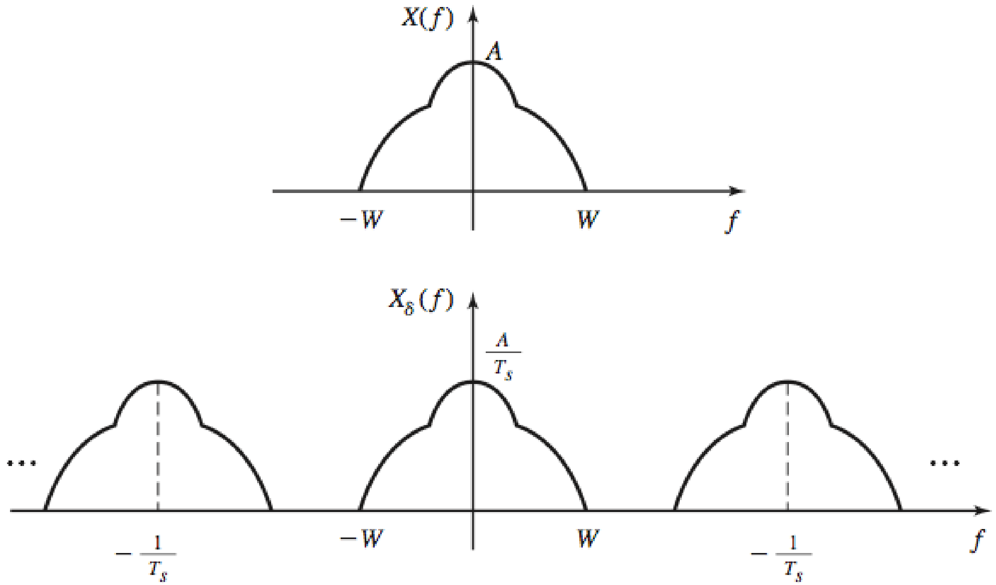


The process of sampling a signal x(t) can be mathematically modelled as the result of multiplying it by a train of impulses.

$$x_\delta(t) = x(t)\sum_{n=-\infty}^{\infty}\delta(t - nT_s) = \sum_{n=-\infty}^{\infty} x(nT_s)\delta(t - nT_s)$$

The quantity T_s is known as the **sampling period** and $f_s = 1/T_s$ is the **sampling frequency**.

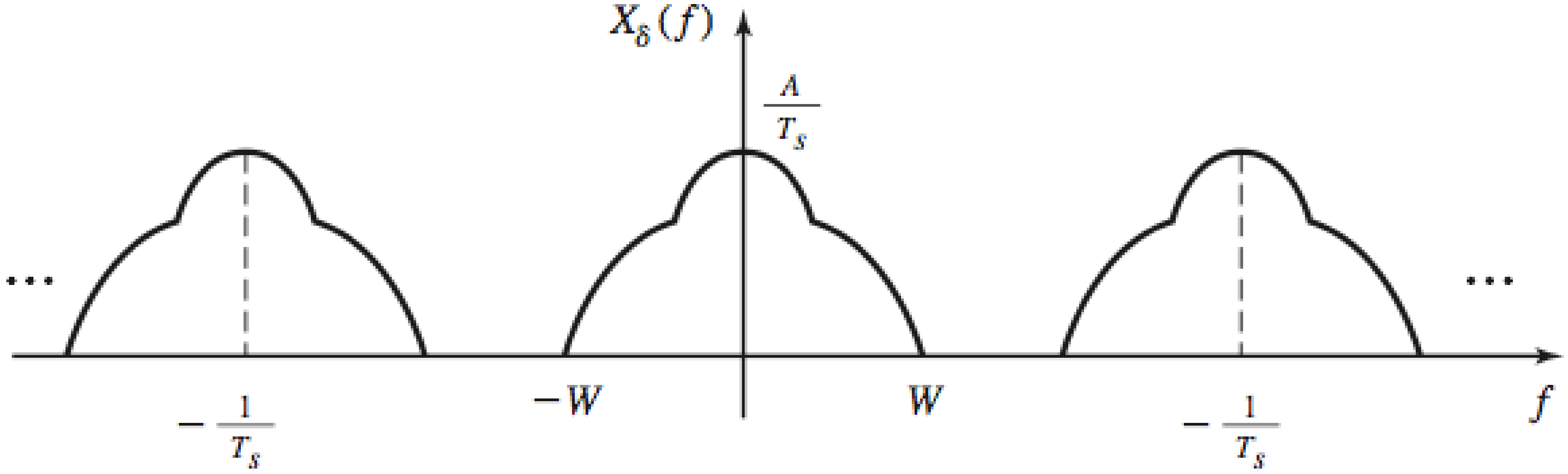
Sampling in the Frequency Domain



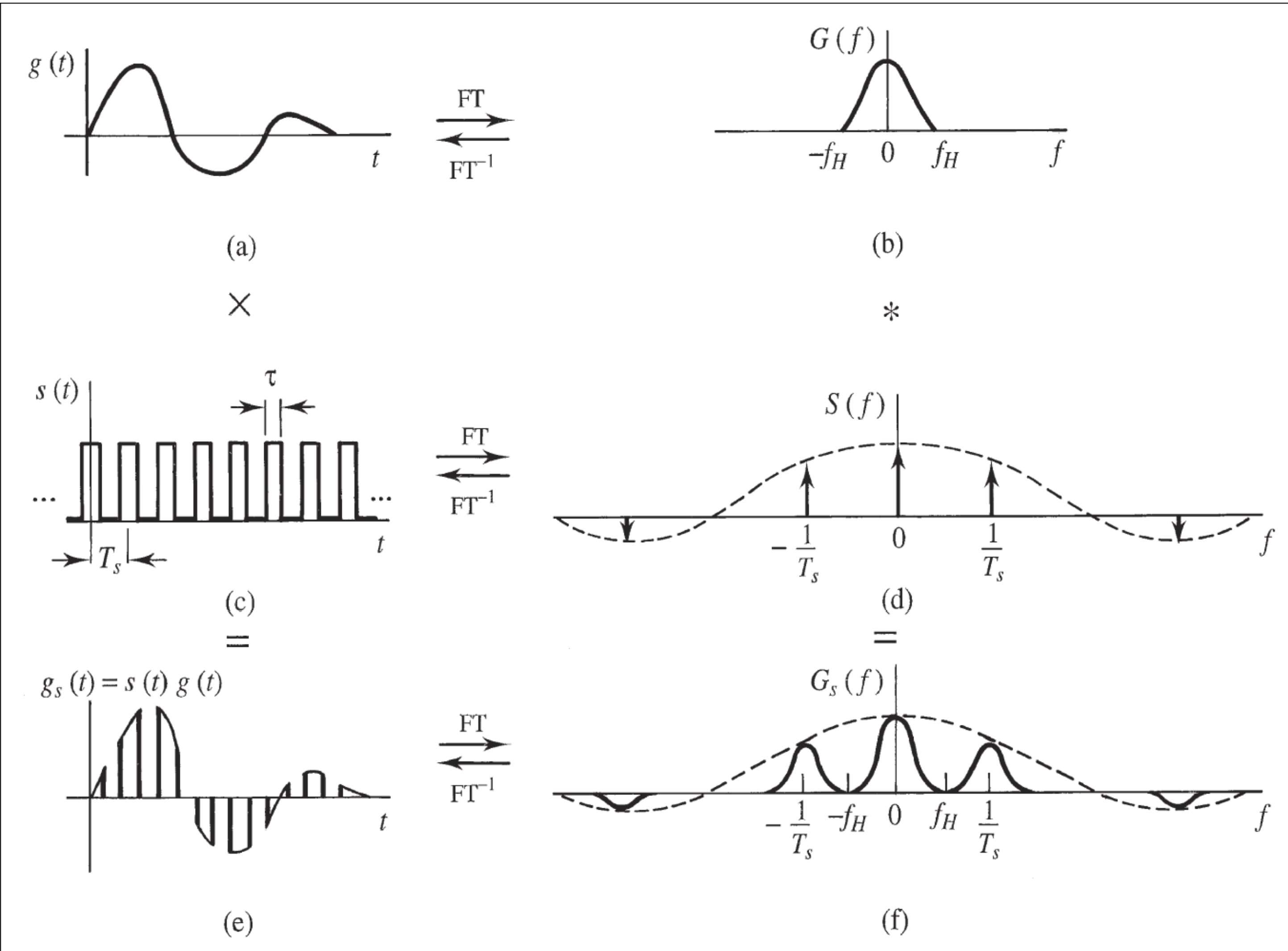
By using the **convolution property**, spectrum of a sampled signal consists of replicas of the original spectrum centred at multiples of the sampling frequency.

$$X_\delta(f) = X(f) \star \frac{1}{T_s} \sum_{n=-\infty}^{\infty} \delta\left(f - \frac{n}{T_s}\right) = \frac{1}{T_s} \sum_{n=-\infty}^{\infty} X\left(f - \frac{n}{T_s}\right)$$

Interpolating for D/A conversion



Natural Sampling (Periodic Pulse Train)



(a) signal g(t);

(b) signal spectrum;

(c) sampling function;

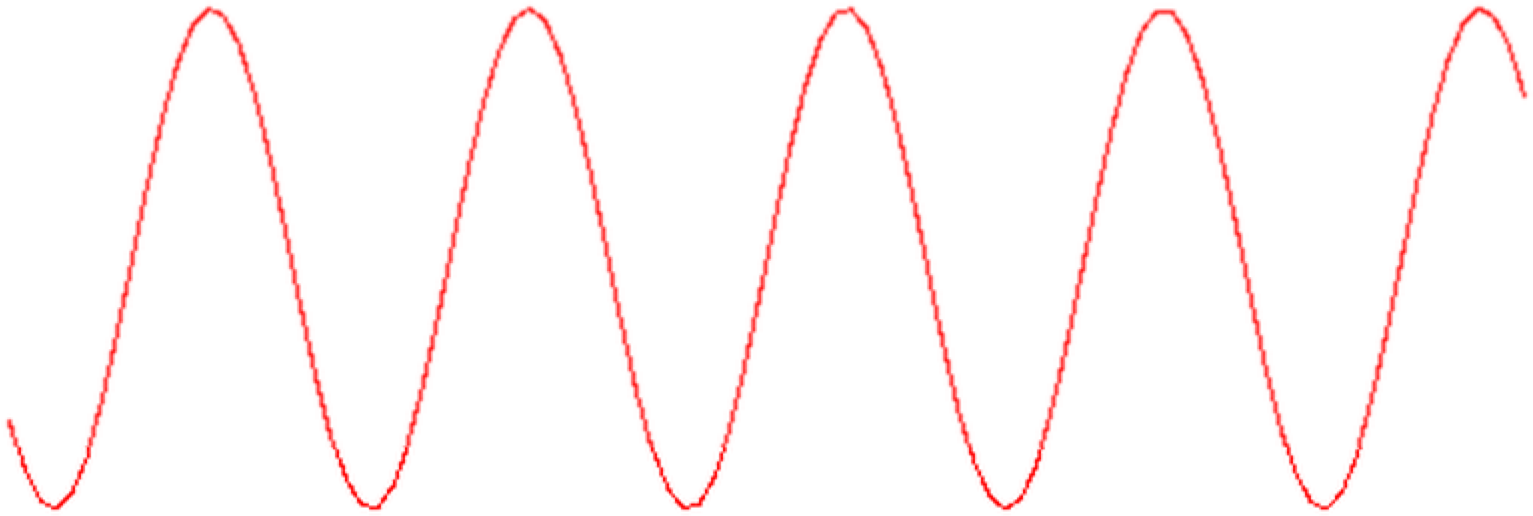
(d) spectrum of sampling function;

(e) sampled signal;

(f) spectrum of sampled signal

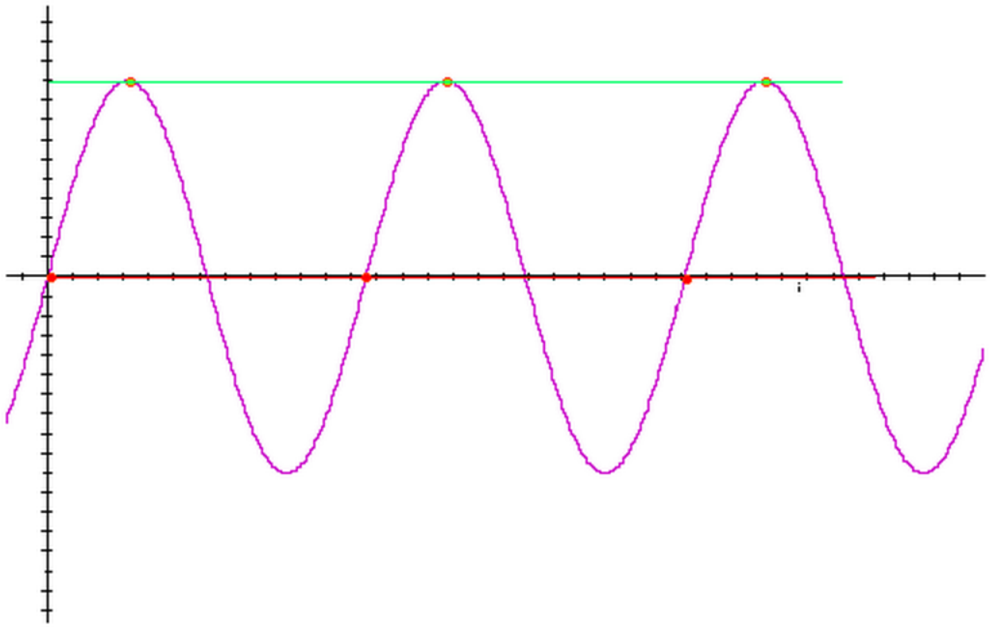
Nyquist's Sampling Theorem

- A sine wave:



