# EBU5303

# Multimedia Fundamentals

## Digitisation

Dr. Marie-Luce Bourguet
marie-luce.bourguet@qmul.ac.uk

# Learning Objectives

- Explain the differences between analogue and discrete signals.
- Discuss how digitisation can cause aliasing and quantisation errors.
- Calculate the signal-to-quantisation-noise ratio.
- Calculate the size of digital images, audio, and video files.

# Reading


http://burg.cs.wfu.edu/TheScienceOfDigitalMedia/Chapter1/Ch1ScienceOfDigitalMedia.pdf

**1.2 Analog to Digital Conversion**

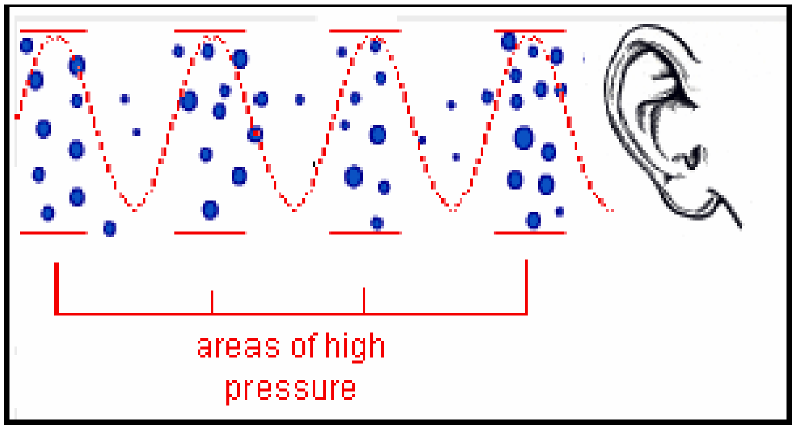
**1.3 Data Storage**

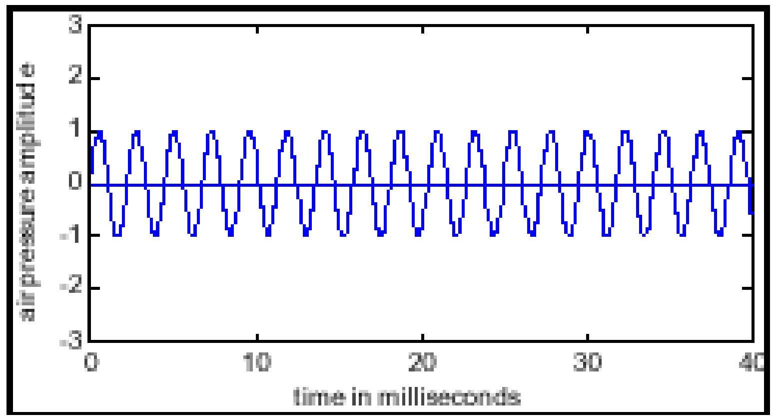
# Agenda

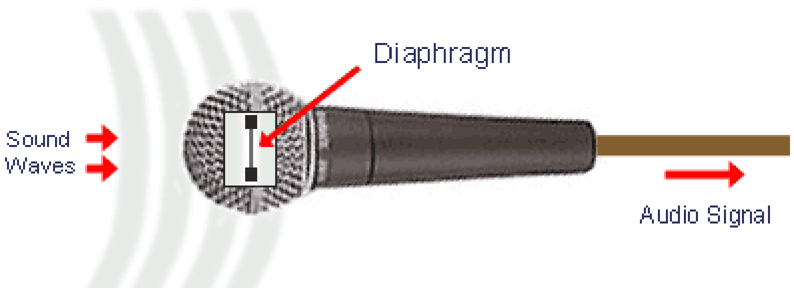
- Sound and images are analogue phenomena that can be represented by complex waveforms
- They must be digitised to be handled by computers: sampling and quantisation
- Sampling and quantisation rates determine the size of the digitised data

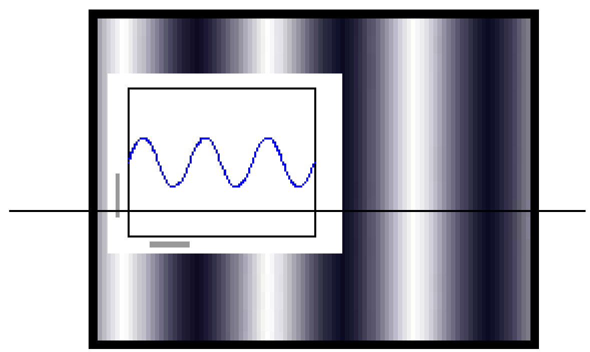
# Analogue versus discrete phenomena


# Sound and images in the analogue domain



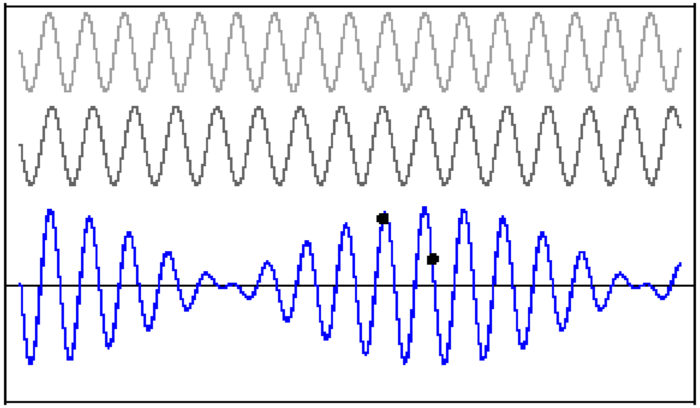



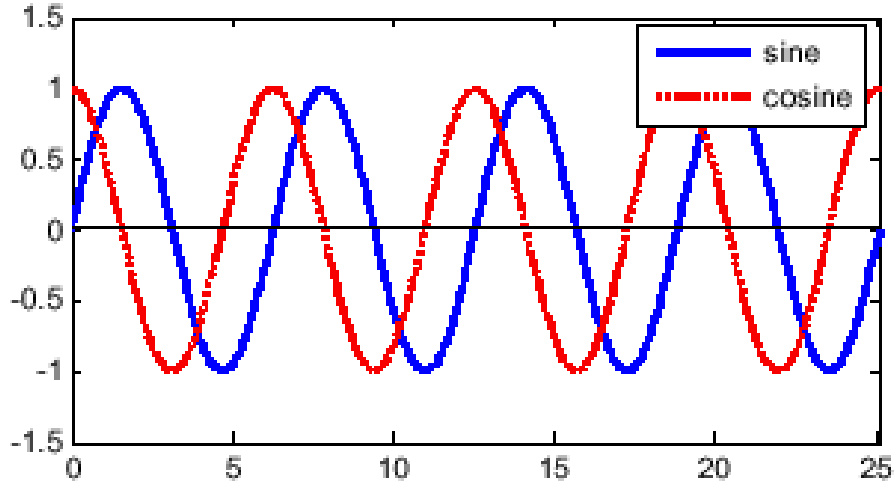



# Complex Waveforms

- Most waveforms are complex, i.e. their shape is the result of adding 2 or more waveforms.
- Each complex waveform may be described as the sum of a number of simple sine waves, each with a particular amplitude, frequency (or wavelength) and phase.



