Chapter 2 Signals and sensors

2.1 Analogue and digital signals

Learning objectives:

→ State the advantages of digital transmission over analogue transmission.

→ Explain what is meant by sampling of an analogue signal.

→ Distinguish a reconstructed analogue signal from the original signal.

Bits and bytes

A digital signal is a signal that has only two states, which it switches between. Digital radio and TV signals are pulses of electromagnetic waves that switch between two states. The two states may be two different amplitudes or two different frequencies. Electronic signals are pulses defined by two voltage values that are usually called high or low (or numerically called 1 or 0).

An analogue signal is a signal that can vary continuously in amplitude. A microphone produces an alternating voltage when sound waves are incident on it. The alternating voltage is an analogue signal which has an amplitude that depends on the amplitude of the sound waves. The lower limit of the signal is zero, and the upper limit depends on the design of the microphone.

Analogue signals produced by a mobile phone microphone are converted to digital signals which are then transmitted and converted back to the original analogue signal at the receiver. The advantages and disadvantages of this process are described on page 19.

Data storage and transmission

The ‘0’s and ‘1’s of a digital signal are called **bits**. A series of bits is called a **byte**. A code must be used when digital pulses are used to send information. For example, ASCII (American Standard Code for Information Interchange) is used to represent every letter, number, and symbol on a computer keyboard by a unique 8-bit byte. When you press a particular key on a computer keyboard, a unique stream of electronic pulses is sent from the keyboard to the microprocessor in the computer.

Notes about binary numbers

The two digits of the binary system of counting are 0 and 1. Any binary number can be expressed as a series of these two digits. Each digit or bit in the series represents the presence or absence of a different power of two in ascending order from right to left. For example, the binary number 101 represents $1 \times 2^2 + 0 \times 2^1 + 1 \times 2^0$, which is 5 in the denary (or base 10) number system.

Table 1 shows the value of the sixteen 4-bit bytes.

Table 1

base 10 (denary)	base 2 (binary)	base 10 (denary)	base 2 (binary)
0	**0000**	8	**1000**
1	**0001**	9	**1001**
2	**0010**	10	**1010**
3	**0011**	11	**1011**
4	**0100**	12	**1100**
5	**0101**	13	**1101**
6	**0110**	14	**1110**
7	**0111**	15	**1111**

Advantages and disadvantages of digital sampling and transmission

One of the key advantages of digital transmission over analogue transmission is the elimination of noise which is unwanted random variations in the signal amplitude due to various effects. (See page 71.) Without noise, the received signals are of the same quality as transmitted signals. For example, digital radio is of higher sound-quality than AM or FM radio programmes because noise is eliminated from a digital signal. In addition, because digital signals can be compressed by using pulses of shorter duration, more information can be sent using digital transmission. Also, many signals can be carried in a single frequency channel by multiplexing the signals, which means sending a pulse from each signal in turn.

Analogue signals can be converted into digital signals for transmission using an analogue-to-digital (A/D) converter and then converted back upon reception using a digital-to-analogue (D/A) converter. In this way, audio and video signals can be transmitted digitally, giving programmes of higher technical quality as well as more channels.

Eliminating noise

In the process of transmission, electrical signals along wires and cables become weaker because of resistance, and signals carried by electromagnetic waves become weaker because of absorption and spreading. Because the noise on a signal increases and the signal becomes weaker as it travels, the **signal-to-noise** ratio decreases the further the signal travels. If the ratio decreases too much, the signal is masked by the noise and cannot be detected. Noise cannot be eliminated from analogue systems, because amplifiers used to boost the amplitude of a signal would also boost the noise amplitude.

Noise can be eliminated in digital systems by using **regenerator amplifiers**, which are designed to produce a clean output pulse from each noisy input pulse, as long as the noise does not mask the signal. With regenerator amplifiers at regular intervals along a transmission link, a digital signal transmitted without noise would be received at the other end of the link without noise. Therefore, because of the elimination of noise, ensuring that the received signals are of the same quality as the transmitted signals, digital transmission is particularly advantageous over analogue transmission for long-distance links.

