5.4 Representing sound in a computer

In this topic you will cover:

- how sound is represented and recorded in a computer
- the difference between analogue and digital data and analogue and digital signals
- sampling, sampling resolution and sampling rate
- the principles of operation of an analogue-to-digital converter
- sound synthesis
- streaming audio.

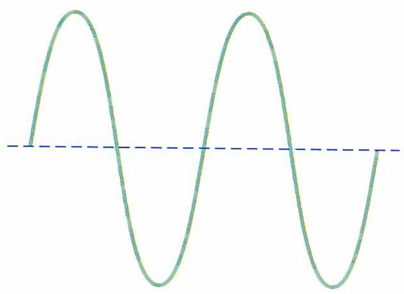

Fig. 5.4.1 *Pure tone vibration*

Fig. 5.4.2 *A superposition of pure tones*

Sound and data

Sound

Sound is an air pressure wave that is sensed by our ears. In an analogue sound system, the pressure wave is captured by a transducer, often a microphone, that produces an electrical voltage or current that varies in proportion to the sound pressure. The electrical signal can then be transmitted by telephone, broadcast by radio, preserved on magnetic tap (audio cassettes) or used in other ways. At the other end, the electrical signal is used to recreate sound by vibrating some mechanical surface in loudspeaker or an earphone, reproducing the original pressure wave with varying degrees of fidelity so that we can hear the sound. The higher the pitch of the sound, the more rapid the vibration. A pure tone is a regular sine wave (Fig. 5.4.1). A more complicated sound is a superposition of these waves and might look like Fig. 5.4.2.

A vinyl LP record exploits the shape of the sound wave by fixing a simila shape in the long spiral groove running across the surface of the record. The groove stores data about the sound in analogue form. When the record is played, a fine needle follows the changes in the groove and creates an electrical signal proportional to the changes. The signal is amplified then fed to a loudspeaker. In the days of wind-up gramophones the needle vibrated a mechanical surface directly to produce the sound.

Today almost all sound systems are digital. The electrical signal from a transducer such as a microphone is converted into a sequence of numerica values proportional to the strength of the signal and those numerical values are stored, transmitted, processed, etc., before they are converted back to an analogue electrical signal and then to sound. For example, an audio CD contains about an hour's worth of sampled voltage values: 44,100 samples per second in each of two stereo channels. Each sample is 16 bits, representing one of 65,536 possible voltage values; multiplying this out gives about 650 MB for an hour, or roughly 10 MB per minute.

What is analogue data?

Physical quantities such as temperature and pressure vary continuously in the real world. They are known as analogue quantities. Fig. 5.4.3 shows how the temperature in a classroom varies with time over a period of 12 hours.

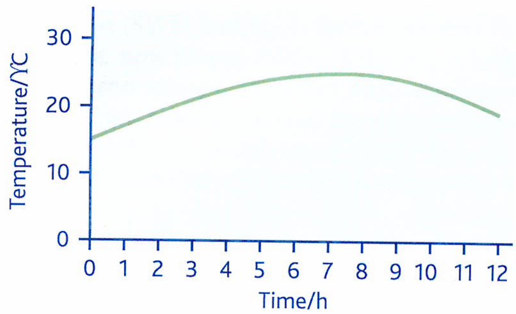

Fig. 5.4.3 *Analogue temperature data*

Data is anything that is collected and submitted for processing. **Analogue data** is data that varies in a continuous manner, such as speech conveyed from speaker to listener by sound waves.

What is digital data?

Digital quantities vary in discrete steps – they are discontinuous. Here are three sets of digital quantities: 1, 2, 3, 4 (set 1); 0, 1, 1, 0, 1, 0 (set 2); A, B, C, D (set 3). When analogue physical quantities are sampled and their values recorded, in the appropriate units, they become **digital data**. Digital data takes the form of discrete values. The denary number 57 is an example of digital data. Another example of digital data is the binary number 10101100. Fig. 5.4.4 shows digital data in the form of temperature readings taken at hourly intervals.