# Sinusoidal functions




- cycle
- wavelength (m)
- f = frequency (Hz)
- T = period (s) = 1/f
- A = amplitude
- $\omega$ = angular frequency (rad/s) = $2\pi f$
- $\phi$ = phase

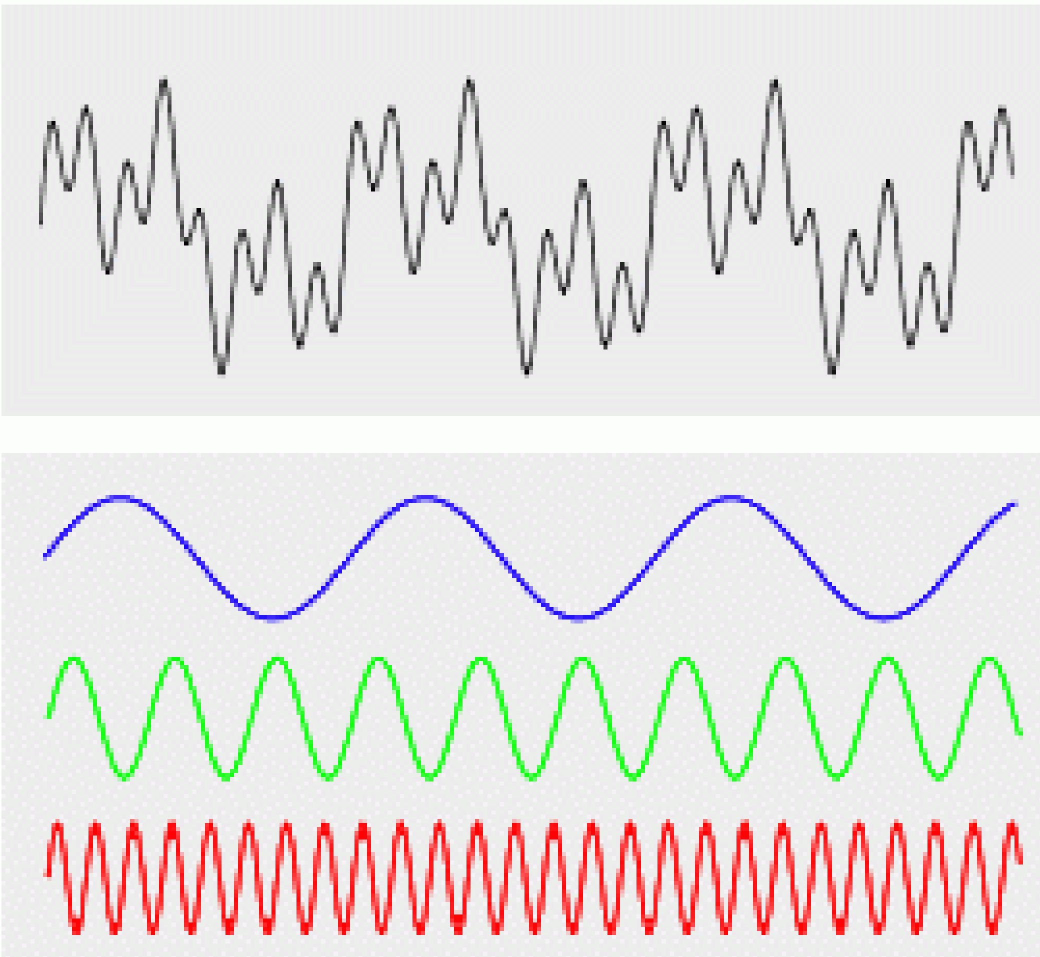
$$s(t) = A \sin(\omega t + \phi)$$

# Fourier analysis

- Fourier analysis attempts to represent a set of data with a series of sines and cosines with different periods, amplitudes, and phases.
- This representation is done by a mathematical process called a transform : the data measurements in the time domain are *transformed* into the period or frequency domain.

$$F(v) = \int f(t) exp(i2\pi vt) dt$$

# Fourier Analysis

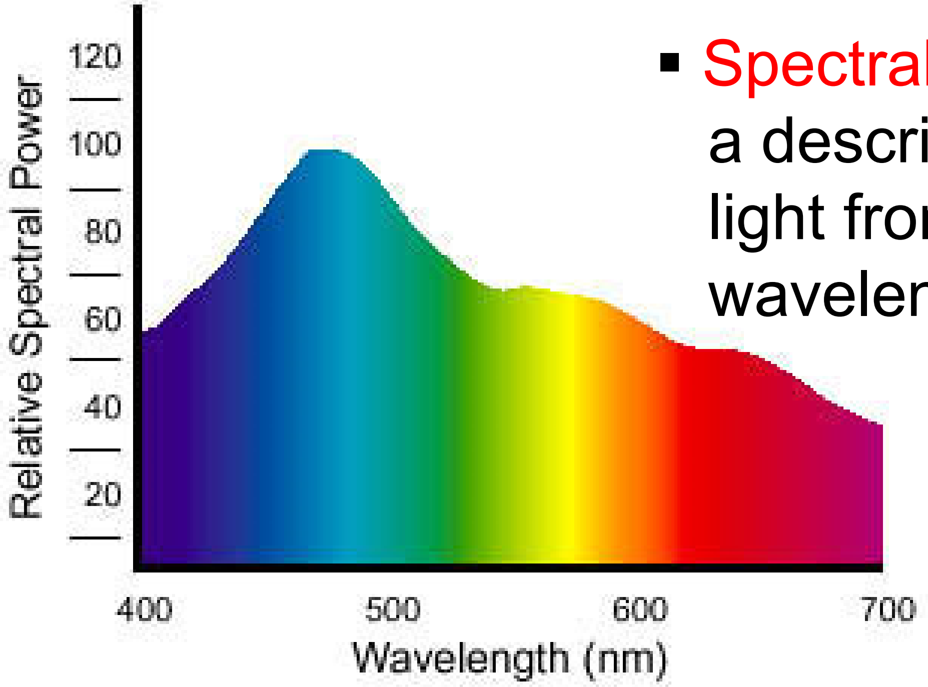


Fourier Analysis:

The complex wave (upper) can be decomposed into the sum of the three simple sine waves (lower).

# Example: Visible Light

- The wavelengths of visible light lie roughly between 400 nm and 700 nm.

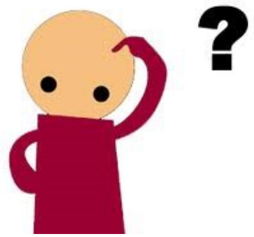



- Spectral Power Distribution (SPD): a description of how the intensity of light from some source varies with wavelength.

