Cambridge IGCSE

Computer Science

Section 1

2

Text, Sound and Images

Unit 1: Data representation

Objectives

- Understand how and why a computer represents text and the use of character sets - ASCII and Unicode
- Understand how and why a computer represents sound (including the effects of sample rate and sample resolution)
- Understand how and why a computer represents an image (including the effects of colour depth and resolution)
- Calculate the file size of an image file and a sound file

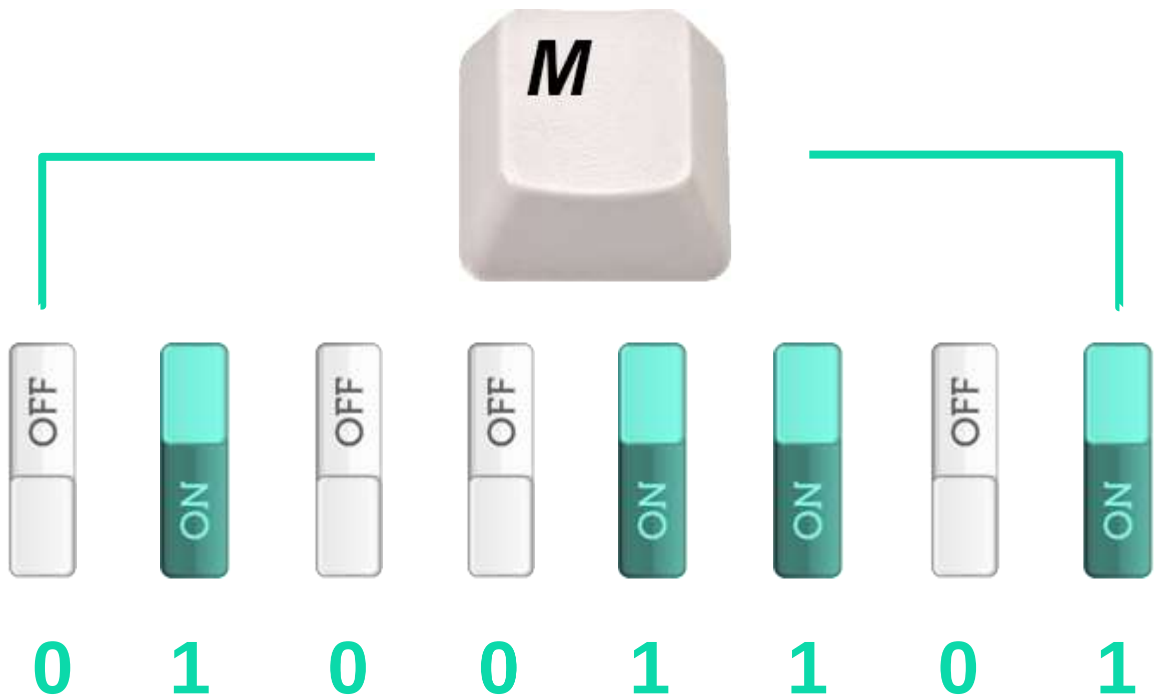
Text, sound and images

Representing text

- We have seen that numbers can be represented in binary
- But what about letters?
- A different system is needed

Representing text

- If a computer understands only 1s and 0s, what happens when the ‘M’ key is pressed on the keyboard?



Representing text

- Every character on the keyboard is represented by a binary value
- Uppercase letters (capitals) have different values from lowercase characters
- Punctuation symbols (!@”-) also have their own binary values
- How many characters are there on a standard keyboard? How many bits would be required to represent this many combinations?

Text - Character sets

A character set shows the characters and symbols that can be represented by a computer system and their binary value.

Each character and symbol is assigned a unique binary value.

We have two main character sets - ASCII and Unicode.

The ASCII code

- **ASCII** (American Standard Code for Information Interchange) became a standard code
- It was originally developed in the 1960s for representing the English alphabet
- It encodes 128 characters into 7-bit binary codes
- Characters include numbers 0 to 9, uppercase and lowercase letters A-Z, a-z, punctuation symbols and a space character

ASCII Table

- In a word processor or Notepad, try pressing ***ALT+65*** on your keyboard (hold down the ***ALT*** key while you type 65 using the numeric keypad on the right of the keyboard)
- Try typing your initials using ***ALT*** key combinations

Decimal	Binary	Character	Decimal	Binary	Character	Decimal	Binary	Character
32	00100000	space	64	01000000	@	96	01100000	'
33	00100001	!	65	01000001	A	97	01100001	a
34	00100010	"	66	01000010	B	98	01100010	b
35	00100011	£	67	01000011	C	99	01100011	c
36	00100100	$	68	01000100	D	100	01100100	d
37	00100101	%	69	01000101	E	101	01100101	e
38	00100110	&	70	01000110	F	102	01100110	f
39	00100111	'	71	01000111	G	103	01100111	g
40	00101000	(	72	01001000	H	104	01101000	h
41	00101001	)	73	01001001	I	105	01101001	i

ASCII representation of numbers

- What is the binary representation of the ASCII character 7? Is this the same as the binary value for 7?
- Why not? What does this mean? **# 37**

Decimal	Binary	Character	Decimal	Binary	Character	Decimal	Binary	Character
48	00110000	0	53	00110101	5	58	00111010	:
49	00110001	1	54	00110110	6	59	00111011	;
50	00110010	2	55	00110111	7	60	00111100	<
51	00110011	3	56	00111000	8	61	00111101	=
52	00110100	4	57	00111001	9	62	00111110	>