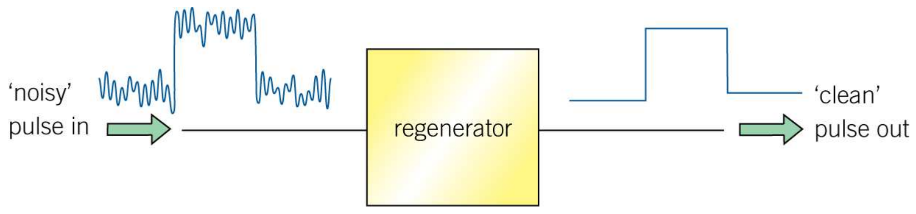


Figure 1 *Cleaning a noisy pulse*

Extra information

Because digital signals can be compressed by using pulses of shorter duration, more bits per second and therefore more information can be sent by using digital transmission. In addition to sending extra data bytes, the extra bits can be used:

- to code the data bytes so that each byte goes to a particular destination. For example, a 16-bit byte may consist of an 8-bit data byte and an 8-bit address byte.
- to check for and correct errors in transmission. For example, an extra bit may be used to check that the parity of a data byte has not changed during transmission. The parity is set at either 0 or 1 according to whether or not there are an odd or an even number of '1's in the data byte when it is transmitted. If one of the data bits changes its state in transmission, the parity check will detect the change.

For the above reasons, digital transmission is more reliable and more cost-effective than analogue transmission.

The disadvantages of digital transmission include the following:

- Extra circuits are needed to sample and convert analogue signals to digital signals and to recreate the original analogue signal.
- Sampling errors may occur due to unwanted voltage variations in the sampling circuit.
- Digital signals need to be synchronised so that the receiver circuit recognises the start of each byte.

Conversion of an analogue signal into a digital signal

Figure 2 outlines how an analogue signal is converted into a digital signal at a single output in three stages. The whole process is called **pulse code modulation**.

Links

For more about pulse code modulation, see Topic 5.2, Amplitude modulation and frequency modulation.

For more about the operational amplifier, see Topic 3.2, Ideal operational amplifiers, and Topic 3.3, Operational amplifier circuits.

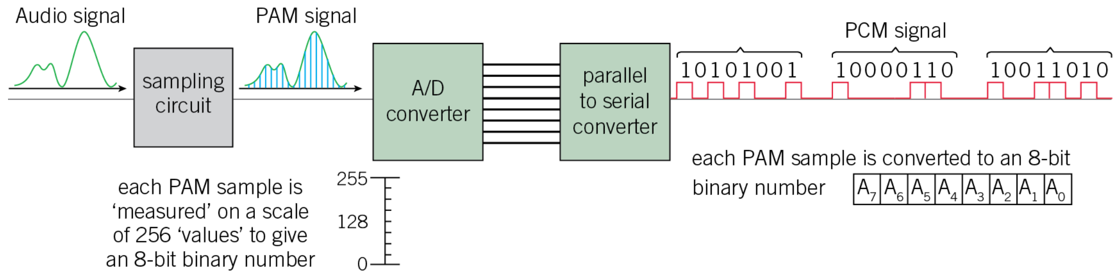


Figure 2 *Pulse code modulation*

Stage 1: Sampling

The analogue signal is sampled at fixed intervals to produce a pulse amplitude modulated (PAM) signal. This can be achieved by supplying the analogue signal to an operational amplifier that is switched on for a brief interval. The output voltage of the amplifier is equal to the sample voltage. You will meet operational amplifiers in Topics 3.2 and 3.3.

Stage 2: Conversion to a data byte

The sample is then converted to a data byte in the form of a binary number on a scale whereby when all the bits are equal to 1, this represents the maximum amplitude that can be sampled. Each sample voltage of the pulse amplitude modulated signal is converted to a binary number with a particular number of bits (e.g., an 8-bit byte or 8-bit word) on a scale whereby the maximum amplitude is represented by 1 for every bit (e.g., 1111 1111 for an 8-bit word). This process is called quantisation and is achieved by using an A/D converter that has a single input and a suitable number of parallel outputs (e.g., 8 for conversion of an 8-bit byte).

Stage 3: Conversion to a digital signal at a single output

Each byte on the parallel outputs of the A/D circuit is then converted, using a parallel-to-serial converter, into a sequence of bits at a single output terminal. The bytes representing successive samples are then transmitted as a stream of pulses, as outlined in Figure 2.