# Summary so far …

- Sound and light (colours and images) are analogue phenomena
- They can be represented by complex waveforms
- Individual components (sine waves) can be extracted from a complex waveform by Fourier analysis

# Questions…

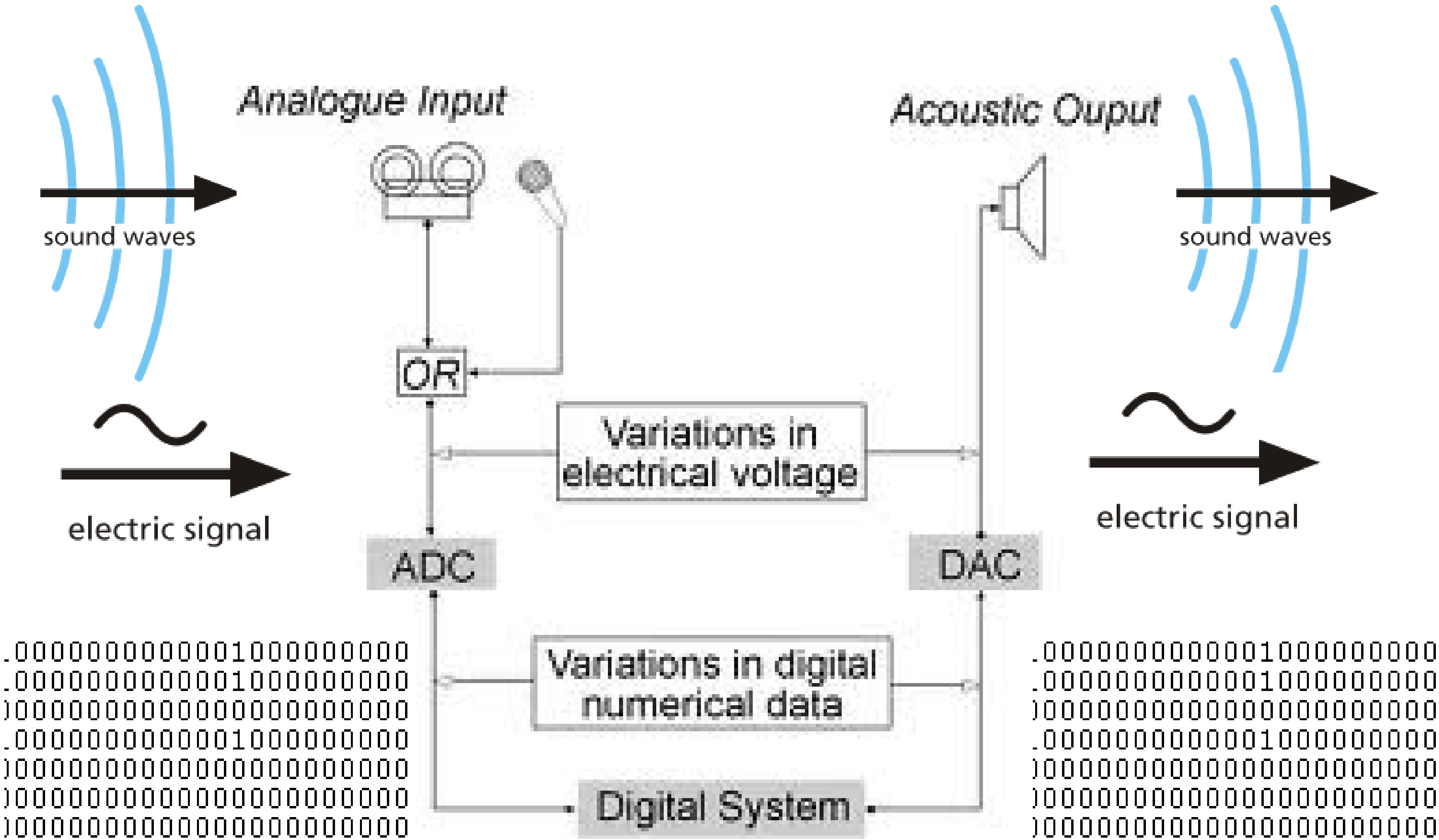



- What kind of sound is represented by a completely regular sine wave ?
- What does the frequency of a wave tell us about the sound it represents ?
- What does the amplitude of a wave tell us about the sound it represents ?
- What information is contained in the wavelength of a light ray ?

# Agenda

- Sound and images are analogue phenomena that can be represented by complex waveforms
- They must be digitised to be handled by computers: sampling and quantisation
- Sampling and quantisation rates determine the size of the digitised data

# Analogue-to-digital conversion



# Analogue-to-digital conversion

= converting the continuous phenomena of images, sound, and motion into a discrete representation that can be handled by a computer.

- Digitised pictures and sound can now be captured in fine detail.
- Digital data communication is less vulnerable to noise than is analogue.
- Digital data can be communicated more compactly than analogue… when compressed!

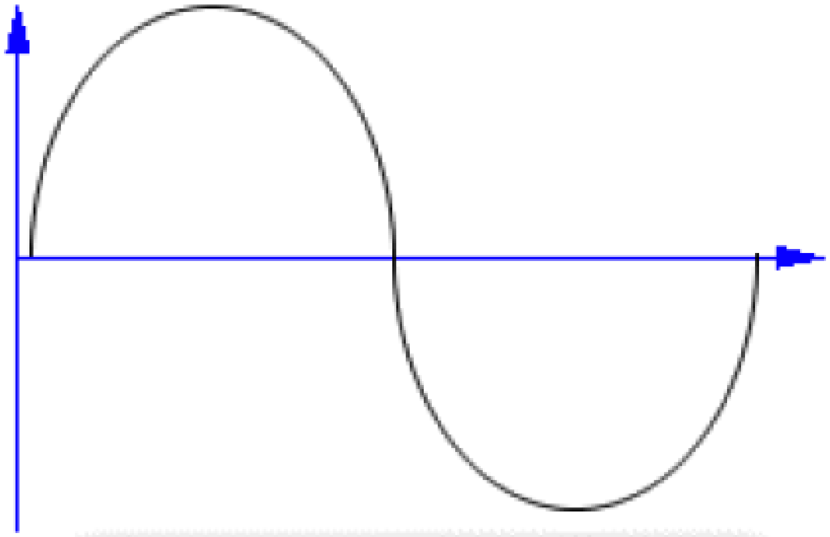
# Personal Research

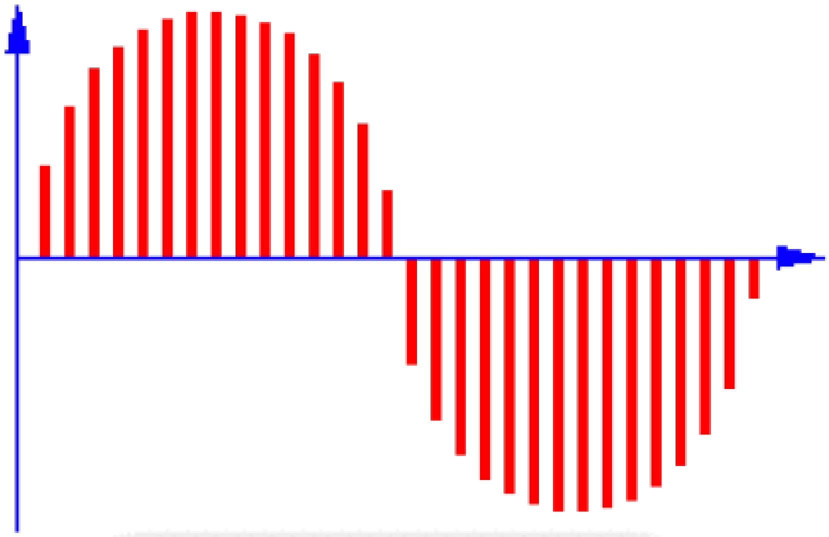
What are the analogue to digital converters available?