Using different alphabets

- To represent other alphabets for different languages, a new code allowing for many more characters was needed
- **Unicode** was developed to use 16 bits, giving 65,536 possible combinations – enough to represent every character in every language
- Example

Владимир Путин

Text - summary

- Text is converted into binary to be processed by a computer.
- Characters are represented using the **ASCII** or **Unicode** character sets.
- A character set is a list of characters, and their binary values, that can be represented by a computer system.
- Seven bits ASCII is enough to represent all the characters and symbols on an English keyboard
  7 bits -> 128 characters

Text - summary

- The extended 8-bit ASCII code allows for an extra 128 special characters such as ©, ®, á
  8 bits -> 256 characters

- Unicode provides a unique binary code for every character in every language
  up to 32 bits -> over 4 billion characters

- The Unicode character set allows for more characters and symbols than ASCII, including different languages and emojis.

- The Unicode character set creates a universal standard that covers all languages and writing systems.

Sound

Sounds must be converted into a digital form in order to be stored and processed by a computer.

Analogue to Digital

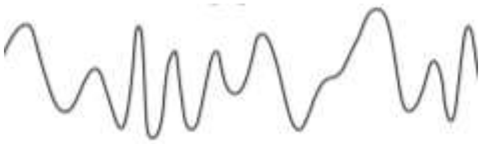
Digital to Analogue

Storing sound

Just like with numbers, text and images, sound is also stored in digital form.

Sound, when produced is in analogue form

- a continuous wave of varying data.

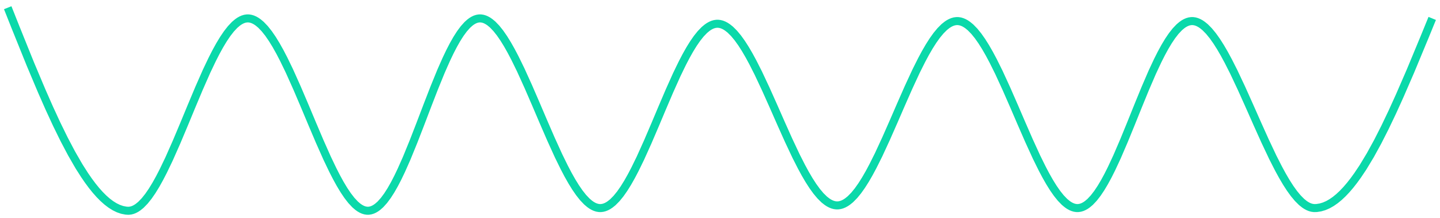


To store sound in computer, it has to be sampled

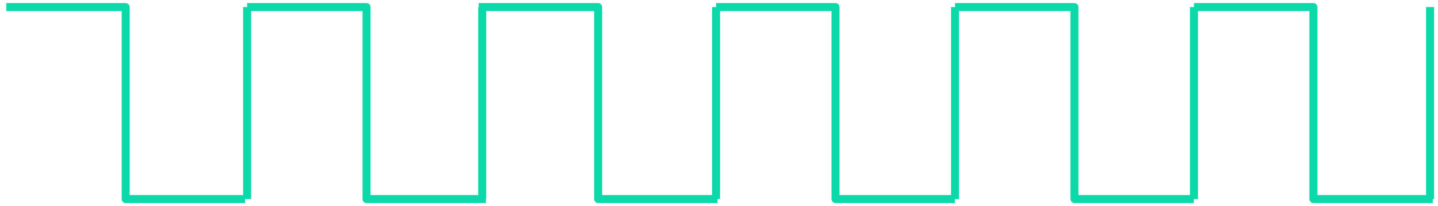
- obtain thousands of samples per second, where each sample corresponds to a binary value.

Analogue to Digital conversion

- Analogue sound signals are in a continuous wave



- Digital signals are discrete



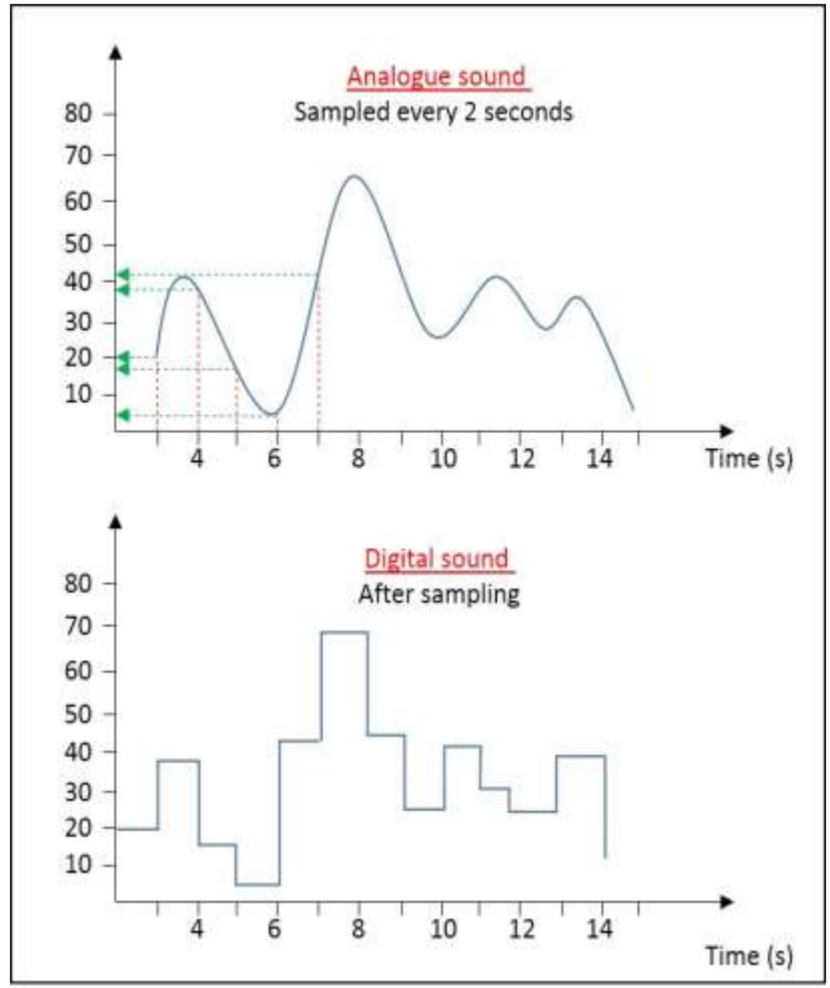
- Sound is digitized by sampling the sound wave thousands of times per second and converting the samples to binary