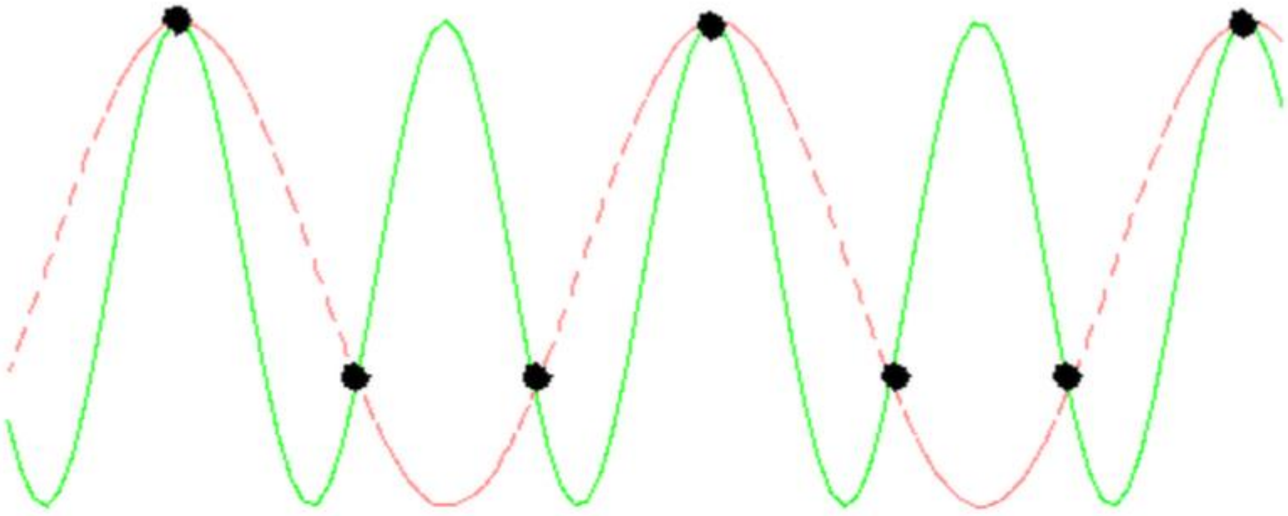
Sampling at 1 time per cycle

- If we sample at 1 time per cycle, we can think the reconstructed signal as a constant:



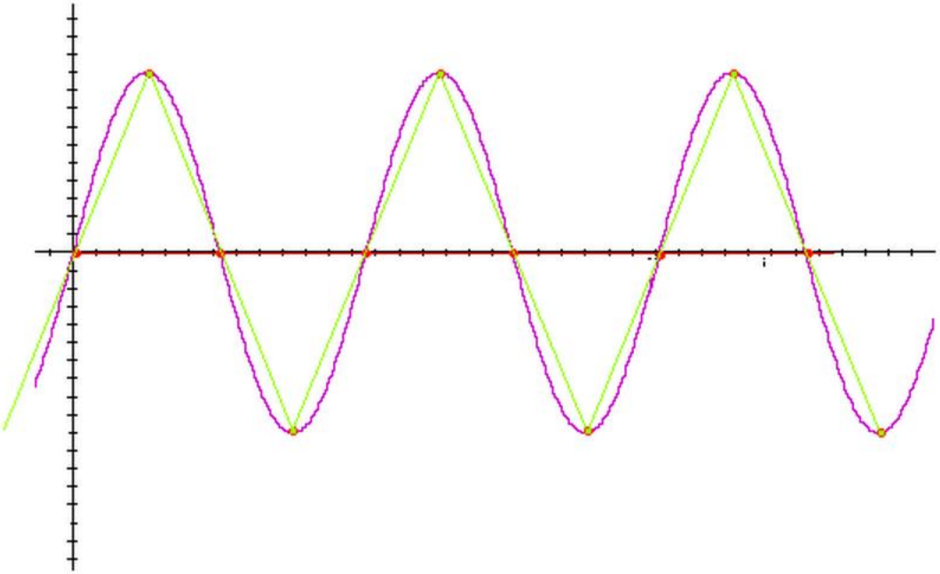
Sampling at 1.5 times per cycle

- If we sample at 1.5 times per cycle, we can think it's a lower frequency sine wave:



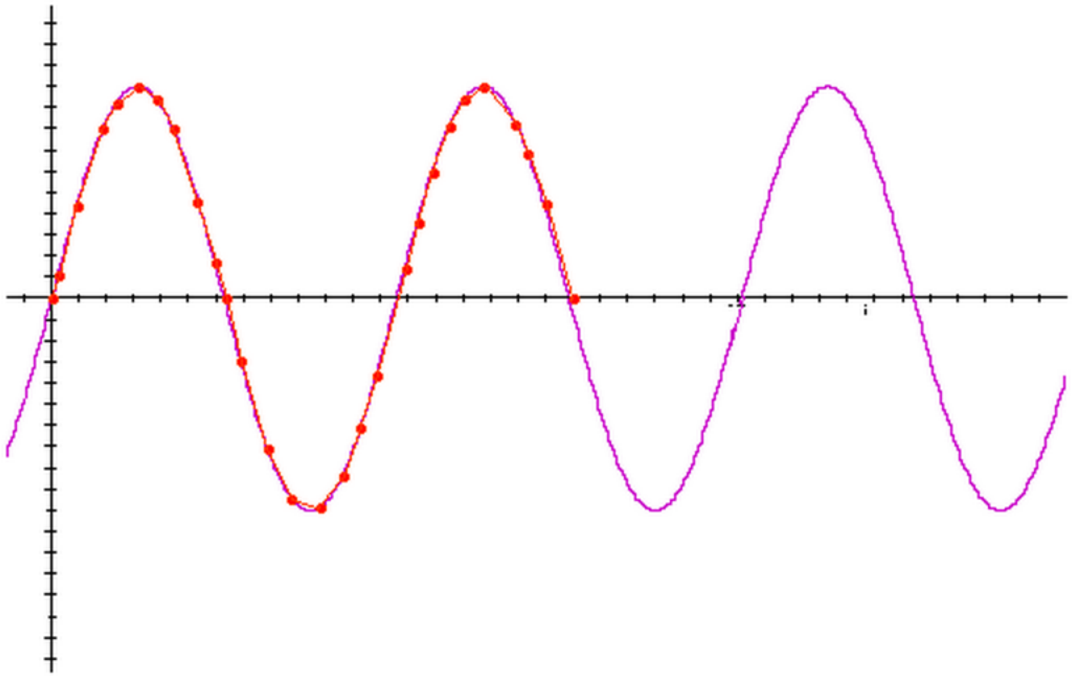
Sampling at 2 times per cycle

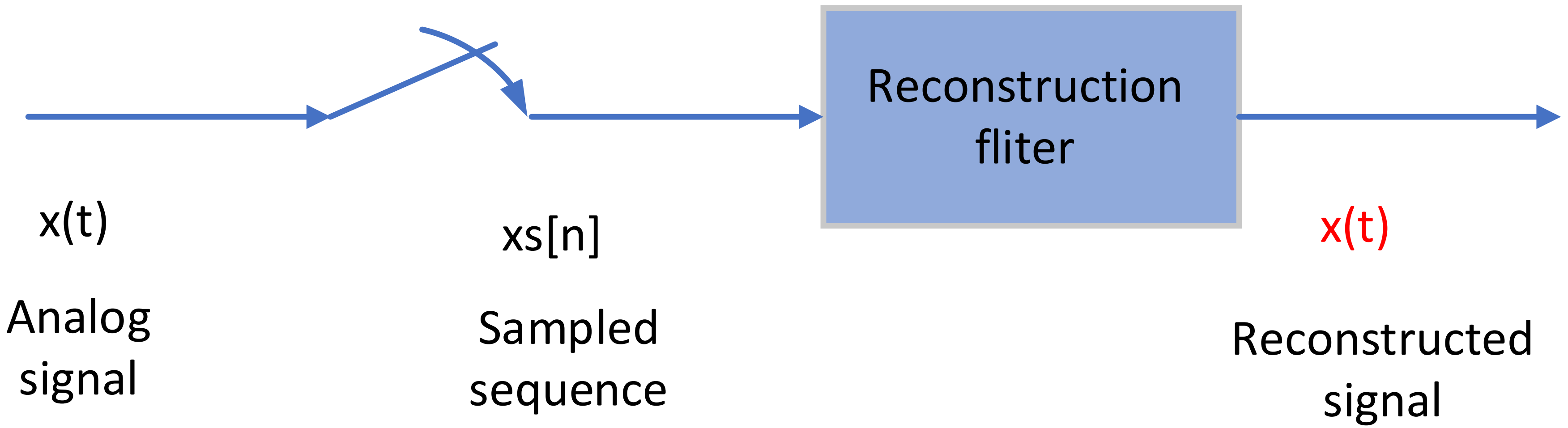
- If we sample at twice the frequency, the **Nyquist Rate**, we start to make some progress in terms of being able to replicate the signal:



Sampling at Many Times per Cycle

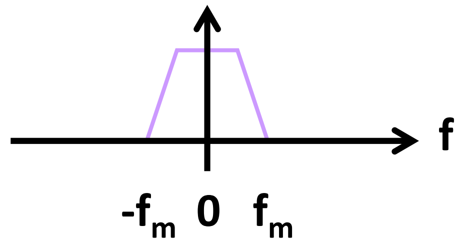
- For lossless digitisation, the sampling rate should be at least twice the maximum frequency responses. Indeed, many times more the better:



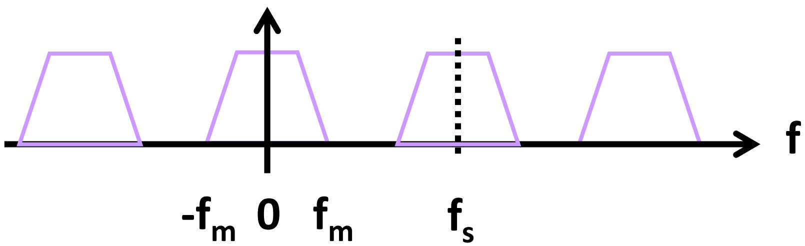

Reconstruction fliter
x(t)
Analog signal
xs[n]
Sampled sequence
x(t)
Reconstructed signal

Aliasing (ideal sampling)