# Sampling

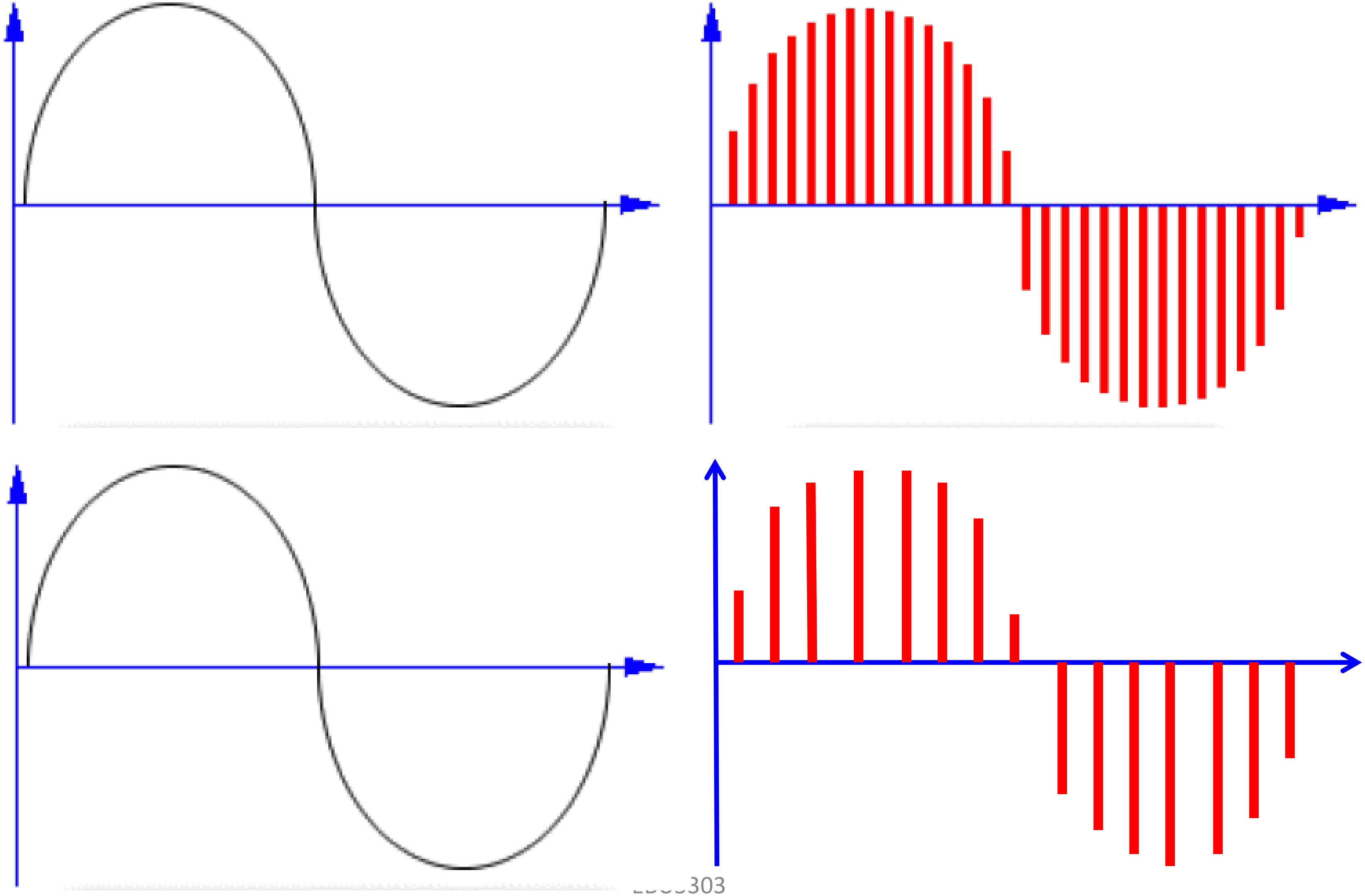
- Sampling chooses discrete points at which to measure a continuous phenomenon (a *signal*).
- In the case of images, the sample points are evenly separated in space.
- In the case of sound, the sample points are evenly separated in time.
- The number of samples taken per unit time or unit space is called the ***sampling rate.***