Hour	Temperature	Hour	Temperature
1	17	7	25
2	18	8	25
3	19	9	24
4	20	10	21
5	21	11	19
6	24	12	18

Fig. 5.4.4 *Digital temperature data*

Key terms

Analogue data: data that varies in a continuous manner.

Digital data: data that takes the form of discrete values.

Analogue signal: an electrical signal that varies in a continuous manner.

The difference between data and signals

To process analogue data, it must be sensed then converted into an equivalent electrical form. The electrical equivalent is called an analogue electrical signal or just an analogue signal. An **analogue signal** is an electrical signal that varies in a continuous manner. The device that does this is known as a transducer. A transducer is designed to convert energy from one form to another. A microphone is an example of a transducer. It converts continuously varying sound pressure waves into a continuously varying electrical signal. Another example of a transducer is a loudspeaker. A loudspeaker converts electrical energy into sound energy.

Fig. 5.4.5 shows a typical electrical circuit for converting sound energy into electrical energy. Fig. 5.4.6 shows the variation in pressure produced by the speaker whistling a pure tone. Fig. 5.4.7 shows the equivalent analogue signal.

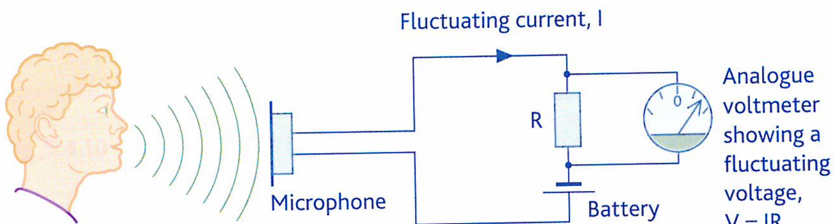

Fig. 5.4.5 *A typical electrical circuit for converting sound energy into electrical energy*

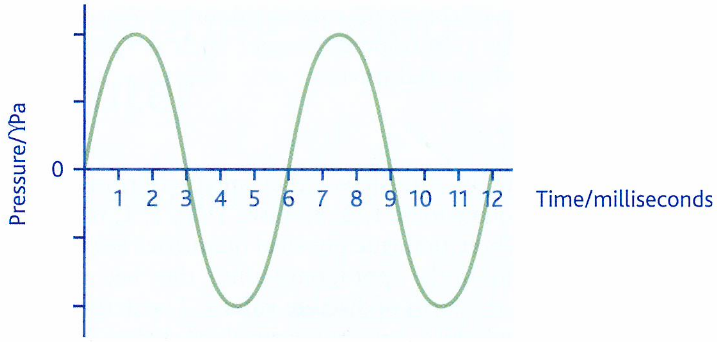

Fig. 5.4.6 *Variation in pressure produced by the speaker whistling a pure tone*

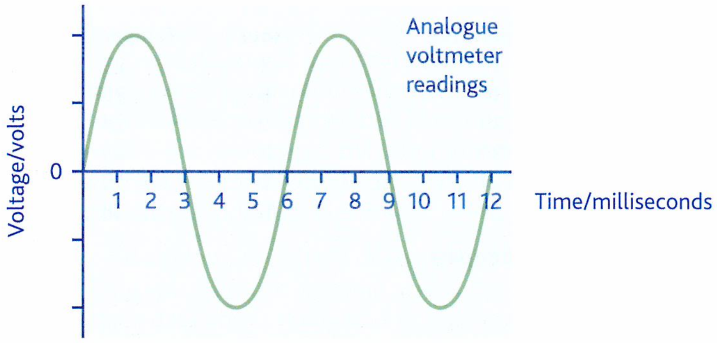

Fig. 5.4.7 *Equivalent analogue electrical signal*

Key terms

Digital signal: an electrical signal with voltage changes that are abrupt or in discrete steps.

Fig. 5.4.8 shows a **digital signal**. In this example the voltage levels available to the signal were –7.5, –5, –2.5, 0, 2.5, 5, 7.5.