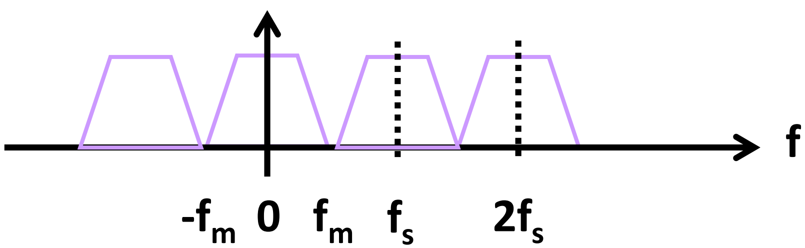
original signal



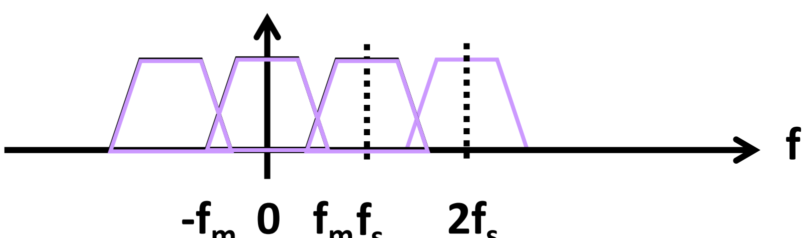
signal sampled with $f_s > 2\ f_m$



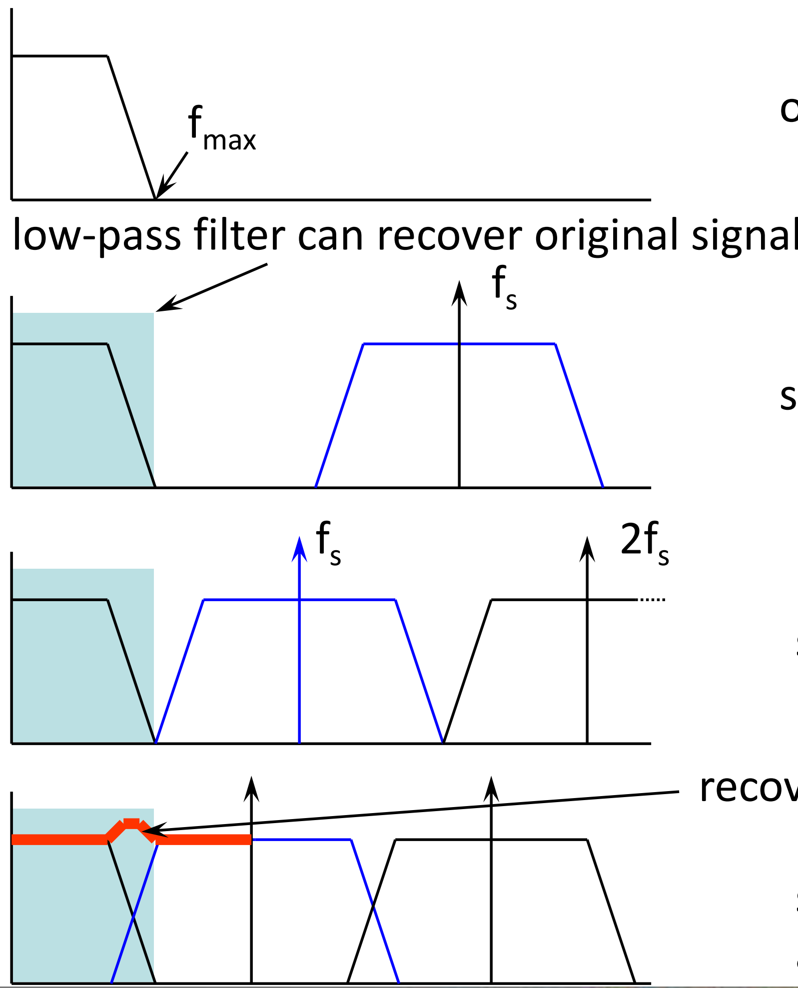
signal sampled with $f_s = 2\ f_m$



signal sampled with $f_s < 2\ f_m$
aliasing occurs



Aliasing in more detail



original signal

signal sampled with $f_s > 2\ f_{max}$

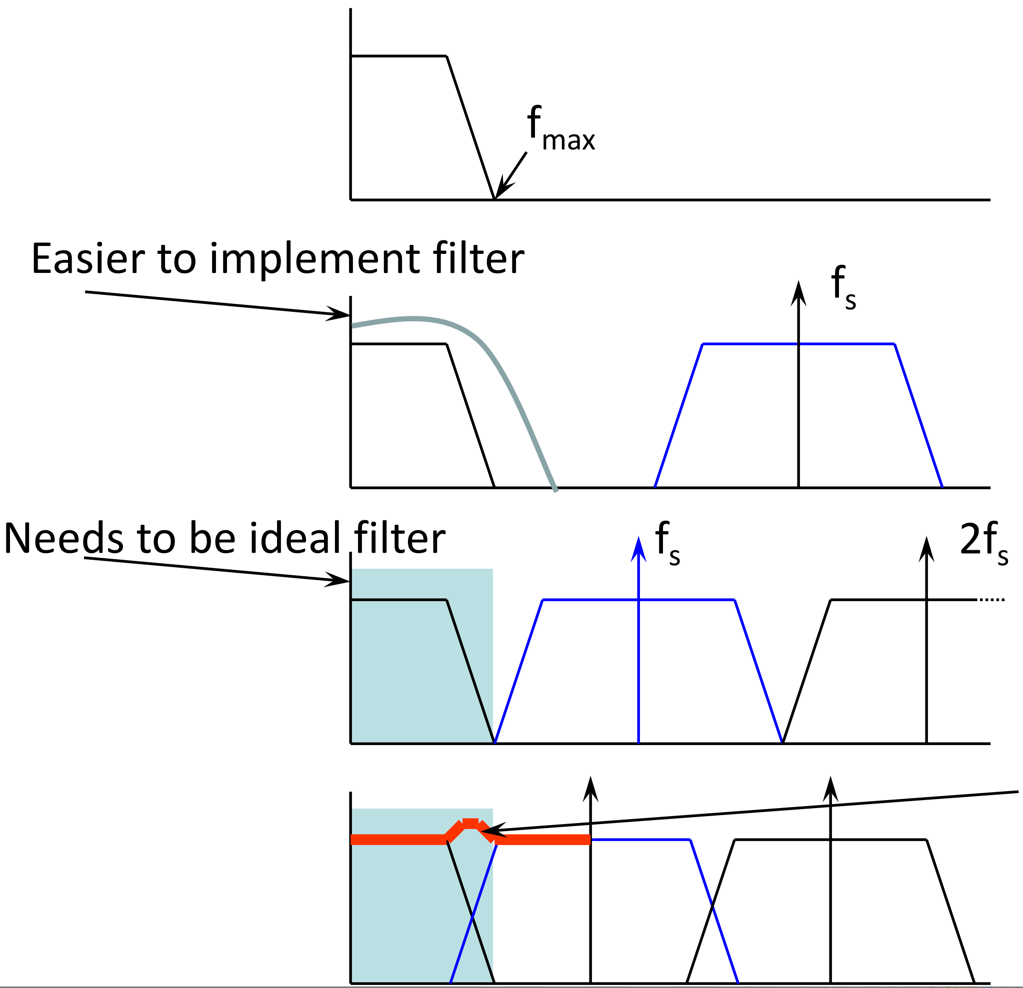
signal sampled with $f_s = 2\ f_{max}$

recovery not possible - spectra interfere

signal sampled with $f_s < 2\ f_{max}$
aliasing distortion occurs

Oversampling

low-pass filter can recover original signal



original signal

oversampling
signal sampled with $f_s > 2\ f_{max}$

signal sampled with $f_s = 2\ f_{max}$

recovery not possible - spectra interfere

signal sampled with $f_s < 2\ f_{max}$
aliasing distortion occurs

Question

- How to avoid aliasing?

1. Using anti-aliasing filter before sampling
2. Sampling over Nyquist sampling.

Sampling Theorem (Nyquist's Criterion)

- To prevent aliasing and hence to allow the original signal to be recovered the sampling frequency (f_s) must be given by:

$$f_s \geq 2\, f_{max}$$

  where f_{max} is the highest frequency present in the original signal.
- This is the SAMPLING THEOREM and is a fundamental theorem.
- Oversampling makes it easier to design a simpler filter to recover the original signal.

Example

- Let $x(t)$ be a band-limited signal to W = 10 kHz, amplitude $0 \leq x(t) \leq 2$. Signal $x(t)$ is sampled at a rate 20% higher than the Nyquist rate to provide a guard band. Calculate the sampling rate for x(t).

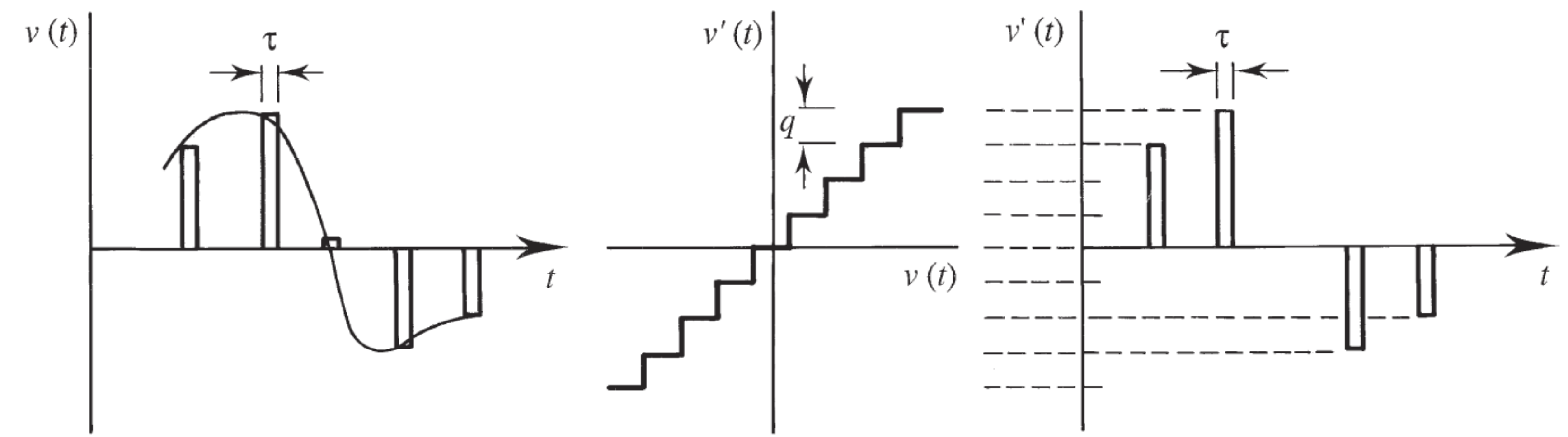
Summary

- A/D, D/A
- Aliasing
- Sampling theorem

QUANTISATION

Quantisation

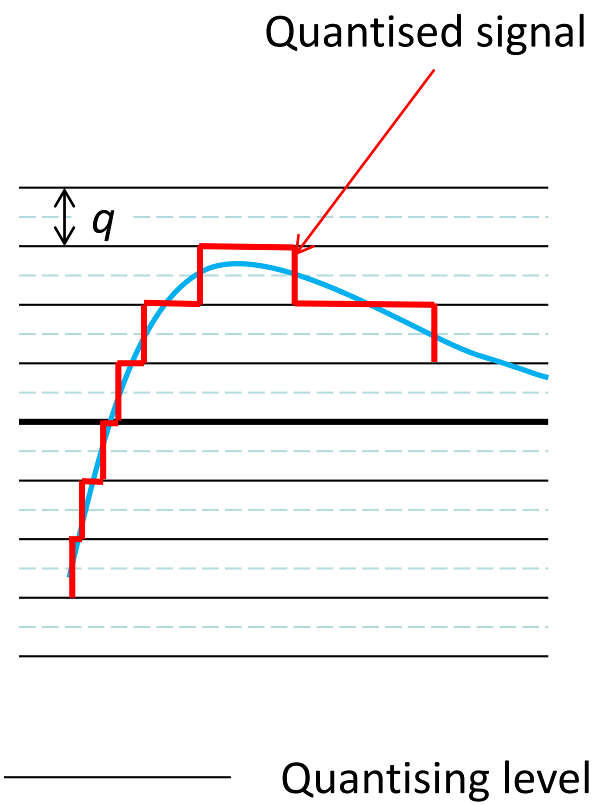
- Results from mapping **continuous analogue values** to **discrete vales** that can be represented digitally.
- May be linear or non-linear
- Pulse-code modulation (PCM) is a method used to digitally represent sampled analogue signals



(a) Analogue input signal (b) Quantiser characteristics (c) Quantised output signal

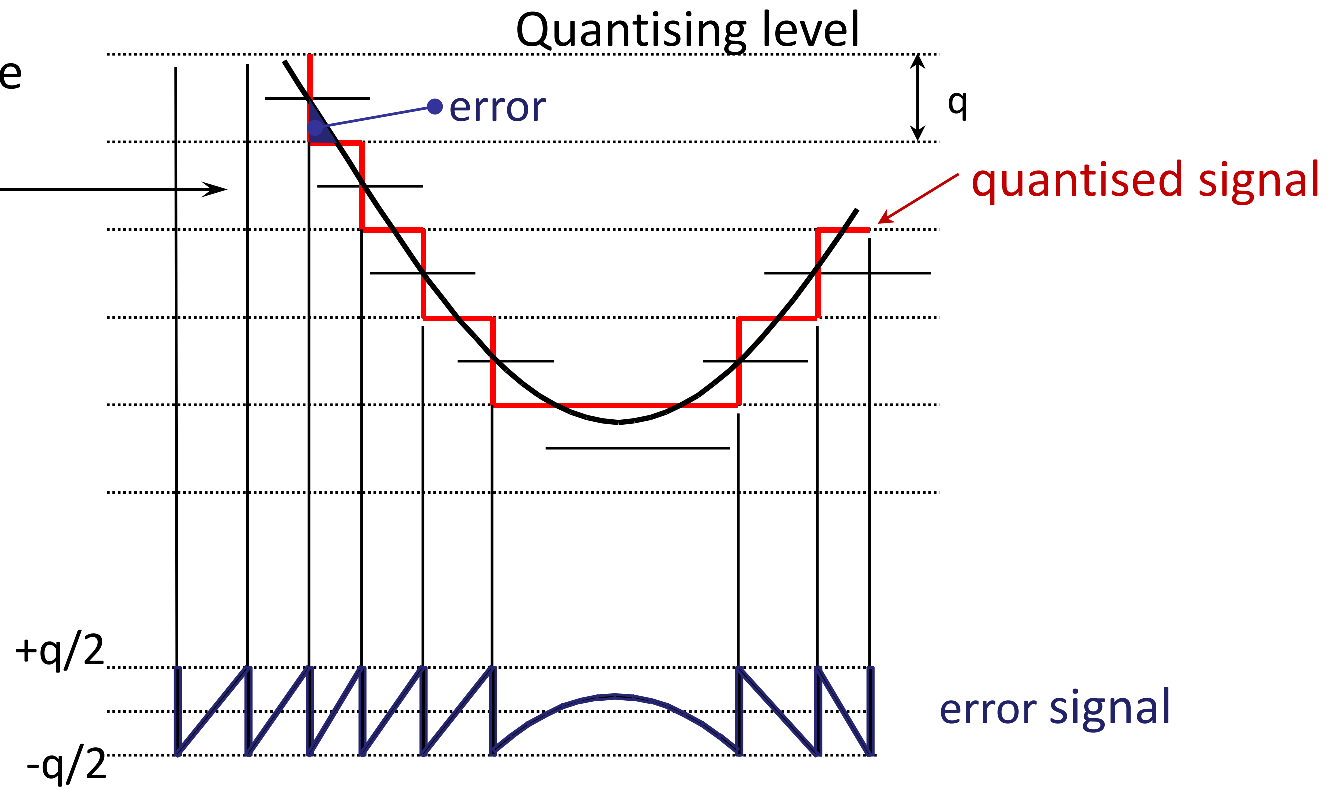
Linear Quantisation

- Peak-to-Peak Voltage, $V_{pp}=V_p-(-V_p) = 2V_p$
- Quantisation interval, q, (step size) uniformly distributed over the full range
- Approximation will result in an error no larger than $\pm q/2$