# Sampling





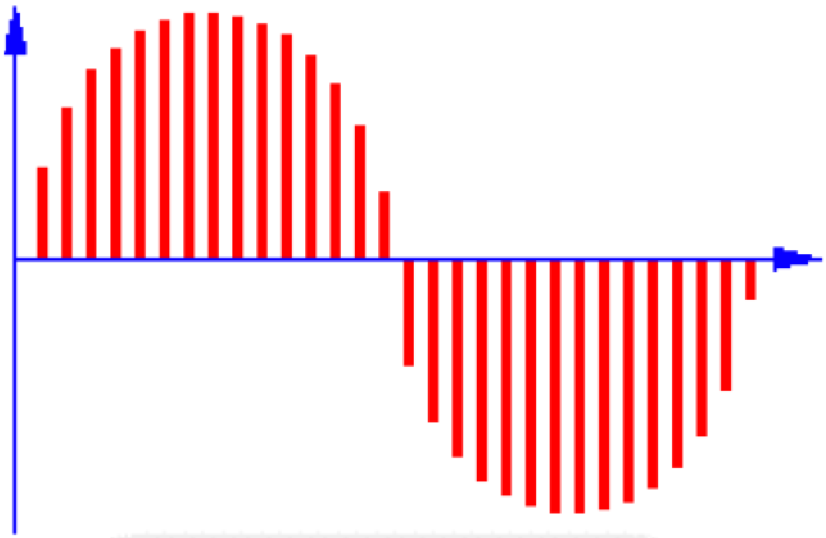

# Sampling

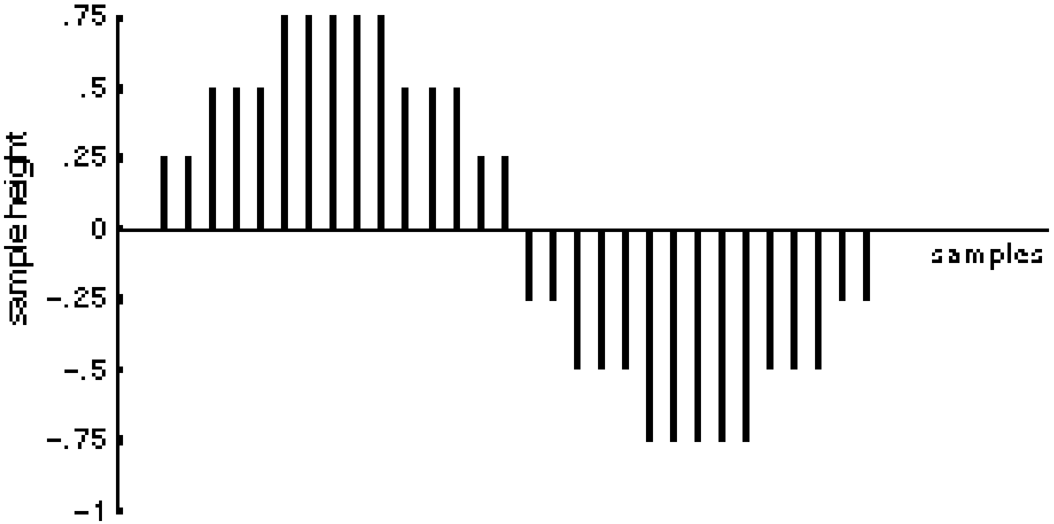


# Quantisation

- Quantisation requires that each sample be represented in a fixed number of bits, called the ***sample size*** or the ***bit depth***.
- In the case of an image, the bit depth is called the ***colour depth****.*
- The bit depth limits the precision with which each sample can be represented.

# Quantisation






Let *n* be the number of bits used to quantize a digital sample. Then the maximum number of different values that can be represented, m, is

$$m = 2^n$$

# Exercise

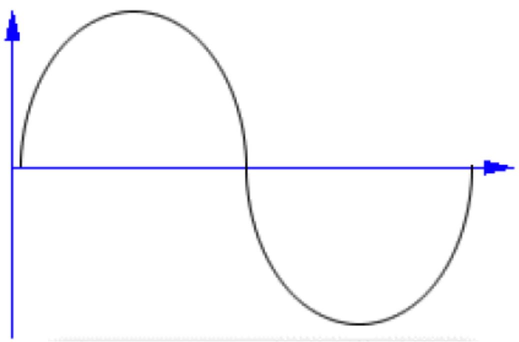
The number of possible colours in an image is determined by the quantisation rate. How many different colours can be represented with 12 bits?

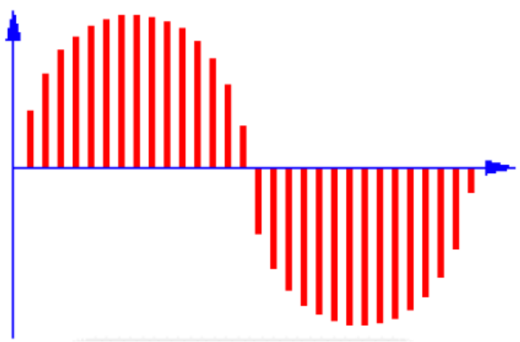
# Exercise

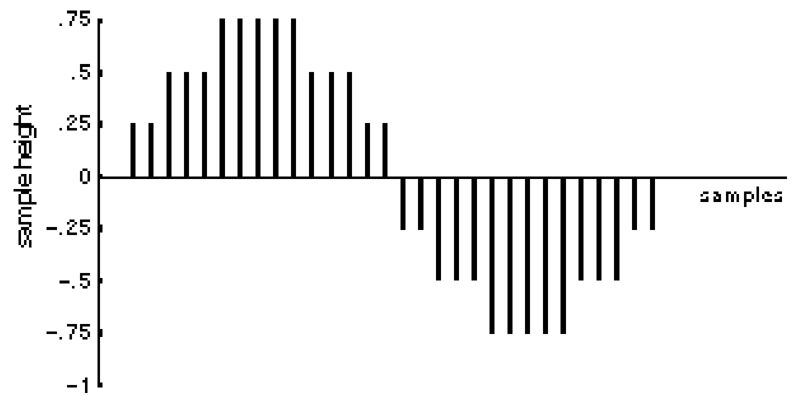
Prove that, if we double the number of bits used to hold a quantised value, then we square the number of quantisation levels.

# Sampling + Quantisation

- The sampling rate relates directly to the frequency of a wave.
- Quantisation, on the other hand, relates more closely to the amplitude of a wave.





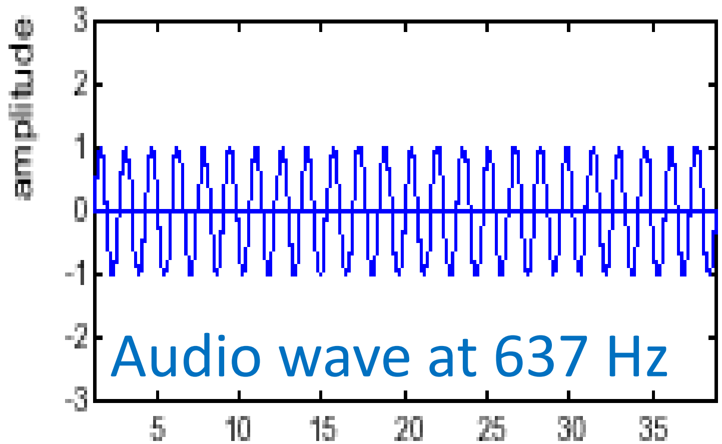


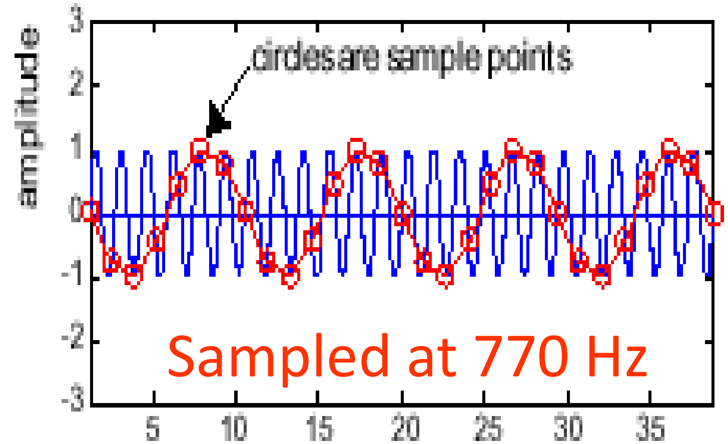
# Aliasing (sampling error)

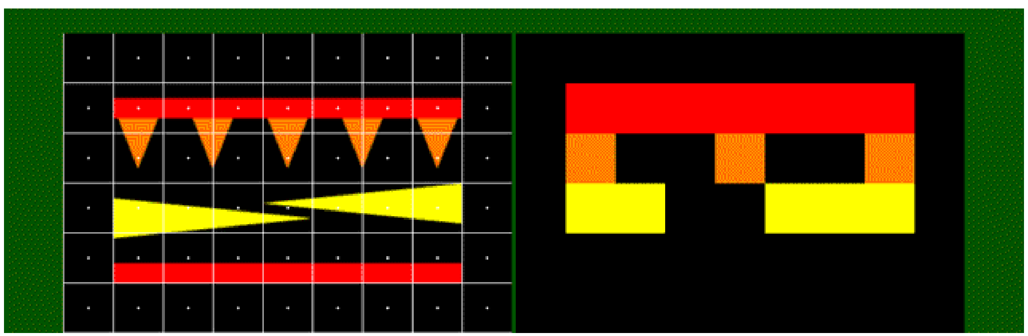
Aliasing refers to a phenomenon that occurs when a continuous signal is under sampled, resulting in distortion or misleading representations of the original signal.