Table 2 shows all the 4-bit binary numbers that result for an analogue signal that can vary between 0 and 15 V. Each row shows the minimum voltage needed to produce the corresponding digital signal. For example, an analogue voltage of 6.8 V would produce 0110 as the digital signal.

Table 2

Analogue voltage / V	Digital signal	Analogue voltage / V	Digital signal
0	**0000**	8	**1000**
1	**0001**	9	**1001**
2	**0010**	10	**1010**
3	**0011**	11	**1011**
4	**0100**	12	**1100**
5	**0101**	13	**1101**
6	**0110**	14	**1110**
7	**0111**	15	**1111**

Notes

1 In the example above, each bit has 2 states (0 or 1) and there are four bits in each digital sample. The analogue signal can therefore be converted into 2^4 (= 16) digital samples as shown in Table 2.

2 The least significant bit of a byte (i.e., the bit at the right-hand end) in Table 2 changes when the analogue voltage changes by 1 V or more.

Quality of conversion

When pulse code modulation is used to transmit a signal, the quality of the reconstructed analogue signal depends on the sampling frequency and the number of bits per sample.

Sampling frequency

If the sampling frequency is too low, the analogue signal cannot be recovered from the digital signal. To recreate the analogue signal, it can be shown theoretically that the frequency of the sampling process (i.e., the sampling rate) must be at least twice the highest frequency of the analogue signal.

For example, an audio signal on the telephone network is restricted to the frequency range 0–4 kHz. So the sample frequency of such an audio signal is 8 kHz. In other words, such a signal should be sampled at a rate of at least one sample every 125 μs. The reason for this minimum sampling rate is that the analogue signal cannot be regenerated accurately from the samples if the sample rate is less than twice the highest frequency of the signal.

Number of bits in a sample

The number of bits in a sample (i.e., the word length) determines the resolution and therefore the quality of the recovered analogue signal. In addition to the data byte representing the signal sample, further bits are added to indicate the arrival of a byte at the receiver and to check that the byte has not been corrupted during transmission. In a sampling process that produces a sample consisting of an *n*-bit byte, the maximum allowed sample voltage is resolved into 2^n levels. The regenerated analogue signal therefore consists of successive samples at 2^n equally spaced levels. For example, a digital signal consisting of 3-bit bytes would resolve the sample into eight levels (= 2^3) and would therefore produce a regenerated analogue signal that has eight equally spaced levels. The more bits there are in each sample, the smaller is the spacing between adjacent levels and the greater is the quality of the recovered analogue signal.

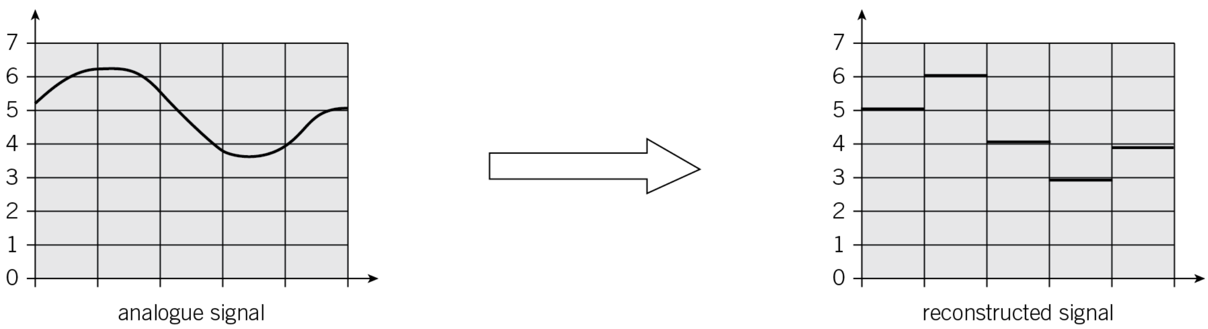


Figure 3 *Sampling levels*

number of bits transmitted per second = sample frequency × number of bits per sample

So the bit rate for a sample frequency of 8 kHz would be 128 kilobits per second for 16-bit words, equal to 1 bit transmitted every 7.8 μs. The maximum bit rate along a wire or a coaxial cable or any other transmission path depends on the type of transmission path. The sample rate depends on the type of signal being transmitted (audio, TV, etc.). So the number of bits per sample, and hence the quality of the recovered analogue signal for a particular type of signal, depends on the choice of the transmission path.