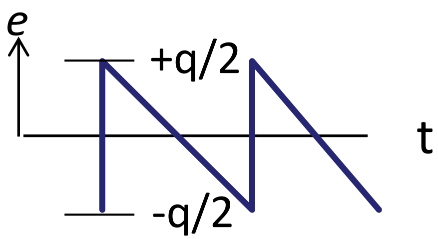
Quantising Distortion

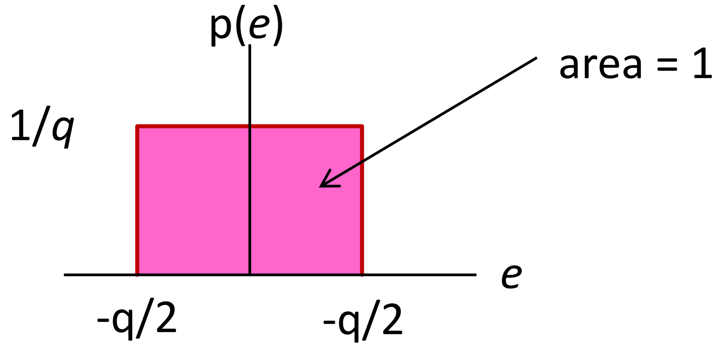
Signal changes state half-way between quantising levels



Quantising error power

Error (e) is approximately sawtooth over the quantisation region, apart from the dwell regions.





A sawtooth waveform has a uniform pdf: all values are equally likely.
The area under the pdf must be 1 so that the amplitude is $1/q$.
Note that p(e)=0 outside the range $+q/2$ to $-q/2$.

Power of the quantisation error is: $P_Q = \int_{-\infty}^{\infty} e^2 p(e)\, de = \int_{-q/2}^{q/2} e^2 \frac{1}{q} de = q^2/12$

Notes

- This holds for reasonable well-behaved signals without frequent dwell regions.
- **Quantising error leads to distortion**, not noise, because the same input will always produce the same output.
- The statistics of the distortion are independent of the statistics of the input.
- This approximation shows that the distortion power is constant and depends only on the step size.

Signal to Quantisation Noise/Error Ratio (SQNR)

The average signal to quantisation noise ratio (SQNR) is:

$$SQNR = \frac{P_s}{P_Q} \qquad P_s = \overline{V^2}$$

q=$2V_{peak}/2^n$

$$SQNR = 4.8 + 6n - \alpha_{dB}$$

n: number of bits per sample

α is the ratio of peak to mean signal power, ${V^2}_{peak} / \overline{V^2}$

Review: What is dB?

- The **intensity of sound** is measured in decibels (dB).
- This is not a linear scale, as the human ear does not perceive volume changes in a linear way.
- The ear's response to sound changes is **logarithmic** and therefore audio volume controls are similarly logarithmic.
- How to calculate logarithmic volume
  - $\mathbf{x=10*log_{10}(y)}$, x in dB and y in decimal
  - $10*\log_{10}(1000)=10*3=30\text{dB}$
  - $10*\log_{10}(2)=10*0.3010=3\text{dB}$

Practise for Quantization

Calculated the bit per simple required to transmit x(t) as a linearly quantised PCM signal maintaining an SQNR of 55 dB. (Assume that the signal's peak to mean ratio is 20dB).

Solution

As for linearly quantised PCM signals, the Signal to Quantisation Noise Ratio in dB is

$$\boldsymbol{SQNR = 4.8+6n-\alpha dB}$$

where α is the signal's peak to mean ratio and n is bit number for each symbol.

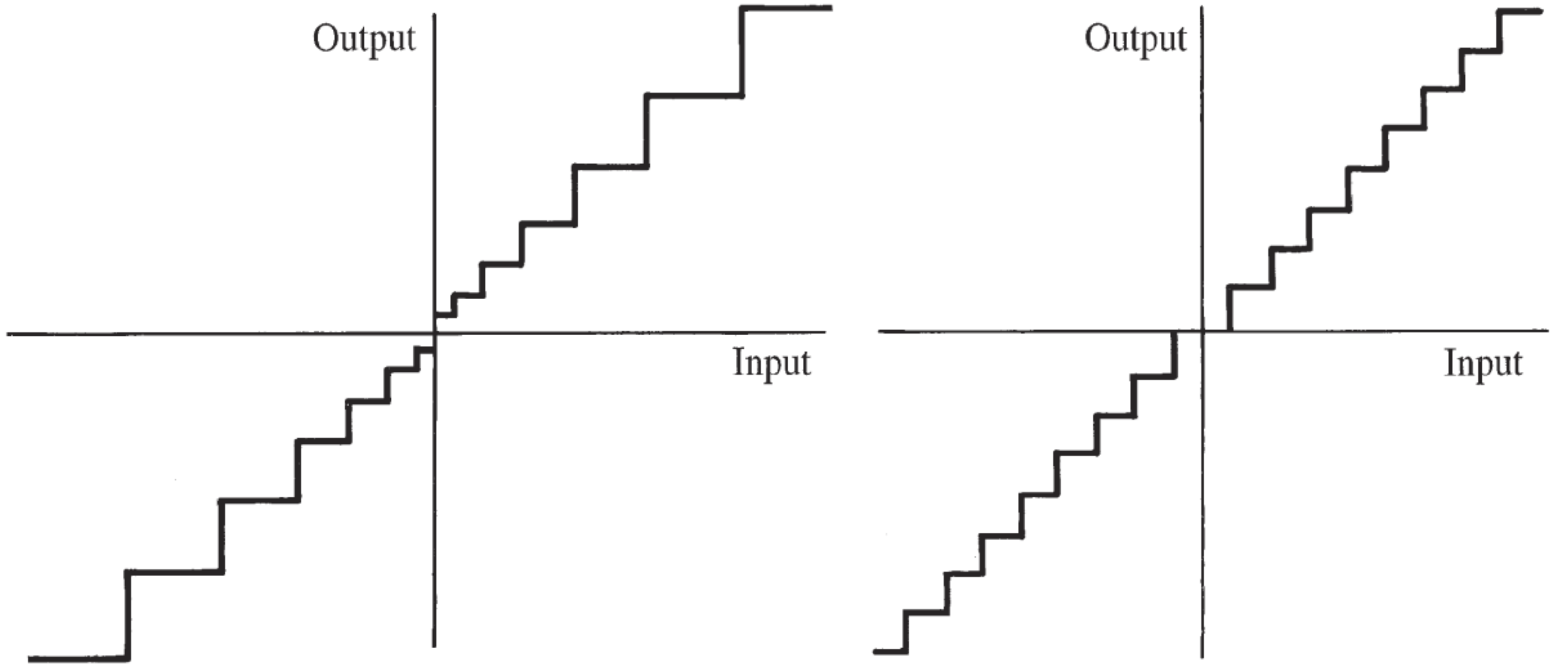
$$\boldsymbol{55=4.8+6n-20 \rightarrow n= 11.7}$$

So 12 bit/symbol is needed for each PCM symbol.

Speech and non-linear Quantisation

- Speech power has a wide dynamic range – whispering, shouting
- SNR is worse for lower powers as quantisation noise is same for all signal magnitudes ($q^2/12$).
- Non-uniform quantisation can provide fine quantisation of the weak signals and coarse quantisation of the strong signals.
- 8-bits per sample not sufficient for good speech encoding with uniform quantisation.
- Solution is to use **non-linear quantisation**:
  - Step-size varies with amplitude of sample.
  - For larger amplitudes, larger step-sizes are used
  - 'Non-linear' because **step-size changes from sample to sample**.

Linear Quantisation and Non-linear Quantisation

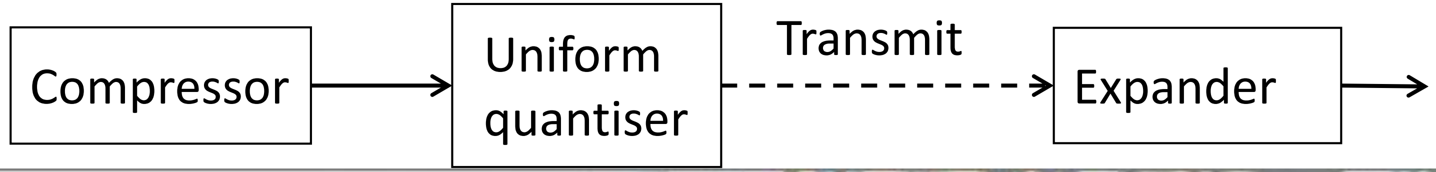


(a) Non-linear quantisation

(b) Linear quantisation

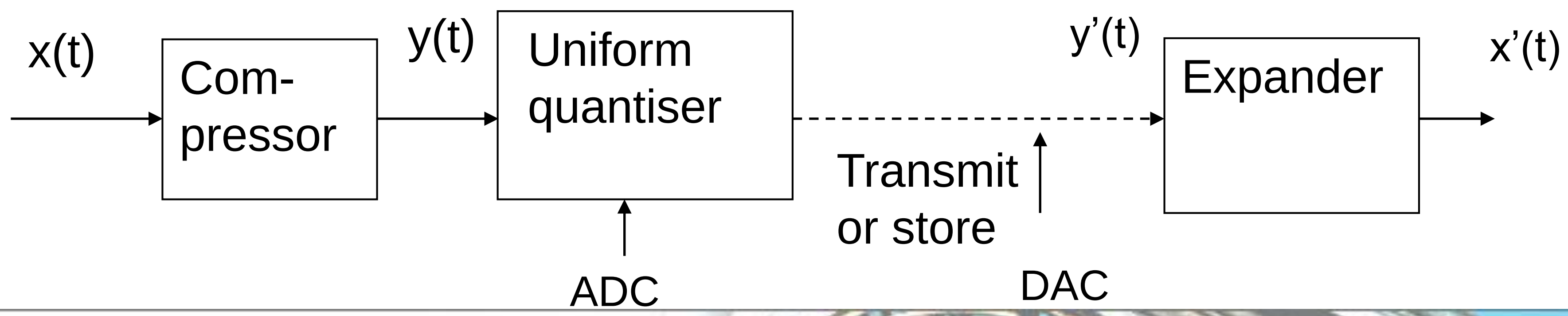
Non-linear quantisation - principle

- Non-linear quantisation uses **logarithmic compression and expansion**.
- Compress at the transmitter and expand at the receiver.
- **Compression changes the distribution of the signal amplitude**.
  - Enlarge the small input and compress the large input
  - As the result, the compressed speech signal is now more suitable for linear quantisation.
- The logarithmic compression and expansion function is also called **Companding**.