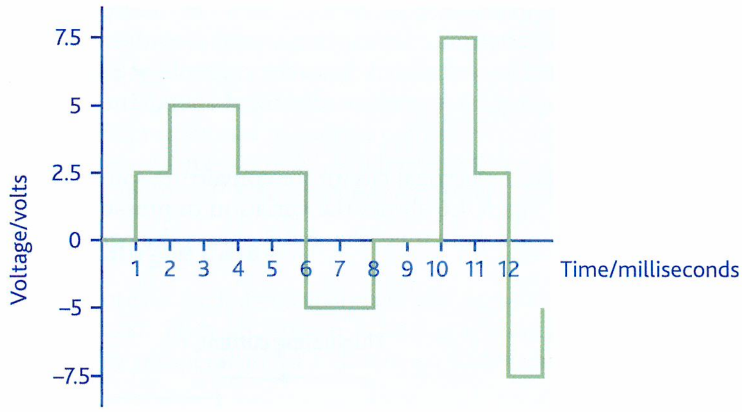

Fig. 5.4.8 *Digital signal*

Table 1 *Coding binary in voltages*

Voltage level	Binary
7.5	011
5	010
2.5	001
0	000
0	100
–2.5	101
–5	110
–7.5	111

Table 2 *Coding binary data by binary signals*

Voltage level	Binary
5	1
0	0

Each voltage level encodes a single item of binary data (Table 1). The most significant binary digit is a sign bit: 0 represents + and 1 represents –. Unfortunately, this leads to two binary patterns for zero.

It is possible to use just two voltage levels, 0 volts and 5 volts (Fig. 5.4.9). The digital signal is then a binary digital signal.

Each voltage represents a single item of binary data: 5 volts is binary 1 and 0 volts is binary 0 (Table 2).

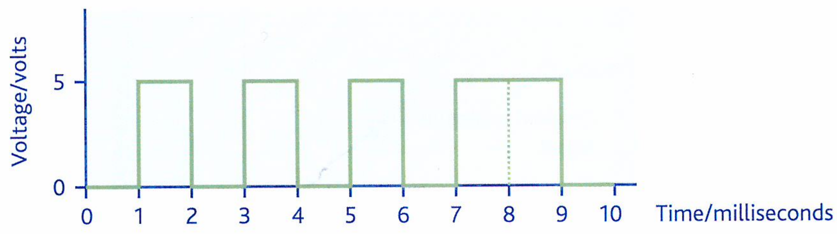

Fig. 5.4.9 *Two-level binary digital signal*

The stream of voltage pulses shown in Fig. 5.4.9 encodes the binary data 0110101010; the least significant digit is the first pulse to be produced.

Analogue to digital conversion

Sound waves may be converted into an equivalent analogue electrical current or voltage using a microphone. The variation in frequency (pitch) and the variation in amplitude (loudness) of the sound are reproduced in electrical form.

A computer may be used to record sound, but first the sound must be converted into digital form. This is done using an analogue-to-digital converter, or **ADC**, to sample an electrical equivalent of the sound and to convert the samples into a digital signal using two voltage levels. To play back the recorded sound, a digital-to-analogue converter, or **DAC** is used to convert the digital signal representing the recorded sound into an analogue electrical signal that is then played through a loudspeaker to produce sound. Fig. 5.4.10 shows how sound could be recorded and played back through the sound card of a computer system. The system relies heavily on a technique called pulse code modulation (**PCM**).

Key terms

ADC: analogue-to-digital converter; converts an analogue signal into an equivalent digital signal.

DAC: digital-to-analogue converter; converts a digital signal into an equivalent analogue signal.

PCM: pulse code modulation; a process for coding sampled analogue signals by recording the height of each sample in a binary electrical equivalent.