Sound sampling

Sampling is the process that converts analogue sound into digital data that can be stored in computer.

An analogue sound signal is shown.

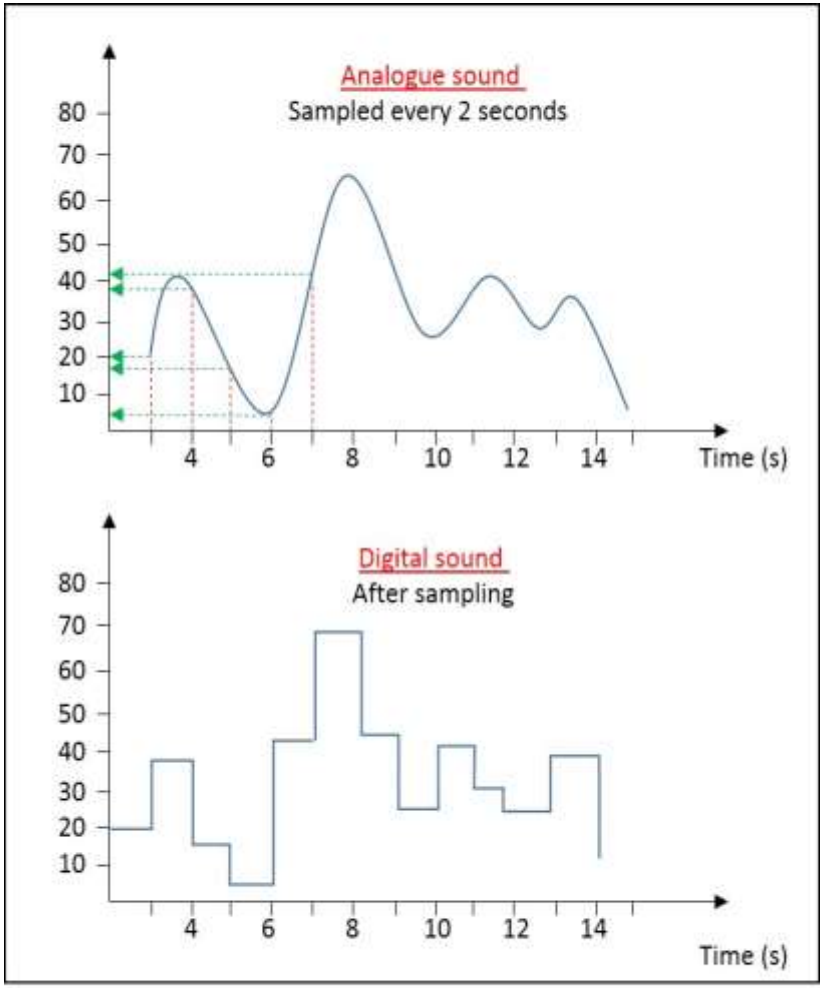
This sound is sampled every two seconds to obtain discrete digital sound.