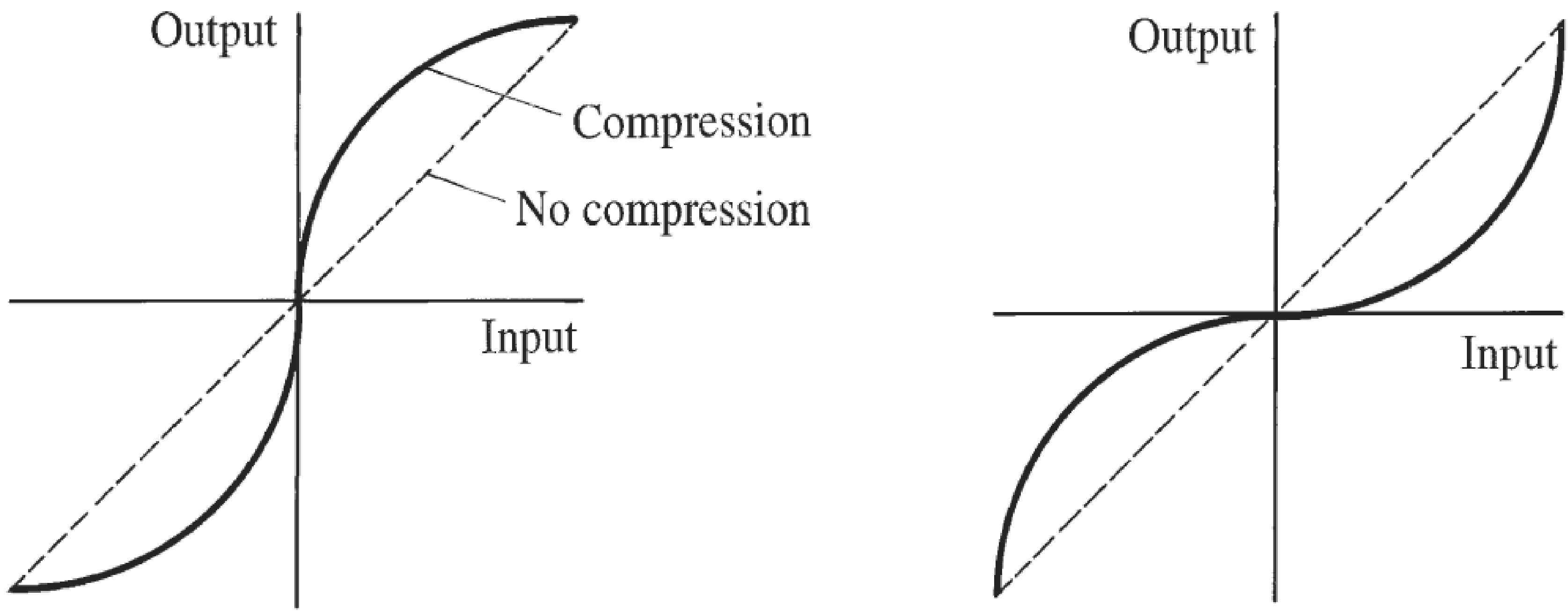


Implementation of Companding (in principle)

- Pass x(t) through **compressor** to produce y(t).
- y(t) is quantised uniformly to give y'(t), which is transmitted or stored digitally.
- At receiver, y'(t) passed thro' **expander** which reverses effect of compressor.
- Analogue implementation is uncommon but shows concept well.



Typical Compander Characteristics



(a) Compression characteristic

(b) Expansion characteristic

Companding

- There are two companding standards for telephony:
  - A-law (G711a - used mainly in Europe)
  - μ-law (G711u - used in North America and Japan).

- 8-bit code consist of

  i) polarity bit P (range is ±V)

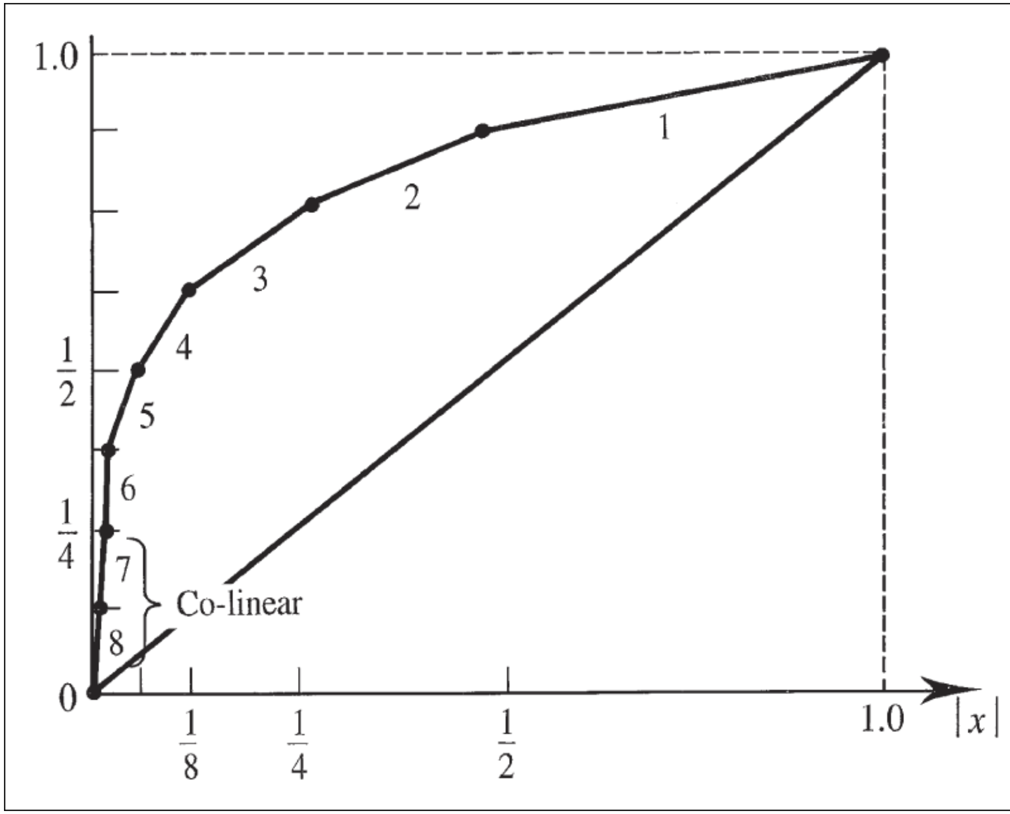
  ii) 3 segment decoding bits XYZ

  iii) 4 bits (abcd) specifying intra segment value on a linear scale

13-segment compression A-law

- Implemented as a segmented, piece-wise linear approximation.

- In the implementation, use straight line to connect any two adjacent points on the curve.



A-law Encoding Table

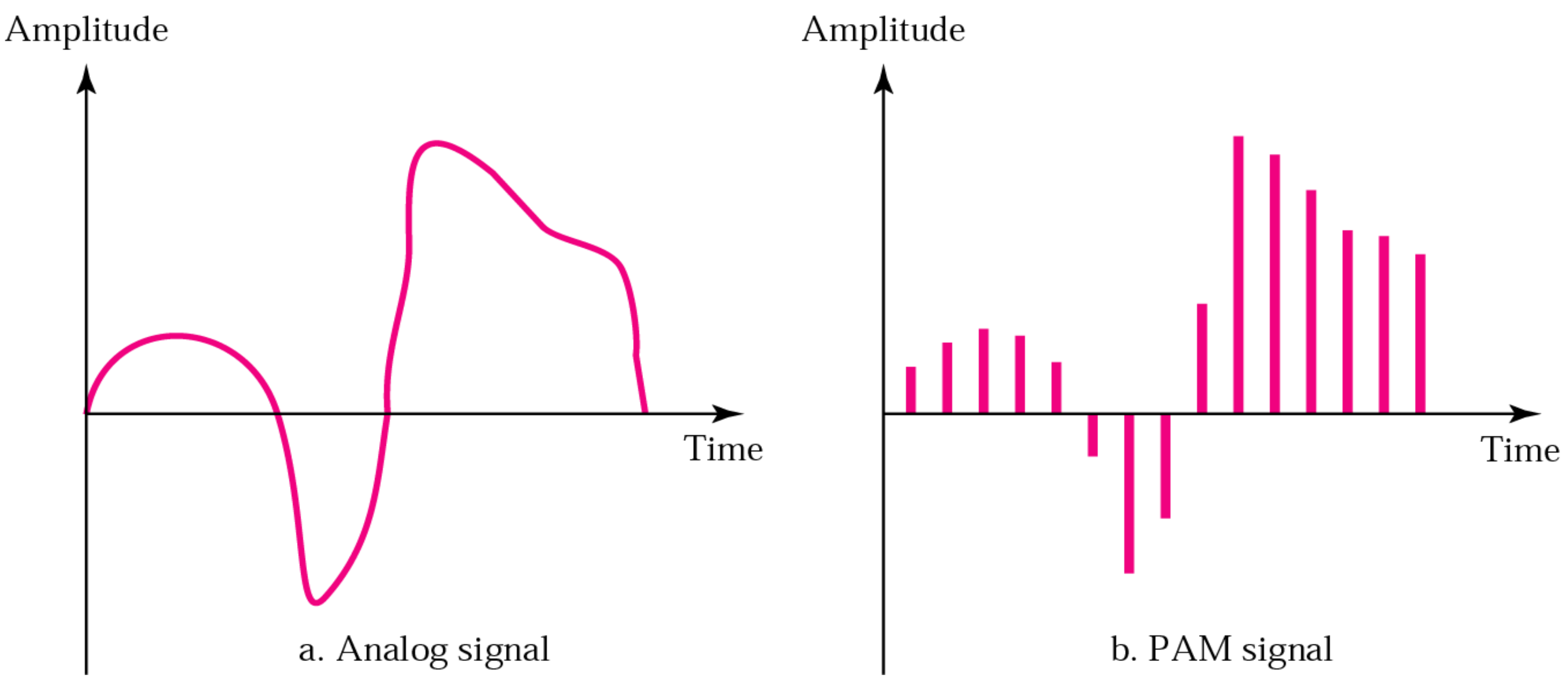
Segment	Coder input range	output code	quantum interval
0	0-V/128	P 000 abcd	V/2048
1	V/128-V/64	P 001 abcd	V/2048
2	V/64-V/32	P 010 abcd	V/1024
3	V/32-V/16	P 011 abcd	V/512
4	V/16-V/8	P 100 abcd	V/256
5	V/8-V/4	P 101 abcd	V/128
6	V/4-V/2	P 110 abcd	V/64
7	V/2-V	P 111 abcd	V/32

P is a polarity bit

abcd is a 4-digit *intra*-segment code – linear quantisation

Pulse amplitude modulation (PAM)

- PAM: Pulses are sent with **equal width** at regular intervals, but the height of the pulse is varied according to the sample value.



a. Analog signal

b. PAM signal

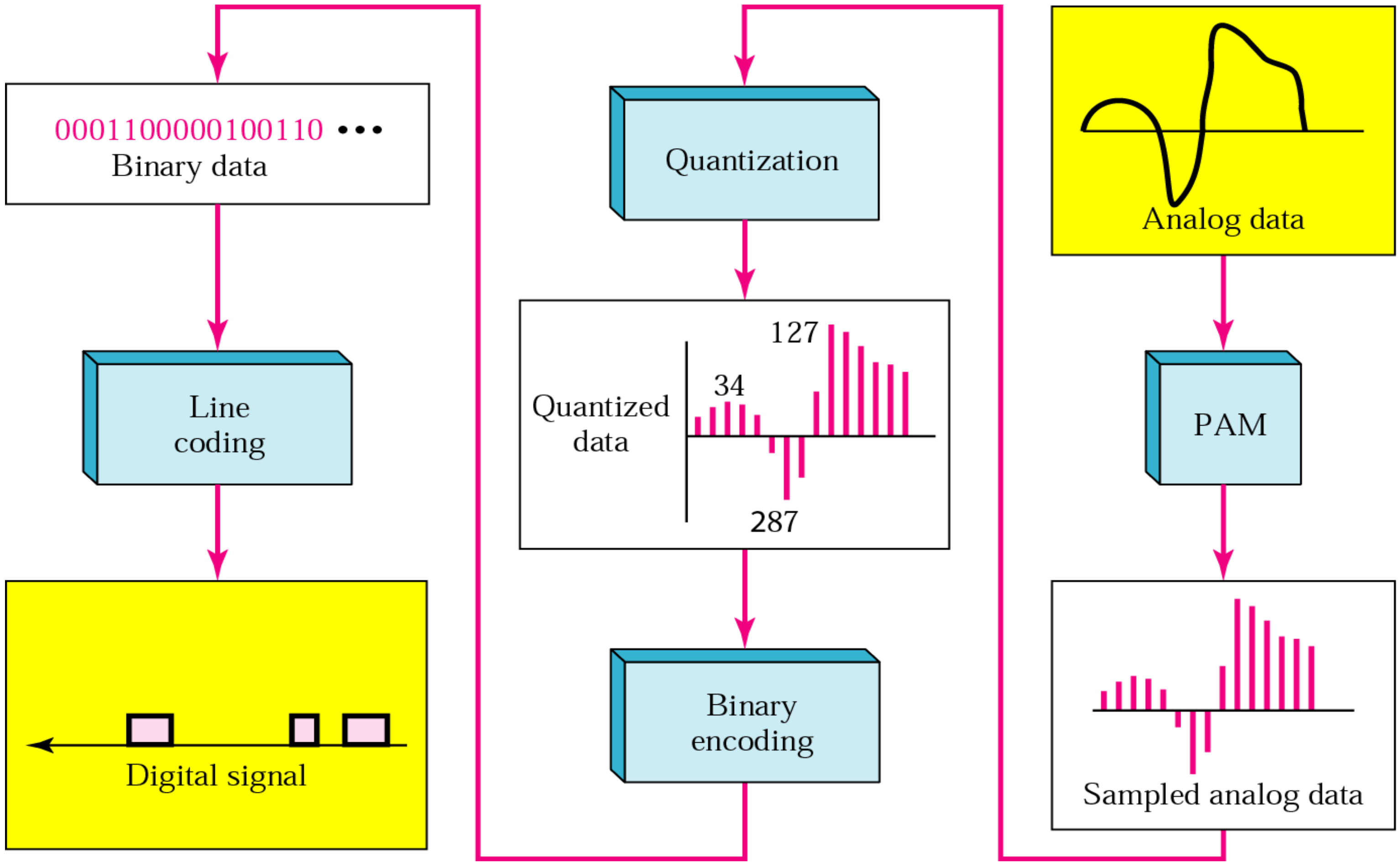
Pulse-Code Modulation (PCM)