When aliasing occurs, high-frequency components in the original signal can appear as lower-frequency components in the sampled or digitised representation. This is because the lower sampling rate fails to capture the rapid changes or details in the signal, resulting in distortions.

# Aliasing (sampling error)



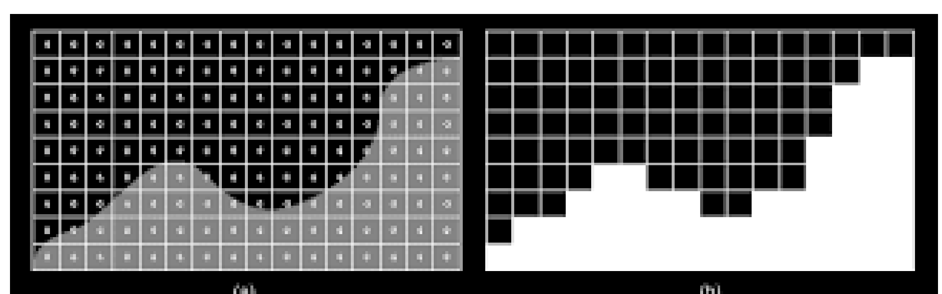


Image of girl | Reduction by 4 | Reduction by 16 | Reduction by 32

# Nyquist theorem

Sample twice as often as the highest frequency you want to capture.

Let $f$ be the frequency of a sine wave. Let $r$ be the minimum sampling rate that can be used in the digitisation process such that the resulting digitised wave is not aliased. Then:

$$r = 2f$$

$r$ is called the ***Nyquist rate***.

# Exercise

If you are recording an audio file and you expect that the highest frequency in the file will be 10,000 Hz, what is the minimum sampling rate you should use to ensure that you will not get audio aliasing?

# Nyquist rate and Nyquist frequency

- The Nyquist rate is the minimum rate at which a continuous signal needs to be sampled in order to accurately reconstruct it without aliasing. It is defined as twice the maximum frequency contained in the signal.

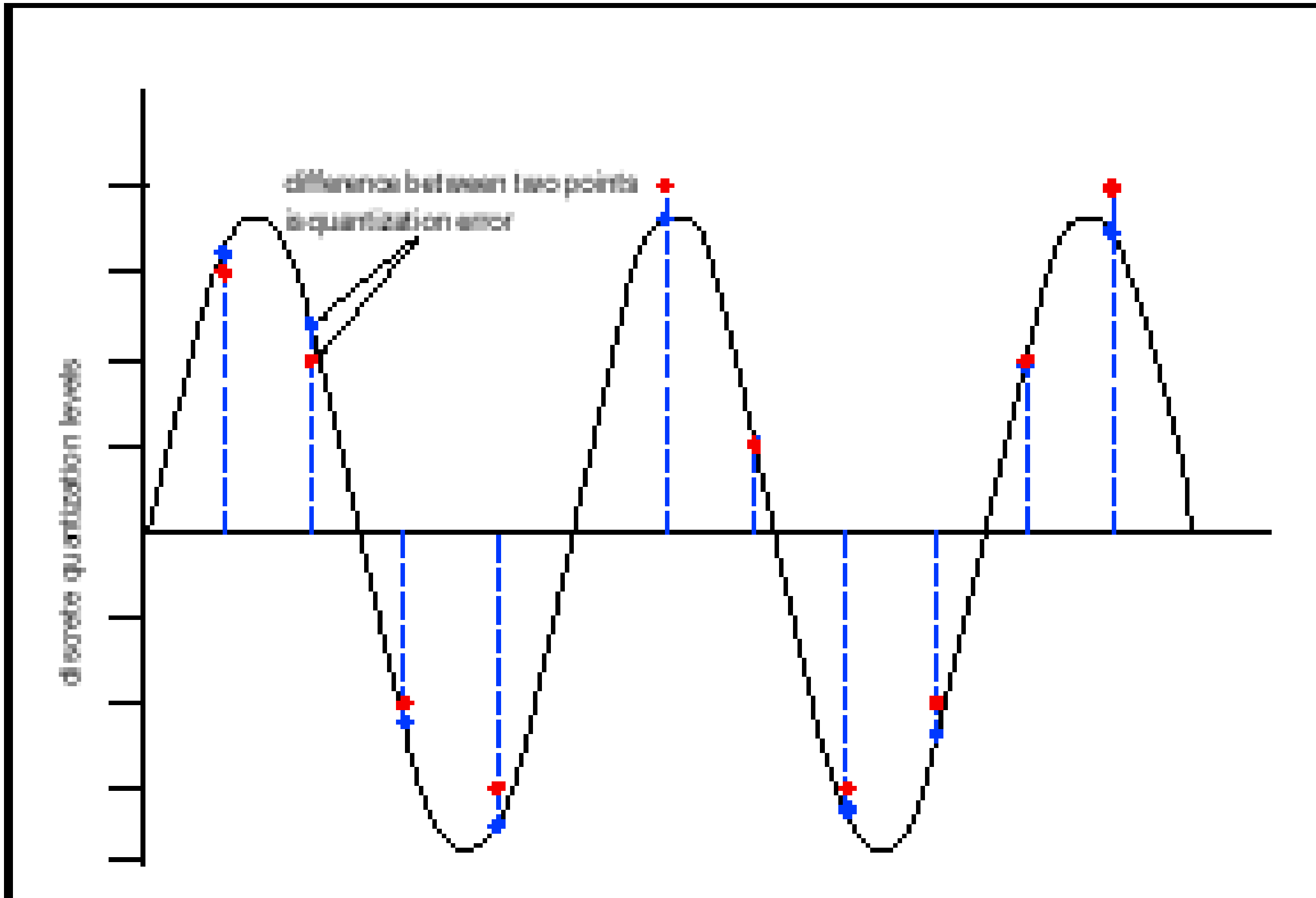
- The Nyquist frequency represents the maximum frequency that can be accurately represented or reconstructed from the sampled signal without aliasing. It is half the sampling rate.