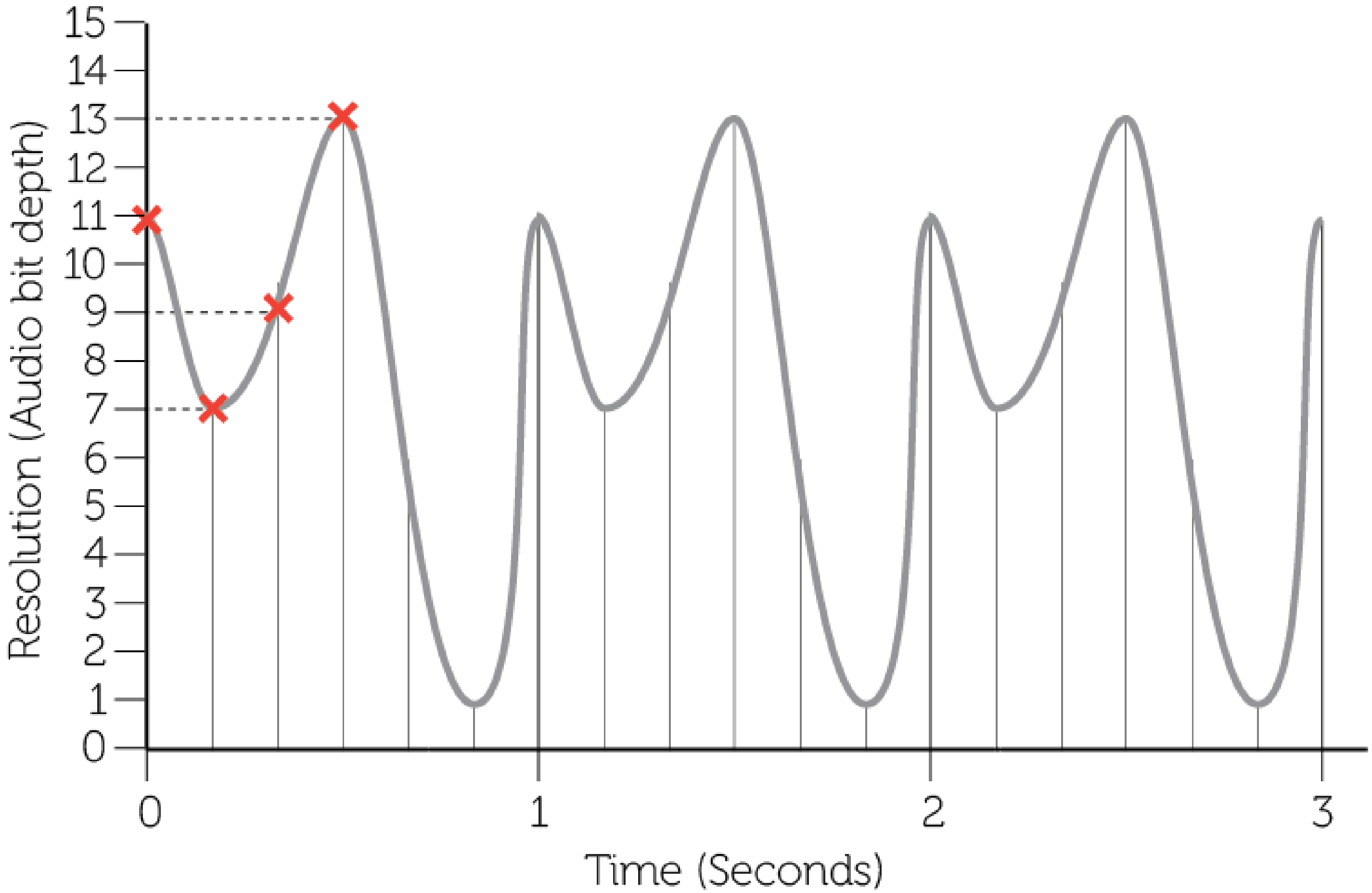
Sound sampling

We can see that that shape of digital sound is quite similar to analogue sound but the curves are not smooth.

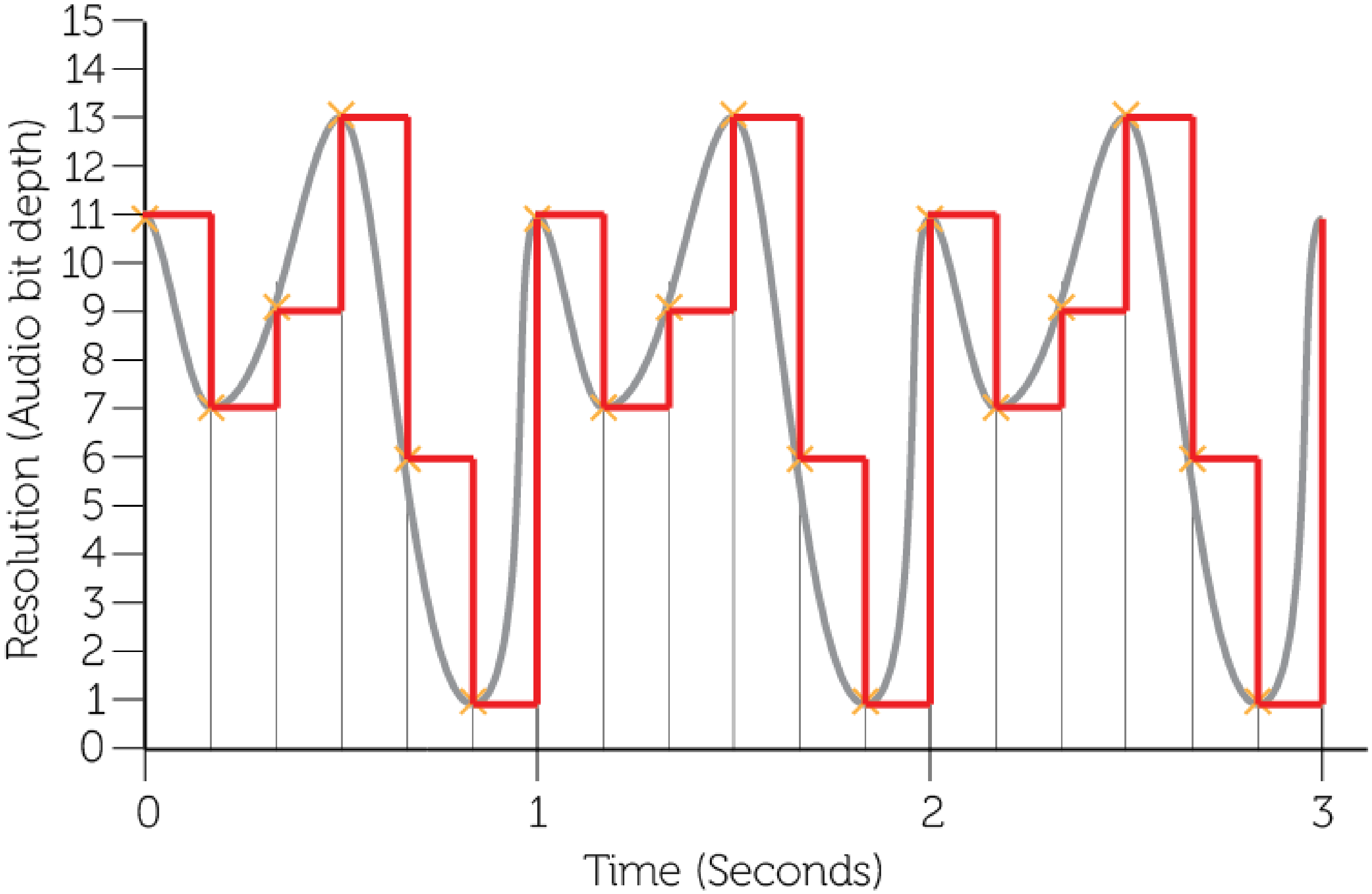
A listener will feel that the digital sound is same as the analogue sound but with reduced quality.