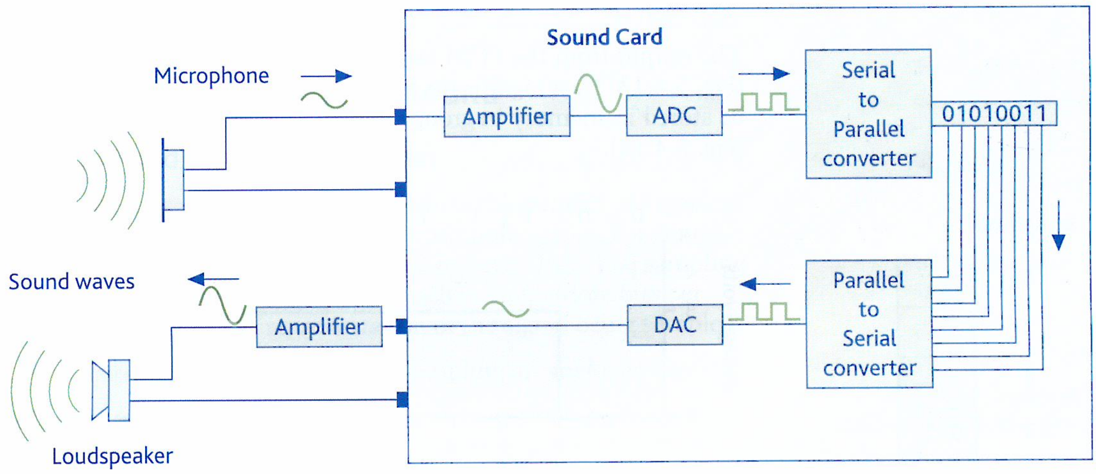

Fig. 5.4.10 *How sound could be recorded and played back through the sound card of a computer*

The PCM form of an analogue signal is produced by an ADC as follows:

1 Samples are taken of the analogue signal at fixed and regular intervals of time (Fig. 5.4.11). The sampling frequency, or sampling rate, must be at least twice the highest frequency in the analogue signal. These samples are represented as narrow pulses of height proportional to the value of the original signal. This process is known as pulse amplitude modulation (**PAM**).

Key terms

PAM: pulse amplitude modulation; a process that samples analogue signals at regular time intervals and produces electrical pulses of height proportional to the original signal's amplitude at the instant of sampling.

Fig. 5.4.11 *Sampling*

2 To produce PCM data, the PAM samples are quantised. That is, the height of each PAM pulse is approximated by an n-bit integer. For example, if $n = 3$, then $8 = 2^3$ levels are available for approximating PAM pulses (Fig. 5.4.12).

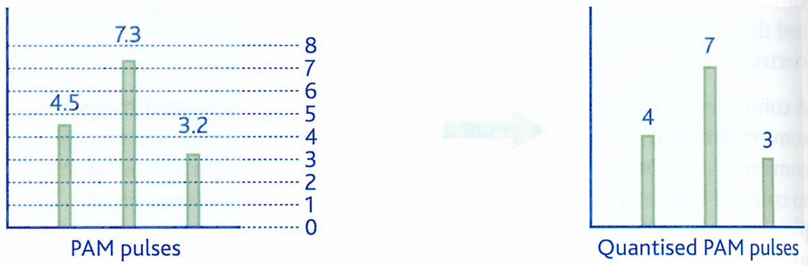

Fig. 5.4.12 *Quantisation*

3 Finally, the height of each PCM pulse is encoded in n bits to produce the digital output in binary signal form. For Fig. 5.4.12, using $n = 3$, the PCM pulses are coded in binary as follows: $4 \rightarrow 100_2$, $7 \rightarrow 111_2$, $3 \rightarrow 011_2$.

The output from the PCM encoder is a sequence of fixed-height pulses (Fig. 5.4.13) least significant bit first. The train of pulses may then be stored in memory in groups, where each group consists of 3 bits (Fig. 5.4.14).

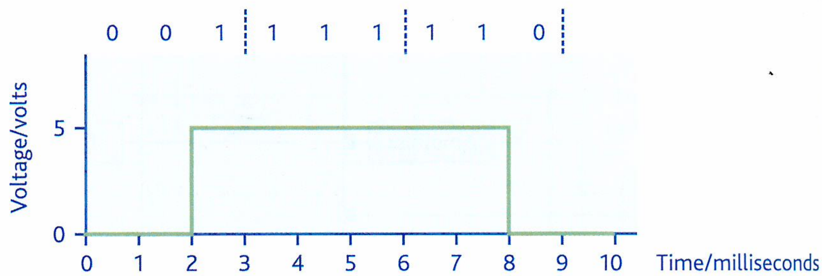

Fig. 5.4.13 *Output from PCM encoder*

Sample no	
6	...
5	...
4	...
3	011
2	111
1	100

Fig. 5.4.14 *Encoding*

Key terms

Quantisation noise: the difference between the original amplitude and its sampled value.