- Nyquist rate = 2 * Nyquist frequency

# Quantisation error

- Quantisation error (also known as quantisation noise) arises because the actual continuous signal can have infinitely many values between two adjacent quantisation levels, but it is approximated by assigning a discrete value.
- The quantisation error is the difference between the actual value of the continuous signal and the assigned quantised value.

# Quantisation error



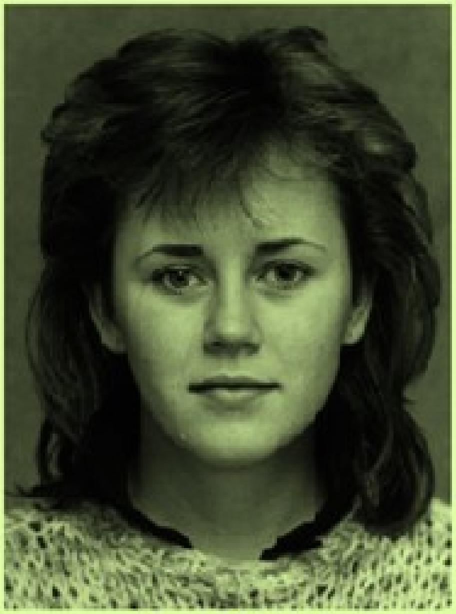


Image of girl: 256 levels

8 levels

4 levels

2 levels

# Signal to Noise Ratio (SNR)

- SNR: the ratio of the meaningful content of a signal versus the associated noise. It is expressed in dB (decibels)

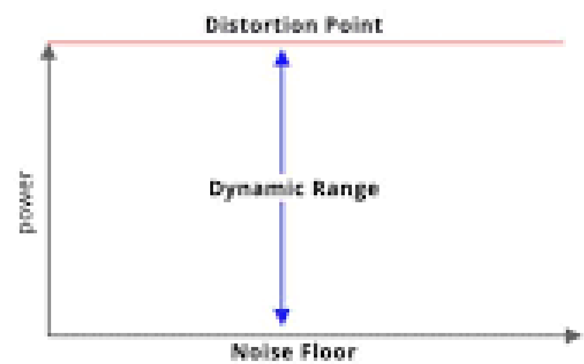
- In analogue data communication, SNR is defined as the ratio of the average power in the signal versus the power in the noise level.

- For a digitised image or sound, SNR, also called ***signal-to-quantisation-noise ratio*** (***SQNR***), is defined as the ratio of the maximum sample value over the maximum quantisation error.

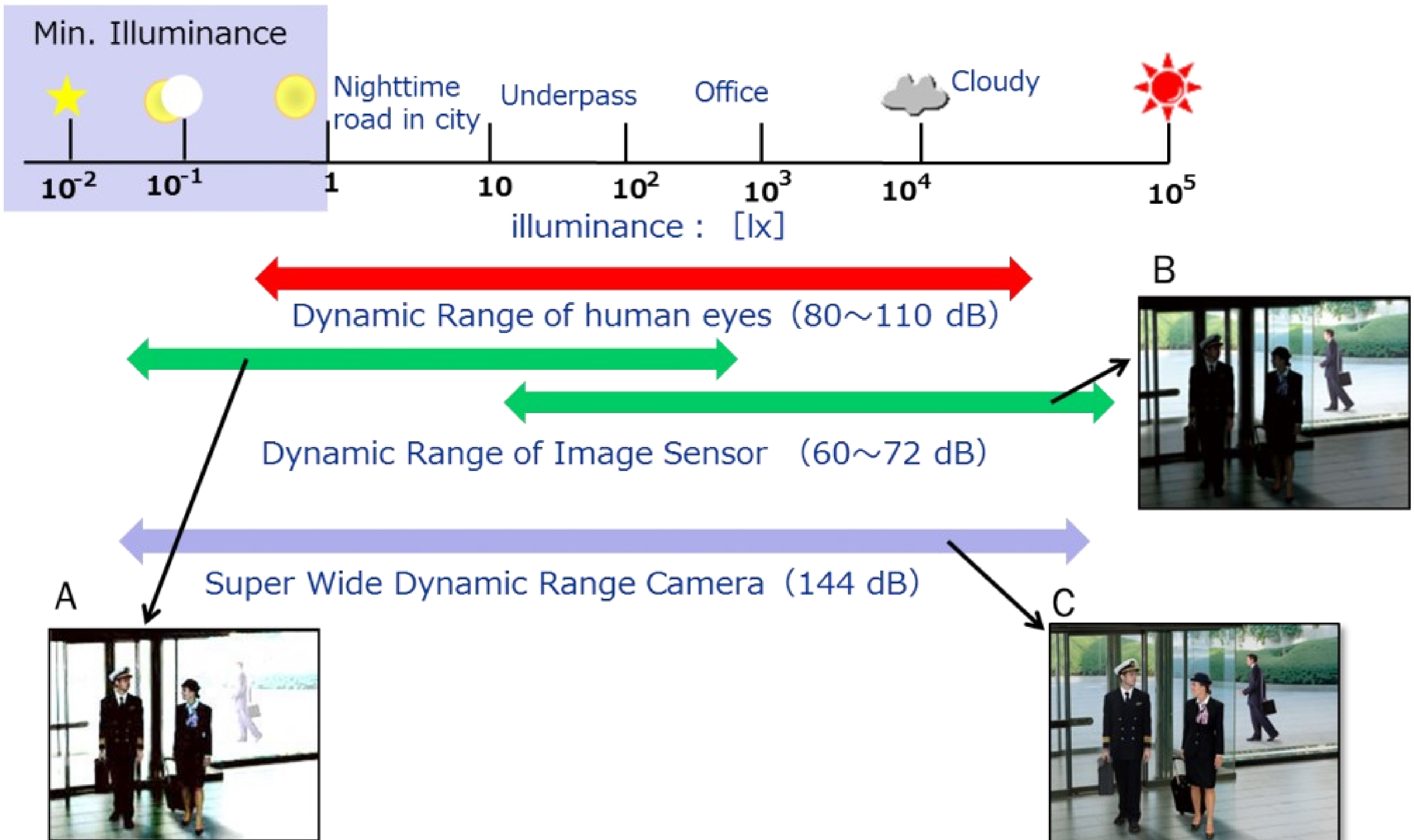
$$SQNR = 20\log_{10}\left(\frac{\max(quantization\ value)}{\max(quantization\ error)}\right)$$

# Dynamic range

- The dynamic range of a signal refers to the difference between the highest and lowest amplitudes that can be accurately represented or measured within that signal.
- For audio signals, dynamic range refers to the span between the softest and loudest sounds in a recording. It is often expressed in terms of the difference between the noise floor (the quietest detectable sound level) and the maximum undistorted signal level.
- In imaging, dynamic range refers to the difference between the darkest and brightest parts of an image.