Sound sampling



Six samples per second



Digitised sound quality

- Recording quality improves:
  - the more frequently we sample the sound
  - the more accurately we record the wave height

- Increasing the sampling rate and sample resolution (bit depth) means recording more data points
  - more data points means more data is recorded and stored, therefore file size increases

Sampling rate	Number of samples per second
Resolution - Bit depth	Number of bits used to represent each sample

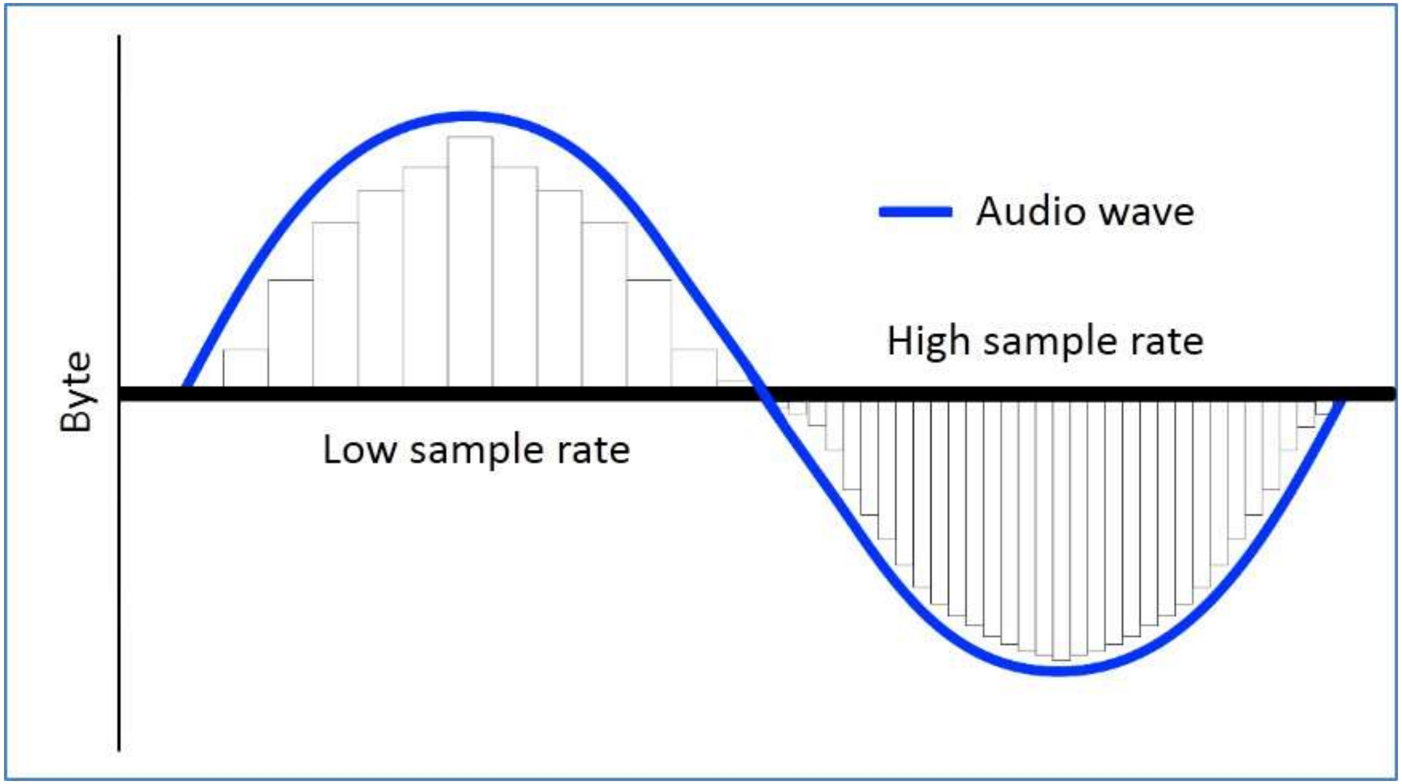
Factors affecting sound quality

Sampling rate

- The sampling rate is the **number of samples taken per second**.
- The higher the sampling rate, the higher the sound detail - the sound wave can be recorded more clearly.
- The unit for sampling rate is represented in Hertz (Hz).
- Each sample represents the amplitude of the wave at a certain point in time.