The process is reversed using a DAC. The DAC produces an analogue signal that is an approximation of the original analogue signal (Fig. 15). The staircase effect is a result of approximation during PCM quantisation. The deviation from the original is known as **quantisation noise**. Fig. 5.4.15 shows a higher sampling rate than was shown in the previous diagrams.

This electrical signal, sometimes after amplification, is used to recreate the sound by vibrating a mechanical surface in a loudspeaker or an earphone.

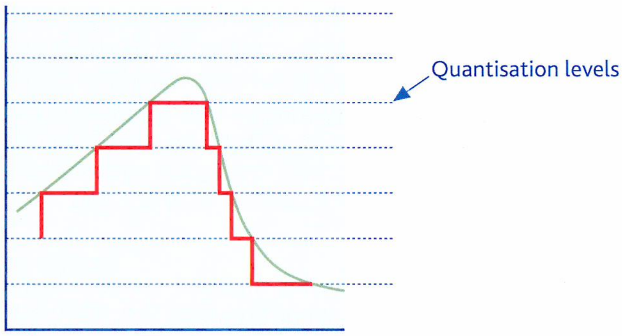

Fig. 5.4.15 *DAC output: an analogue signal that approximates the original analogue signal*

Worked example

Problem

An analogue signal of frequency 1,000 Hz is converted to a PCM digital signal by sampling at a frequency of 2,000 Hz = 2,000 samples per second. Each sample is encoded in 8 bits using PCM coding. How many bytes of storage are required for the PCM-coded result if recording 10 seconds of analogue signal?

Solution

A sampling frequency of 2,000 Hz means that 2,000 samples are taken per second. In 10 seconds this means that 20,000 samples are taken. One byte (8 bits) is required for each sample. Therefore 20,000 bytes of storage are required.

Questions

1. A sound wave is digitised using a circuit containing an ADC that uses PCM coding. The sampling frequency is 40 kHz. (a) What device is required in addition to the ADC? (b) Each sample is coded in 16 bits. How much storage is required in bytes for one minute of recorded sound?
2. Explain by example the difference between analogue and digital data.
3. Explain the difference between an analogue signal and analogue data.
4. What converter will convert a digital signal into an analogue signal?

Managing and manipulating sound

Sampled sound and Nyquist's theorem

The sampling rate, or sampling frequency, defines the number of samples that are taken per second when digitising a continuous signal. For time-varying signals, the sampling rate is measured in hertz (Hz). The sampling resolution is the number of bits assigned to each sample. More bits per sample means a more accurate representation of the signal being sampled.

Here are some questions to consider when sampling an analogue signal.

- How often do we sample? Sampling more often means that we can track the changes in a rapidly varying signal more accurately. If we don't sample often enough, we miss meaningful changes in the signal. **Nyquist's theorem** states that we must sample at a frequency at least twice the rate of the highest frequency in the sampled signal.
- How accurately do we sample? More bits per sample means a more accurate representation of the signal wave.
- How much space do we need to store the sound? More frequent sampling and more accurate samples mean more data, which can add up very quickly.
- How much bandwidth do we need to transmit sound in electrical form? Bandwidth is information-carrying capacity. It takes more capacity to preserve all the low and high frequencies that might be

Key terms

Nyquist's theorem: sample at a frequency at least twice the rate of the highest frequency in the sampled signal.

needed, perhaps for hi-fi sound, whereas it takes much less bandwidth if only some frequencies need to be transmitted, perhaps for telephone speech.

- How do we convert back from numbers to an analogue waveform? If the original waveform is to be reproduced with reasonable fidelity, we have to do this carefully too.

Storing sound in files

There are several file formats for storing digitised sound. The most notable is the WAV format supported in Windows 3.1, 95, 98, NT, 2000, ME, XP and Vista. One minute of sound saved to a WAV file requires 2.5 MB of disk space. The WAV format is used when audio is stored on CDs.