Note

The bit rate can be increased considerably by multiplexing the bits of several signals in a process known as **time-division multiplexing**. This is explained in more detail in Topic 5.1. In effect, the bits are shortened considerably, so the corresponding bits from each signal can be transmitted in sequence before the next bits from each signal.

Conversion of a digital signal to an analogue signal

To convert a digital signal to an analogue signal, the digital signal must be decoded at the receiver. A digital signal in the form of *n*-bit bytes (where *n* is the number of bits per byte) in sequence can be converted to an analogue signal by converting each *n*-bit byte into an analogue voltage on a scale where the maximum byte value (e.g., 1111 1111 for an 8-bit byte) produces the maximum analogue voltage that is allowed.

Assuming the signal is a stream of bytes in a single channel:

- each byte must first be converted from serial to parallel form,
- the *n* bits of each byte are then supplied as '0' or '1' to an *n*-bit D/A converter, which is a summing amplifier with *n* input terminals. See Topic 3.3 Figure 3. Each input resistance is chosen so that the output voltage from the operational amplifier is in proportion to the voltage level the byte represents.

In this way, each *n*-bit byte is converted to an analogue voltage. The analogue signal can then be recreated from successive analogue voltages. To understand how a D/A converter works, consider the 4-bit D/A converter in Figure 4.

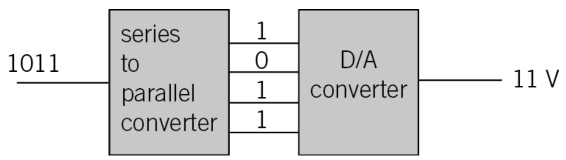
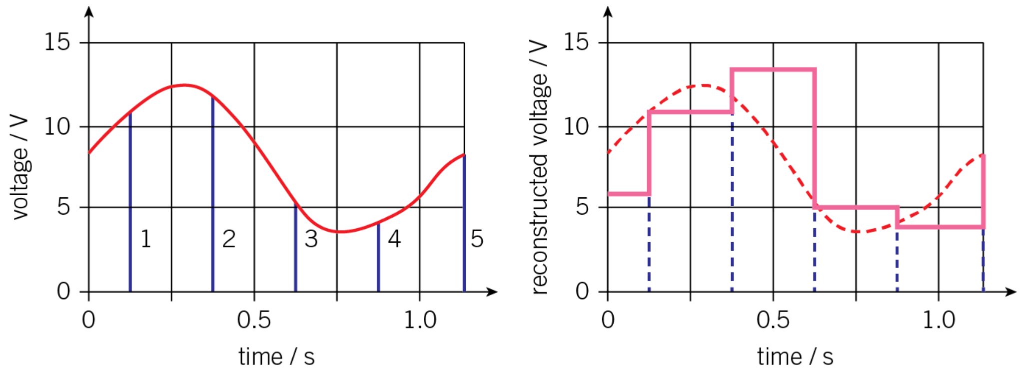


Figure 4 *Digital-to-analogue conversion*

The output voltage depends on the state of the input voltages and can have any one of 16 values on a scale from 0 to 15 V in 1 V steps. In Figure 4, a 4-bit byte digital signal produced by sampling an analogue signal is converted back into the analogue signal. The recreated analogue signal is not smooth like the original signal. This is because the digital signal has only four bits and so can create only 2^4 samples. In other words, the voltage of the recreated analogue signal can have only 16 values. Using an A/D converter that quantises the analogue signal into more bits per byte would result in a recreated signal with smaller steps. For example, an A/D converter that produces 8-bit digital signals would give 2^8 or 256 values of the digital signal and therefore a much smoother recreated signal.

QUESTION: An 8-bit D/A converter gives an analogue output in the range 0 to 5.1 V. Calculate the least change of the output voltage.



original sample	1	2	3	4	5
voltage / V	11.0	13.5	5.2	4.0	7.8
digital value	1011	1101	0101	0100	0111
reconstructed voltage	11.0	13.0	5.0	4.0	7.0

Figure 5 *Reconstructing an analogue signal*

Summary questions

1 State the difference between an analogue signal and a digital signal and give one example of each.

2 **a** Explain what is meant by noise in a communication system and explain why it can be problematic.

b State the function of a regenerator amplifier in a digital system.