Factors affecting sound quality

Sampling rate



Factors affecting sound quality

Resolution (Bit depth)

- Sample resolution (bit depth) is the **number of bits available for each sample**.
- The higher the resolution (bit depth), the higher the quality the audio will be.
- *A CD has a bit depth of 16 bits and a DVD has a bit depth of 24 bits.*
- *A n bit system can have 2^n different values. Hence, a CD can represent values from 0 to 65535(2^{16} -1).*

Sound - summary

A sound wave is sampled for sound to be converted to binary, which is processed by a computer.

The sample rate is the number of samples taken in a second - measured in Hertz (Hz).

The sample resolution is the number of bits per sample (the bit depth).

The accuracy (quality) of the recording and the file size increases as the sample rate and resolution increase.

Sound

File size =

sample rate (Hz) x resolution (bits) x length (seconds) bits

(for a stereo sound file, multiply this answer by 2)

Sound

9 A 32-second sound clip will be recorded. The sound will be sampled 16000 times a second.

Each sample will be stored using 8 bits.

Calculate the file size in kilobytes. **You must show all of your working.**

File Size .... kB

[3]

Images

- Images are made up of **PIXELS**
- A pixel is the smallest identifiable area of an image
- Each pixel is a single colour and is given a binary value which represents that colour

e.g. 11111111 00000000 00000000 = Red

#	FF	00	00
	Red	Green	Blue

- A pixel's colour can be changed by changing this value

Image resolution

- **Resolution** is the number of pixels in an image
- The resolution of an image is expressed as its width and height in pixels e.g. 3264 x 2448
    - Screen resolution is expressed as **pixels per inch** (PPI)
    - Print resolution is expressed as **dots per inch** (DPI)

1x1 2x2 5x5 10x10 25x25 50x50 72x72 300x300

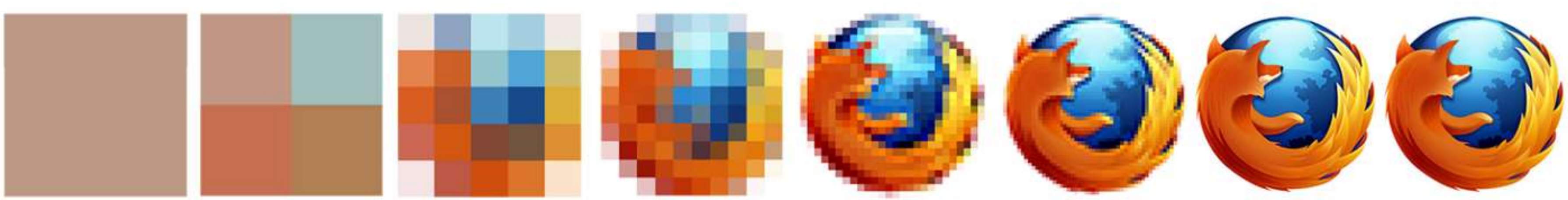


Image resolution

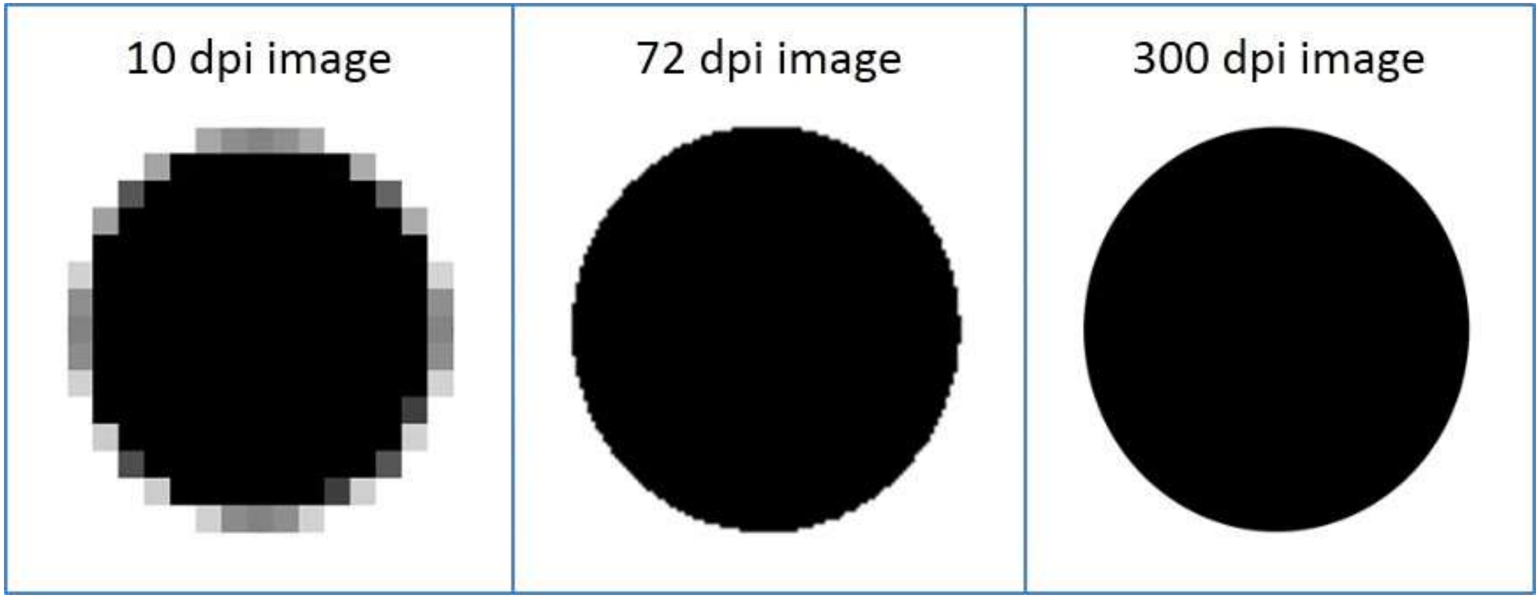


Image resolution

- Magazines and books have higher resolution compared to the images on a computer screen.
- An image on a website usually has a resolution of 72 dpi.
- An image in a book has a resolution of 300 or even 600 dpi.
- An image with a resolution of 300 dpi contains a grid of 300 pixels wide and 300 pixels high in a grid.