Another file format is MPEG audio, which uses extensions .mp2, .mpa, .mp3, .mp4. MPEG is primarily a compression algorithm that can be applied to a number of audio formats, e.g. WAV. However, it also exists as an audio file format. MPEG is based on psychoacoustic modelling that removes frequencies the brain and ear will not miss. This substantially reduces the file size before other compression techniques are applied. Files of one-tenth their WAV size with no apparent loss in quality are achievable with MPEG CD quality. For MP3, one minute of CD-quality audio can be cut down from 2.5 MB to about 0.25 MB (**www.MP3-tech.org**). MP3 is the compressed format that is used for most music available on the Internet.

We can also use the properties of human speech to compress telephone speech significantly (**www.data-compression.com/speech.html**). This is used in mobile phones. Compression techniques for speech do not work as well for music.

When sound is converted into digital form it is possible to mix sounds from multiple sources and to store the combination in a single file. There is a lot on the Web about sound. Visit **www.music.arts.uci.edu/dobrian/digitalaudio.htm**.

Synthesising sound

According to **www.webopedia.com/TERM/MIDI.html**, Musical Information Digital Interface, or **MIDI**, is a widely used representation for instrumental music that does not store sound waves at all. Instead it stores a digital representation of the notes to be played, including what note, what instrument and what duration. The resulting form is very compact and very flexible for many purposes. It is easy to transpose into a different key, play on different instruments and synthesise musical notation from it.

The device that is going to produce the sounds has to create or **synthesise sounds** from an internal definition of what they sound like. For example, a typical low-end synthesiser can approximate 64 instrumental sounds. Depending on the synthesiser, the approximation to the real instrument may go from very good to very bad. Piano sounds can be approximated fairly well but the human voice is not well approximated.

Streaming audio

Many Internet audio sources stream the sound to a streaming client, such as RealPlayer, on your machine. As the streaming client receives the audio data, it puts it in a buffer to be stored until it is used. A few seconds after the client starts receiving audio data, the client player starts

Did you know?

Stagecoach wheels appear to rotate backwards in films because the frame rate of the film camera recording the scene is less than twice the rotation rate of the wheel spokes.

Key point

More bits per sample means a more accurate representation of the signal wave. More frequent samples means a more accurate representation of the signal wave. More bits per sample means more data. More frequent samples means more data. More bits per sample means more bandwidth required for transmission. Higher sampling frequencies mean more bandwidth required for transmission.

Key point

For a given duration of audio in seconds, MIDI files require less storage space than MP3 files and MP3 files require less storage space than WAV files. MP3 is used for most music available on the Internet. WAV is used for audio CDs.

Key terms

MIDI: Musical Information Digital Interface; a way of representing the sounds made by an instrument.

Synthesise sound: use digital means to generate audio signals resembling instrument sounds or the human voice.

reading the data from the buffer and playing it. As long as the player is not trying to access data that hasn't been received yet, streaming will be successful. If the buffer runs out of audio data, the player will pause until it receives more. Streaming avoids the need to download and store large sound files. The sounds come in as they are needed but without occupying file space on your hard drive. The other reason that many sources stream audio is that, in theory, it prevents copying.

Editing sound

Once a sound wave is stored in digital form, it may be edited to remove notes, add new notes, change the frequency of notes, and reduce or eliminate background noise. A sound-editing package such as Goldwave or Audacity can take separate digital recordings of sound waves then merge or mix them into one file in different ways.

In this topic you have covered:

- the difference between analogue and digital data
- the difference between analogue and digital signals
- sampling and digitising using PAM and PCM
- sampling rate and sampling resolution
- analogue-to-digital converter and quantisation noise
- sampled sound and Nyquist's theorem
- factors that affect the quality of recorded sound
- file formats for storing sound data in files
- sound synthesis, streaming audio, editing sampled sound.

Questions

5 State Nyquist's theorem.

6 State two factors likely to affect the quality of the recorded sound and the file size of a digital audio recording.

7 Take each factor in your answer to Question 2 and explain how it affects the quality and the file size.

8 What is meant by synthesising sound?

9 Explain the process of streaming audio.