- PCM includes three components: **sampler, quantiser and encoder.**
- The encoder converts the sequence of quantised amplitudes into a sequence of bits. Hence, the bit rate R_B can be calculated as

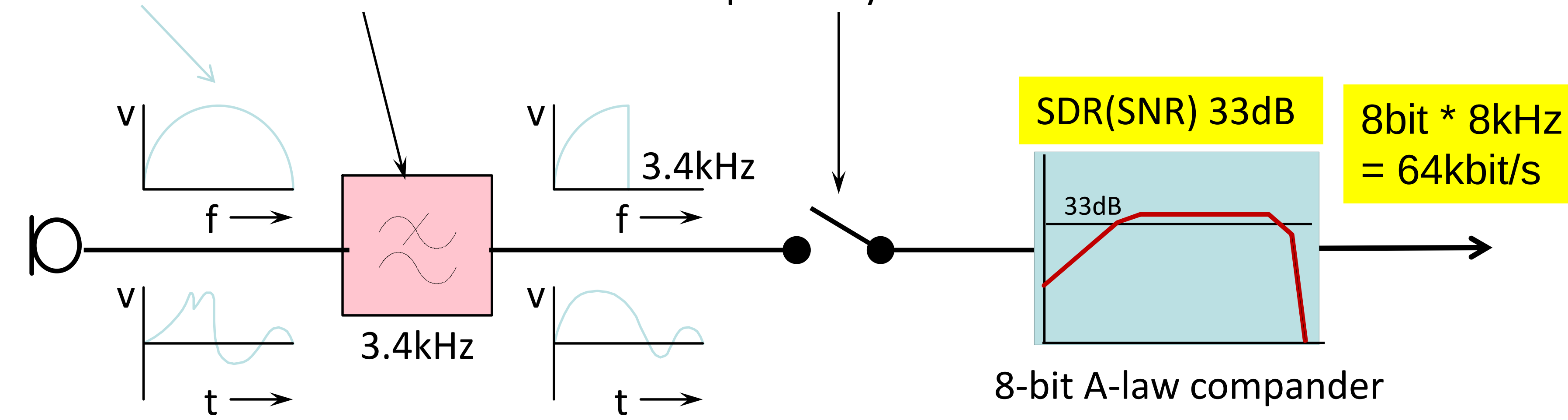
$$R_B = f_s \times N_B \ bps$$

where $\boldsymbol{f_s}$ **is the sampling frequency** and $\boldsymbol{N_B}$ **is the number of bits per quantisation level**. The quantiser can be either uniform or non-uniform. In telephony, speech f_s = 8 kHz, N_B = 8; hence R_B = 64 kbps.

Pulse-Code Modulation



Pulse-Code Modulation Telephony



By bandlimiting the incoming speech signal to 3.4kHz and sampling at 8kHz, the sampling theorem is satisfied

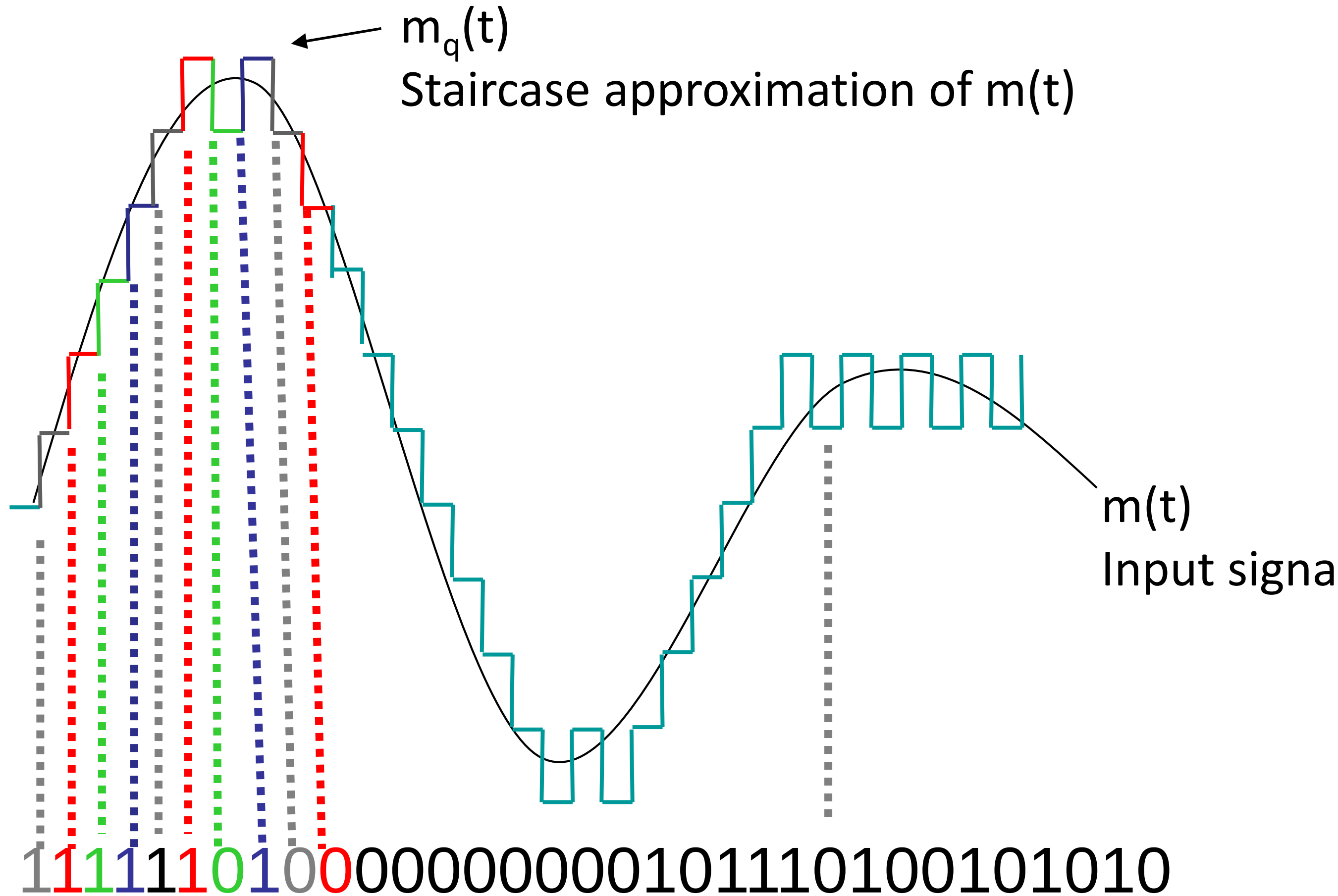
You need to remember this diagram:
- anti-aliasing filter gives bandwidth of 3.4kHz to the sampler
- sampling rate of 8kHz – 1 sample every 125 µs
- 8-bit (a-law) non-linear quantisation

DELTA MODULATION

Delta Modulation (DM)

- Simpler than PCM.
- Provides a **staircase version** of the message signal by referring to the difference between the input signal and its approximation.
- Quantization is done using 2 levels:
  - Positive difference: $+\Delta$
  - Negative difference: $-\Delta$
- Provided that the **input signal does not change too rapidly** from sample to sample, this approximation works well.

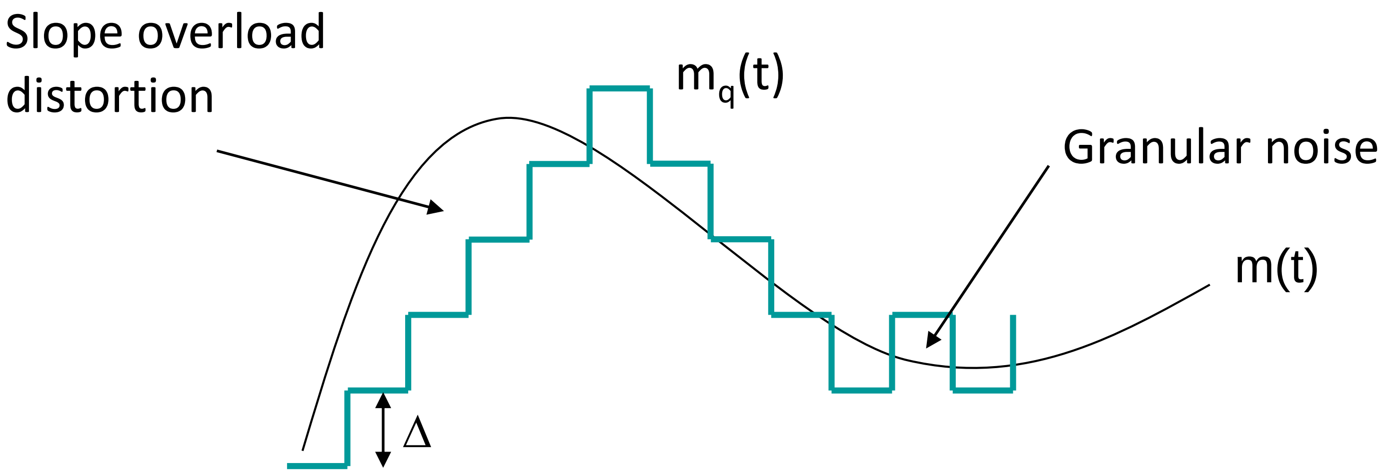
DM illustration



Binary sequence at modulator output

DM Quantisation Error

- Slope overload distortion
- Granular noise



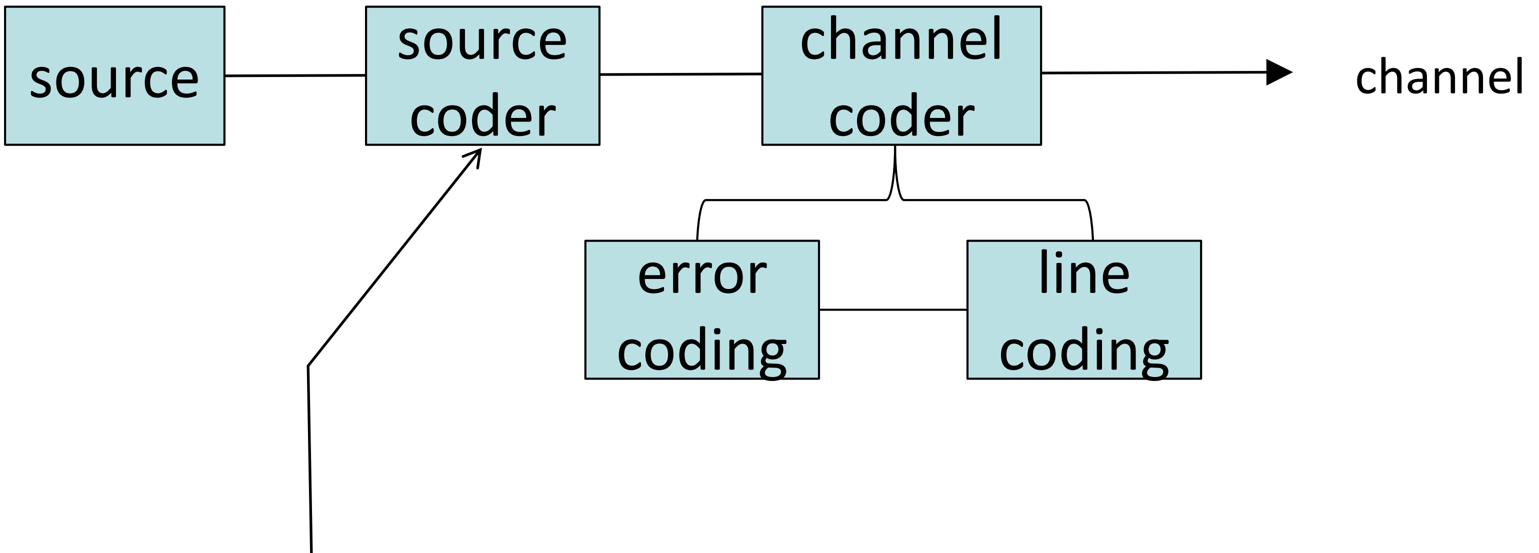
To minimise slope overload distortion $\frac{\Delta}{T_s} \geq \max\left|\frac{dm(t)}{dt}\right|$

Summary

- Linear vs non-linear quantization
- PAM
- PCM
- DM

INFORMATION THEORY

Coding Elements



Used to compress the source so that it uses less resources in transmission – use Information Theory to do this

Memoryless Source

- Probability of an event occurring does not depend on what came before.

- Produces *symbols* (e.g. A, B, C, D) from an *alphabet*.
- Probability of a particular symbol is fixed.

- Sum of probabilities is 1.