# Dynamic range



# SQNR and dynamic range

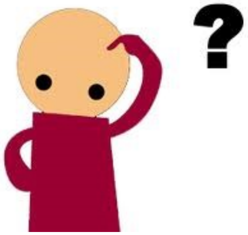
- SQNR is directly related to ***dynamic range:*** the ratio of the largest sound amplitude (or colour, for images) and the smallest that can be represented with a given bit depth.
- With a higher bit depth, the quantisation levels are closer together, resulting in a smaller quantisation error and, therefore, lower quantisation noise.

Let *n* be the bit depth of a digitised media file (e.g. digital audio). Then the signal-to-quantisation noise ratio *SQNR* is:

$$\boldsymbol{SQNR = 20 \log_{10}(2^n)}$$

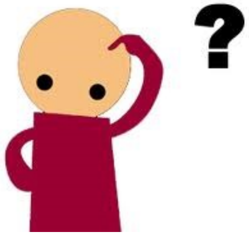
Approximation: SQNR (dB) ≈ 6 * n

# Question




Estimate the SQNR of a sound file encoded using 8 bits per sample.

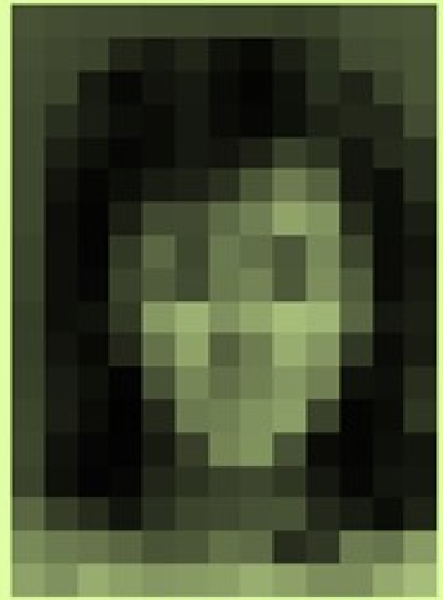
# Question



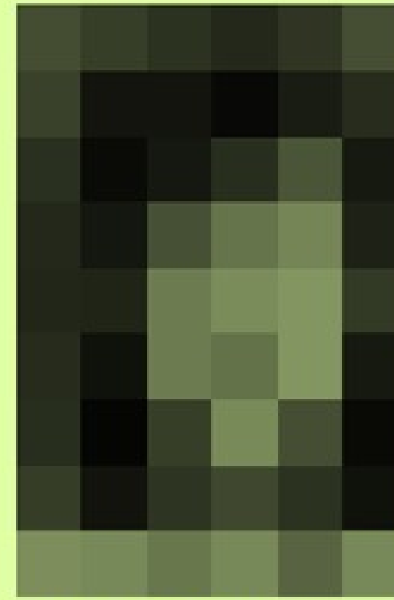

You have a digital image in 8 bit colour which lacks the subtle gradations from one colour to the next that you would like to see. Is this a matter of aliasing or quantisation error? Explain.

Image of girl

Reduction by 4



Reduction by 16



Reduction by 32

Image of girl: 256 levels

8 levels

4 levels

2 levels

# Summary so far …

- Analogue sounds and images must be digitised
- Digitisation includes two processes: sampling and quantisation
- Undersampling causes aliasing (the original data cannot be reproduced)
- Quantisation errors cause loss of details and determine the Signal to Quantisation Noise ratio
- The Nyquist theorem helps us choose an adequate sampling rate

# Agenda

- Sound and images are analogue phenomena that can be represented by complex waveforms
- They must be digitised to be handled by computers: sampling and quantisation
- Sampling and quantisation rates determine the size of the digitised data

# Data size (example): audio file

Sampling rate: 44.1 kHh (44,100 samples/s)
Bit depth: 32 bits per sample (16 for each of two stereo channels)
Number of minutes: 1 minute
Total number of bits:

44,100 * 32 * 60 = 84,672,000 bits
84,672,000 bits = 10,584,000 bytes

File size: ~ 10.58 Mbytes for 1 minute
Data rate of the file: ~ 1.41 Mbits/s

# Data size : image file

Example:
Sampling: 1024 pixels x 768 pixels (samples)
Bits per pixel: 24

1024 x 768 = 786,432 pixels
786,432 x 24 = 18,874,368 bits

18,874,368 / 8 = 2,359,296 bytes
File size ~ 2.36 MB

# Data size : video file

Sampling: 720 pixels x 480 pixels
Bits per pixel: 24
Frame rate : 30 frames/s
Length: 1 minute
Audio : 44,1 KHz, 32 bits

pixels per frame : 720 x 480 = 345,600 pixels
bits per frame : 345,600 x 24 = 8,294,400 bits
Video images size/s : 8,294,400 x 30 = 248,832,000 bits
Total video image size : 248,832,000 x 60 = 14,929,920,000 bits

# Data size : video file

Sampling: 720 pixels x 480 pixels
Bits per pixel: 24
Frame rate : 30 frames/s
Length: 1 minute
Audio : 44,1 KHz, 32 bits

Audio size/s : 44,100 x 32 = 1,411,200 bits
Total audio size : 1,411,200 x 60 = 84,672,000 bits

# Data size : video file

Sampling: 720 pixels x 480 pixels
Bits per pixel: 24
Frame rate  : 30 frames/s
Length: 1 minute
Audio : 44,1 KHz, 32 bits

Audio size/s : 44,100 x 32 = 1,411,200 bits
Total audio size : 1,411,200 x 60 = 84,672,000 bits

Total video size : images + audio = 15,014,592,000 bits
File size ~ 1.88 GB   !!!

# Exercise

Compute the number of bytes needed for 1 minute of video that has 720 x 576 pixels per frame, 25 frames per second, 3 bytes per pixel, and CD-quality stereo audio.

# Storage Media

| Storage medium | Maximum capacity |
|---|---|
| **Portable Media** | |
| CD (Compact Disk) | 700 MB |
| DVD (Digital Versatile Disc or Digital Video Disk), standard one sided | 4.7 GB standard; 8.5 GB dual-layered |
| DVD video or high capacity | 17–27 GB |
| Memory stick or card | 8 GB |
| HD-DVD (High Definition DVD), standard one-sided | 15 GB standard; 30 GB dual-layered |
| Blu-ray Disk | 25 GB standard; 50 GB dual-layered |
| Flash drive | 64 GB |
| **Permanent Media** | |
| Hard disk drive | 1 terabyte (1000 GB) |

# Personal Research

Digital storage media and their capacities are evolving very fast…

Currently, what are the most common types of digital storage media and their typical capacities?

# Summary

- Sound and images are analogue phenomena that can be represented by complex waveforms
- They must be digitised to be handled by computers: sampling and quantisation
- Sampling and quantisation rates determine the size of the digitised data

# Problem

You have found a large box full of old photographs in your grandparents' attic. You decide to digitise the photographs so you can easily share them with the rest of your family.

Q1: In this specific scenario, what other advantages are there of digital images over the analogue photographs?

# Problem

Q2: You counted 1230 photographs! What would be the most practical method for digitising that many photographs (think quality, time and cost)?

Q3: How/where will you store the digital images? (consider how much data is going to be generated …)