Colour depth

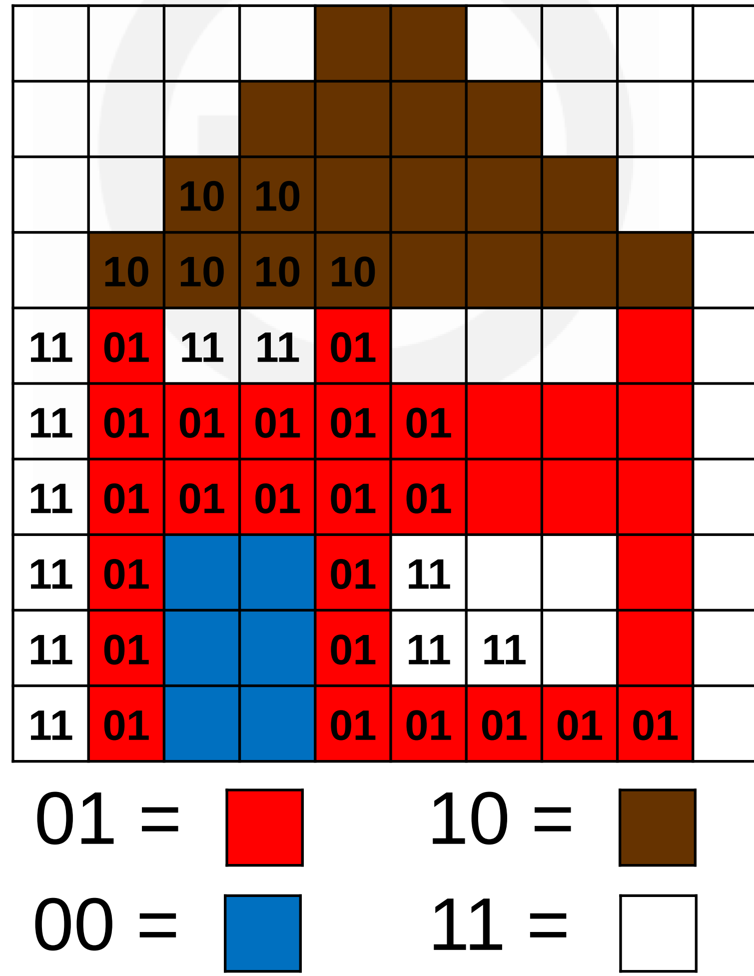
- Each pixel is given a binary value
- Each value represents a different colour
- Using one bit per pixel allows only 2 values, 0 and 1

  1 = Black, 0 = White

0	0	0	0	1	0	0	0	0	0
0	0	0	1	1	0	0	0	0	0
0	0	1	1	1	0	1	0	0	0
0	1	1	1	1	0	1	1	0	0
1	1	1	1	1	0	1	1	1	0
0	0	0	0	1	0	1	0	0	0
1	1	1	1	1	1	1	1	1	1
0	1	1	1	1	1	1	1	1	0
0	0	1	1	1	1	1	1	1	0
0	0	0	0	0	0	0	0	0	0

Colour depth

- More bits per pixel = more possible colour combinations
  - 1 bit = 2 Colours
  - **2 bits = 4 Colours**
  - 3 bits = 8 Colours
  - 4 bits = 16 Colours
- How many bits per pixel required for 256 colours?



Colour depth

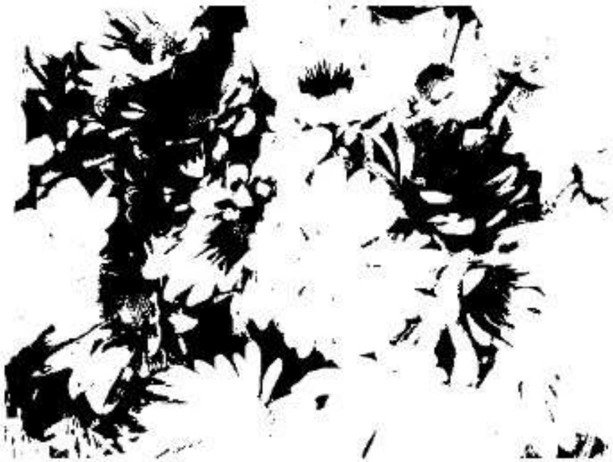
- Colour depth is the number of bits used to indicate the different colours of a pixel.
- In case of a black-and-white image, the colour depth is 1.
- A 2-bit colour depth can represent four different colours

Binary code	Colour
00	White
01	Light grey
10	Dark grey
11	Black

Colour depth

- As the number of bits increases, more colours can be used.
- An image with colour depth n can represent 2^n different colours.
- Most computer systems and digital systems use a 24-bit system that can represent over 16 million colours per pixel.
- With an increase in colour depth, the size of the file also increases.

Colour depth

2-bit colour depth	24-bit colour depth
	

Images - summary

A pixel is the smallest identifiable area of an image.

An image is a series of pixels that are converted to binary, to be processed by a computer.