Source with Memory

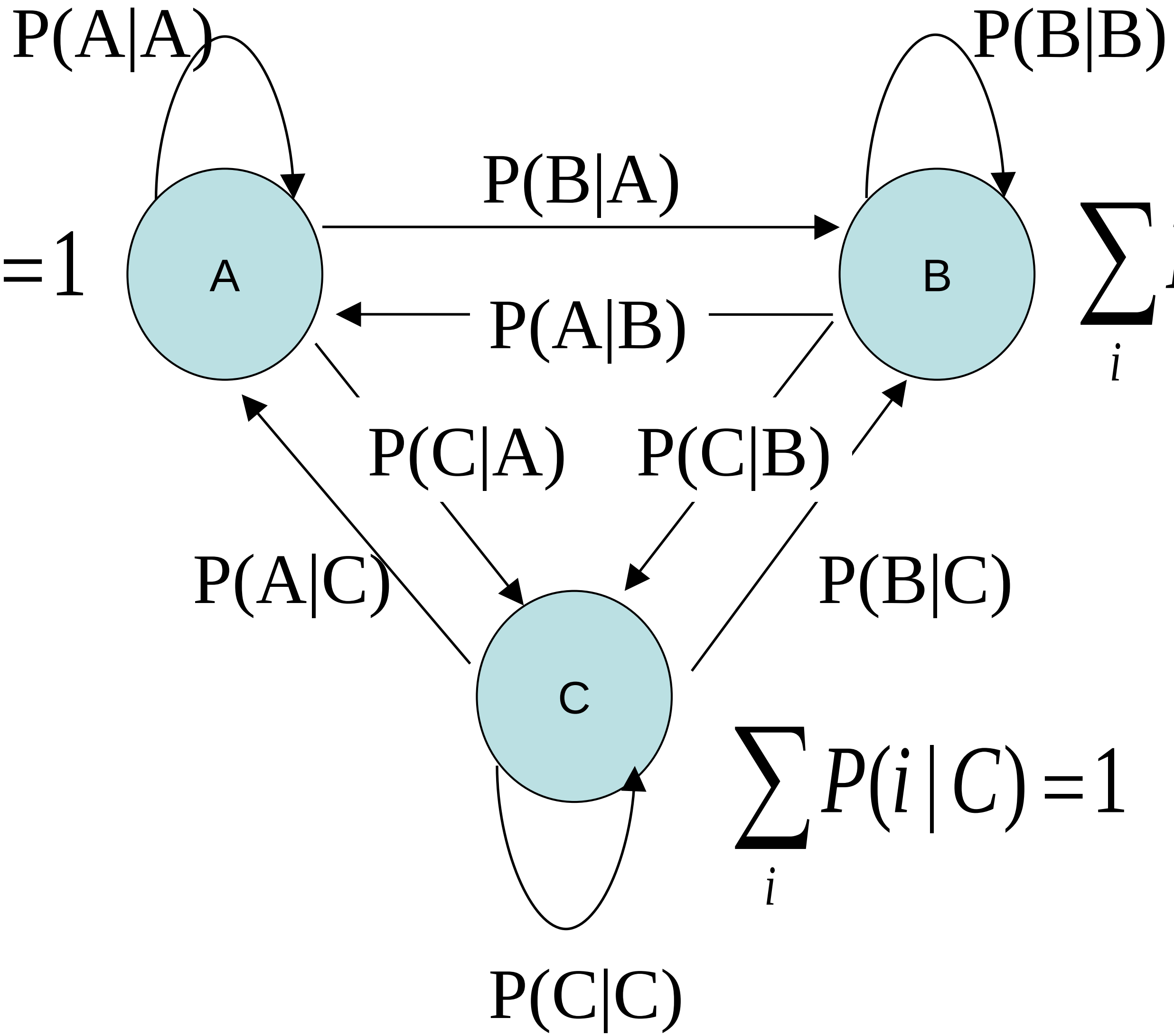
- Probability of a symbol depends on previous symbol.
- In English the probability of “u” occurring after “q” is higher than after any other letter.
- Lots of real sources in this category and coding like JPEG and MPEG exploits this to produce **smaller file** sizes.
- Can be very complex.

Probability Reminder

- P(A) is probability of event A
- P(B|A) is conditional probability – probability of event B occurring given that A has occurred.
- Joint probability $P(A \cap B)$ is probability that A and B occur – this is also written P(A,B)
- Here the notation P(A;B) is used to denote the probability of the event pair A followed by B – note this is not the same as P(B|A)

Probability Reminder

- Independent sources
  - $P(A \cap B) = P(A)P(B)$
    - condition for being statistically independent
  - $P(B|A) = P(B)$
- Conditional probability
  - $P(A \cap B) = P(A|B)P(B)$
  - $P(A|B) = P(A)P(B|A)/P(B)$



$$\sum_i P(i \mid A) = 1$$

$$\sum_i P(i \mid B) = 1$$

$$\sum_i P(i \mid C) = 1$$

$$P(A) = P(A|A)P(A) + P(A|B)P(B) + P(A|C)P(C)$$

$$P(A)(1 - P(A|A)) = P(A|B)P(B) + P(A|C)P(C)$$

And similarly for the others

$$P(B)(1 - P(B|B)) = P(B|A)P(A) + P(B|C)P(C)$$

$$P(C)(1 - P(C|C)) = P(C|A)P(A) + P(C|B)P(B)$$

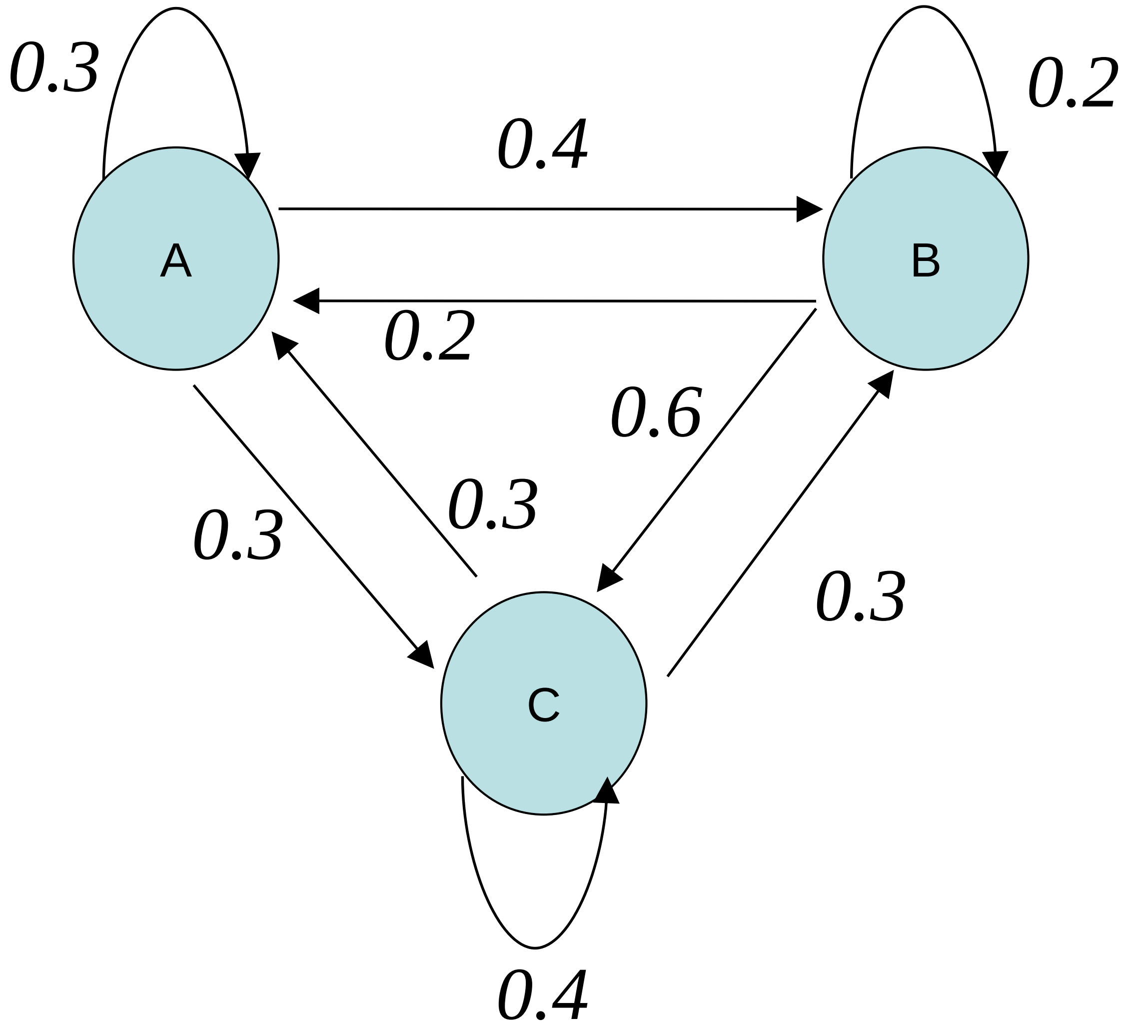
Also we must get one of the symbols so:

$$1 = P(A) + P(B) + P(C)$$

Taking this last equation plus 2 others gives 3 simultaneous equations that can be solved for P(A), P(B) and P(C)

Example

P(A|A)	0.3
P(B|A)	0.4
P(C|A)	0.3
P(A|B)	0.2
P(B|B)	0.2
P(C|B)	0.6
P(A|C)	0.3
P(B|C)	0.3
P(C|C)	0.4



$$P(A)(1-P(A|A)) = P(A|B)P(B) + P(A|C)P(C)$$
$$P(B)(1-P(B|B)) = P(B|A)P(A) + P(B|C)P(C)$$
$$P(C)(1-P(C|C)) = P(C|A)P(A) + P(C|B)P(B)$$

- 0 = - 0.7 P(A) + 0.2 P(B)+ 0.3 P(C)
- 0 = 0.4 P(A) - 0.8 P(B) + 0.3 P(C)
- 0 = 0.3 P(A) + 0.6 P(B) – 0.6 P(C)
- 1 = P(A) + P(B) + P(C)

To solve need to include this line plus 2 others

Solving gives:

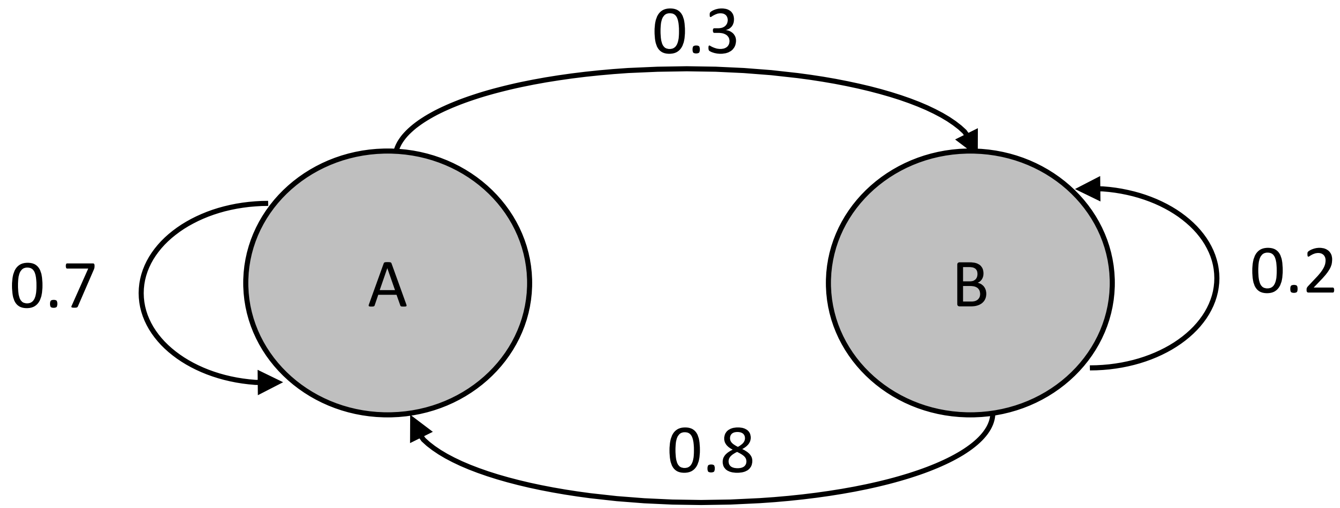
P(A)	0.2703
P(B)	0.2973
P(C)	0.4324

- Need to calculate the probabilities of pairs of symbols.
- $P(B;A) = P(A|B)P(B)$ and similarly for all other pairs

$P(A;A)=P(A\|A)P(A)$	0.0811
$P(A;B)=P(B\|A)P(A)$	0.1081
$P(A;C)=P(C\|A)P(A)$	0.0811
$P(B;A)=P(A\|B)P(B)$	0.0595
$P(B;B)=P(B\|B)P(B)$	0.0595
$P(B;C)=P(C\|B)P(B)$	0.1784
$P(C;A)=P(A\|C)P(C)$	0.1297
$P(C;B)=P(B\|C)P(C)$	0.1297
$P(C;C)=P(C\|C)P(C)$	0.1730
Total	**1**

Practice for Probability

A Markov process is shown in Figure below. Calculate the probability of event A occurring and event B occurring



$$P(A)=P(A)P(A|A)+P(B)P(A|B)$$
$$=0.7P(A)+0.8P(B)$$

$$0.3P(A)=0.8P(B)$$

$$P(A)+P(B)=1$$

$$P(A)*11/8=1$$
$$P(A)=8/11=0.727$$
$$P(B)=3/8=3/11=0.273$$

Information Theory

- What is information?
  - Everything has information.
  - Everywhere has information.

- Information in communication.
  - Signals: carry the information, a physical concept.
  - Symbols: describe the information by mathematics

Information Theory

- Measure of information is a **measure of uncertainty** - more certain data contains less information.

- Which has more information?
  - Now the weather outside is sunny .
  - You tell me it is sunny now. 100% happened
  - You tell me it is snowing now. Can’t happen

- A lie contains infinite information.

Information Theory

- Unit of information now generally in bits
  - p is the probability of the event and I is the information content.
  - Notice that we take logs to base 2 to get the units in bits.

$$I = \log_2(1/p)$$

  - Other bases give other units - e is Nat, 10 is Det (or Hartley)

- Remember that $\log_2(x) = \dfrac{\log_{10}(x)}{\log_{10}(2)}$

- Excel function LOG(number,base) takes logs to any base.

Entropy

- If a source emits a number of symbols, the **entropy** (H) is the *average information content per symbol:*

$$H = \sum_i p_i \, log_2(1/p_i)$$

p_i is the probability of the i'th event occurring

$$\sum_i p_i = 1$$

Example

- Determine the entropy of a source that transmits 4 symbols with the probability of each symbol being emitted being:

Symbol:	*A*	*B*	*C*	*D*
Probability	0.15	0.15	0.2	0.5

$$H = \sum p_i \log_2 (1/p_i)$$

$$H = 0.15\log_2(1/0.15) + 0.15\log_2(1/0.15) + 0.2\log_2(1/0.2) + 0.5\log_2(1/0.5)$$

which gives H=1.785 bit/symbol

Maximum Entropy

- This occurs **when all events have the same probability**.
- Max entropy for N symbols is

$$H = \sum_{i=1}^{N} p_i \log_2 \left(1/p_i\right)$$

$$\text{but } p_i = 1/N$$

$$\text{so } H = \sum_{\mathrm{i}=1}^{N} (1/N) \log_2 \left(N\right) = \log_2 \left(N\right)$$

- In the previous example this would be when the probability for each event is 0.25.
  - which would give H= 4* 0.25 $\log_2(4)$ = 2 bits/symbol

Source Coding

- Aims to **reduce the number of bits transmitted**.
- Previous example:

Symbol:	*A*	*B*	*C*	*D*
Probability	0.15	0.15	0.2	0.5

- 4 symbols means 2 bits required ($2^2 = 4$ combinations) if no account is taken of probability of event
- Entropy of 1.785 bit/symbol means that is all that is required
- So the coding efficiency is 1.785/2=0.875 (87.5%)
- How??? - many methods of source coding such as
  – Video compression (MPEG)
  – Think of other examples …..

Principle of Source Coding

- The event with high probability uses short code
- The event with low probability uses long code
- Consider the same example:

Symbol:	*A*	*B*	*C*	*D*
Probability (p_i)	0.5	0.25	0.125	0.125
code	0	10	110	111
Number of bits (n_i)	1	2	3	3

- H=1.75 bit/symbol
- Average code length $=\sum p_i n_i = 0.5 \times 1 + 0.25 \times 2 + 0.125 \times 3 \times 2 = 1.75$ bit/symbol

Huffman Coding

- Put events in DESCENDING order of probability
- Combine the lowest 2 and re-order
- Repeat until only one value of "1"
- At each split, mark top as "0" and bottom as "1" (or vice-versa as long as same throughout)
- Go backwards from final "1" to each symbol in turn

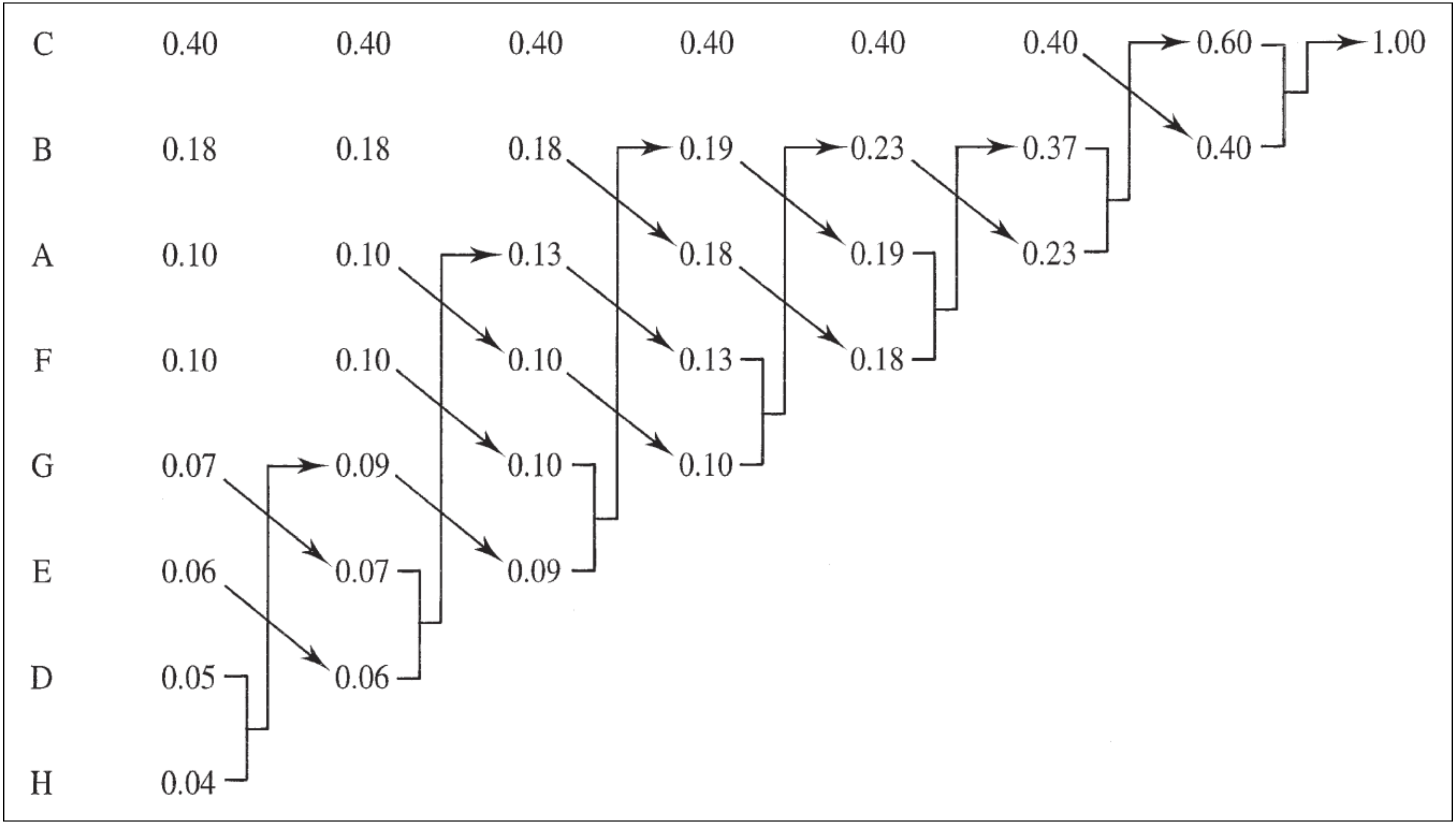
Huffman Coding

- Add 1/0 at each split as appropriate
- Write down:
  - pattern: these are bits transmitted
  - No. of bits for each symbol
  - probability*no of bits for each symbol (pi×Ni)
- $\Sigma\ (p_i \times N_i)$ is average no. of bits transmitted.
- Compression ratio= no. source coding bits / average number of bits transmitted.
- Code efficiency=H / average code length.

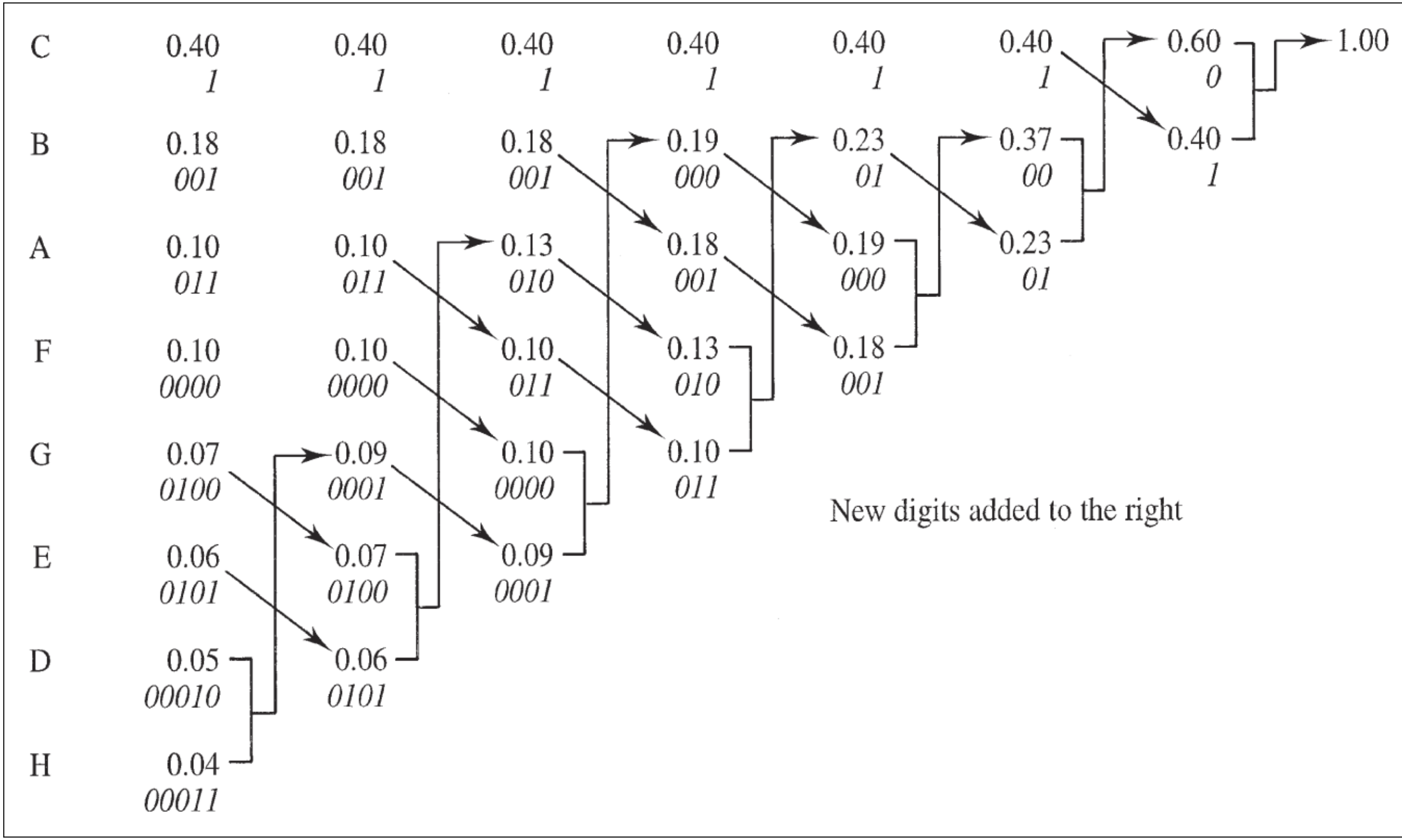
Huffman Coding Procedure (8-symbol)

Symbol	C	B	A	F	G	E	D	H
Probability	0.40	0.18	0.10	0.10	0.07	0.06	0.05	0.04
Codeword								

Huffman Coding – Reduction



Huffman Coding — Splitting



Huffman Calculation Example

- What is the source entropy *H*?
- What is the source efficiency $\eta_{source} = \frac{H}{H_{max}}$?
- What is the code efficiency $\eta_{code} = \frac{H}{L}$?

Determine the Huffman code for each Symbol below

Symbol (S)	S0	S1	S2	S3	S4
Probability (P_k)	0.25	0.25	0.125	0.3	0.075

Symbol (S)	S0	S1	S2	S3	S4
Probability (P_k)	0.25	0.25	0.125	0.3	0.075
Huffman Code	01	10	110	00	111
Huffman Code (v2)	10	01	001	11	000

Unique Decodability

- Variable length codes must be unique
- Codes A=0 B=011 C=110 are NOT unique
  - 0110 could be AC or BA
- Codes A=0 B=10 C=110 are unique but errors are a problem
  - 0110 can only be AC
  - 0110 (AC) corrupted to 0010 is read as 0 0 10 AAB

Summary

- Modern communications
- Channel impairment
- Sampling
- Quantization
- Information Theory