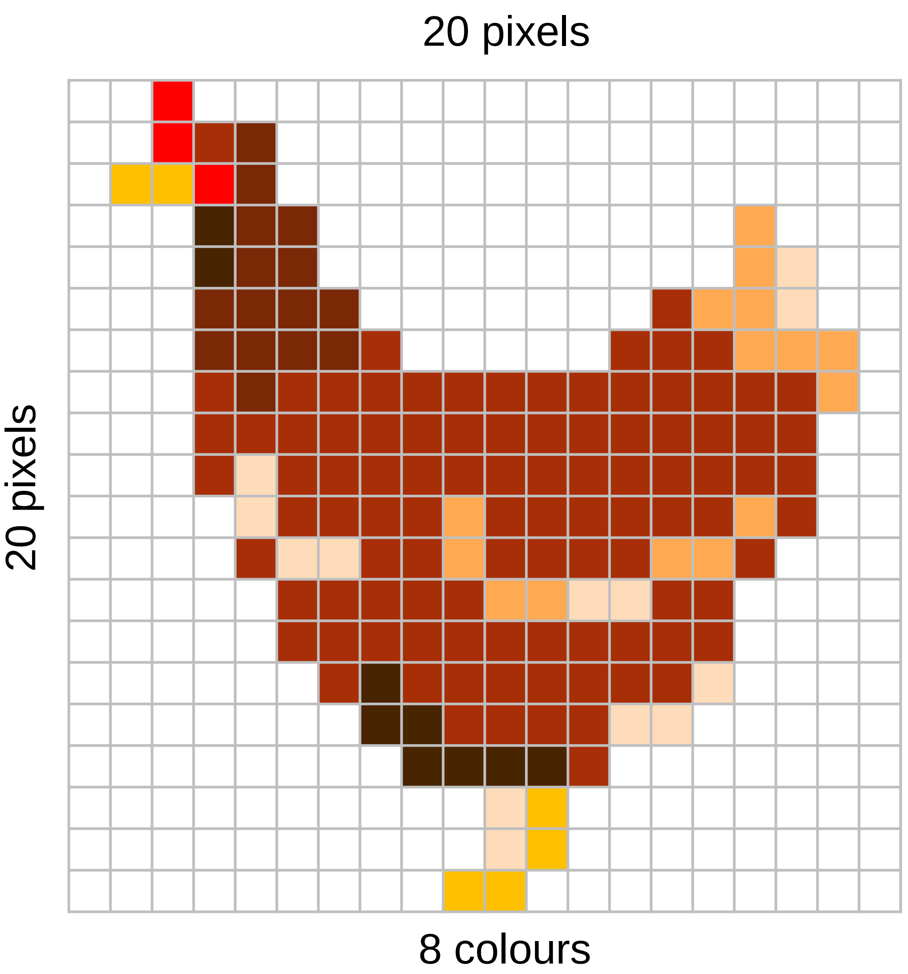
The resolution is the number of pixels in the image.

The colour depth is the number of bits used to represent each colour.

The file size and quality of the image **increases** as the resolution and colour depth **increase**.

Colours and resolution vs file size

- How does the number of colours affect file size?
- How does the size of the image affect file size?



Images

File size =

resolution (pixels) x colour depth (bits)   bits

Extras

PPQs:

- sound -> page 186
- images -> pages 203, 224, 227, 289
- general -> page 263

Digitized sound quality

- Recording quality improves:
  - the more frequently we sample the sound
  - the more accurately we record the wave height
- Increasing sampling rate and resolution means recording more data points
  - What happens to the size of the sound file?

Factor	Definition
Sampling rate	Number of samples per second
Bit depth	Number of bits used to represent each clip

Factors affecting sound quality

Bit depth

- High-quality audio files are created as pulse-code modulation (PCM).
- PCM is the process for digitising a sound file and creating an uncompressed file.
- **WAV** and **AIFF** are a few examples of uncompressed audio file formats.

Factors affecting sound quality

Bit rate

- Bit rate is the **number of bits of data used to store data sampled every second**. Measured in kilobits per second (kbps).

Bit rate = Sampling rate × bit depth × channels

- An audio file has 44,100 samples per second, bit depth of 16-bits and 2 channels (stereo audio). Bit rate of this file can be calculated as:

Bit rate = 44100×16×2 = 1,411,200 bits per second = 1411.2 kbps

- The higher the bit rate, the higher the quality of the recording

Factors affecting sound quality

Bit rate and file size

- A three-minute audio file with sampling rate of 44,100 samples per second, bit depth 16 bits and 2 channels, has a bit rate of 1411.2 kbps per second.
- For 3 minutes, the number of bits required is 1,411,200 × 180 = 254,016,000 bits.
- This value is equal to 254016000 ÷ 8 = 31752000 bytes = 31.75 megabytes (MB). This is the file size of three-minute audio file.

Sound file formats

- **.WAV** – uncompressed files
- **.FLAC** or **.M4A** lossless compression, slightly smaller files
- **.MP3** – Lossy compression, much smaller files

MP4 file format

- MP4 is a digital multimedia format most commonly used to store video and audio
  - It can also be used to store subtitles and still images
  - It allows different multimedia streams (video, audio, text) to be combined into one file

MIDI files

- MIDI stands for Musical Instrument Digital Interface
- A MIDI file:
  - is not a recording of a live music source
  - is a set of instructions for digital instruments to play synthesised sounds
  - can be used to synchronise an orchestra of digital instruments to play simultaneously
  - uses up to 1000 times less disk space than a conventional recording
  - is commonly used for mobile phone ringtones