3 An audio signal with a maximum amplitude of 6.4 V is converted into a digital signal. Each sample of the audio signal is converted into an 8-bit byte.

a State the byte that is produced when the voltage of the audio signal is:

i 6.4 V

ii 3.0 V.

b State the least change of the audio voltage that would produce a different byte.

4 An analogue signal is converted to a digital signal which is then used to reconstruct the analogue signal.

a Explain why a digital signal with 8-bits per byte would give a better reconstructed signal than signals with 4 bits per byte.

b When the analogue signal is sampled and converted to a digital signal, the sampling frequency must be at least twice the maximum signal frequency. State and explain the result of using a lesser sampling frequency.

2.2 Sensors

Learning objectives:

→ Describe a sensor circuit that includes a thermistor or a light-dependent resistor.
→ Explain how a temperature or a light sensor works.
→ Calculate the output pd of a temperature or light sensor.

Sensor-operated systems

When you approach an automated door, an **electronic sensor** detects your presence and causes the door to open. Sensors are used in any electronic system that responds to an external change. A list of such electronic systems would be very long and very diverse and would include everyday items such as smoke alarms and touch screens and much more sophisticated devices such as remote-controlled weather stations and robotic machines in vehicle assembly plants.

A sensor supplies a signal to a **processing unit** which can operate an **output device** such as an indicator lamp or a relay. The processing unit may receive signals from more than one sensor, and it is programmed to respond according to the signals it receives.

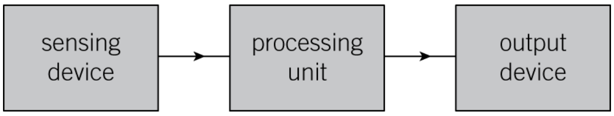


Figure 1 *A sensor-operated electronic system*

An electronic sensor contains a **sensing device** which responds to an external change by producing:

- an analogue signal such as the change of pd from a **potential divider** that includes a **thermistor** or a **light-dependent resistor**, or
- a digital signal such as the switch in the pd from a **Hall effect proximity sensor**, as explained in Topic 1.4.

Sensing devices in potential dividers

A potential divider consists of two resistors in series of resistances R_1 and R_2 connected to a source of fixed potential difference V_S as shown in Figure 2.

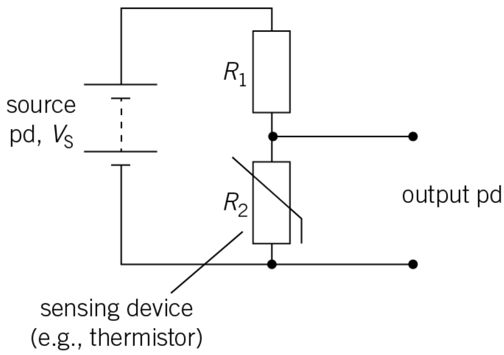


Figure 2 *Using a potential divider*

Link

Potential dividers were looked at in Topic 13.5, The potential divider, in Year 1 of the *AQA Physics* student book.

The current I through the resistors is equal to the source pd divided by the total resistance of the two resistors: $I = \frac{V_S}{R_1 + R_2}$

The pd across each resistor is equal to its resistance × the current. Therefore:

$$\text{the pd across } R_1 \text{ is equal to } IR_1 = \frac{R_1}{R_1 + R_2} \times V_S$$

$$\text{the pd across } R_2 \text{ is equal to } IR_2 = \frac{R_2}{R_1 + R_2} \times V_S$$

With a sensing device such as a thermistor or a light-dependent resistor as one of the resistors and the other resistor having a fixed resistance, when the resistance of the sensing device changes, the pd across one resistor decreases and the pd across the other one increases. If R_1 and R_2 are the resistances of the fixed resistor and the sensing device respectively, as shown in Figure 2, when the resistance of the sensing device increases, the current decreases, so the pd across the fixed resistor decreases. Therefore, the pd across the sensing device increases.

Note

In Figure 2, the pd across the sensing device provides the output pd from the potential divider. If the resistors were interchanged in the circuit with the output pd provided by the pd across the fixed resistor, the output pd would increase when the resistance of the sensing device decreases.

Temperature sensor

A temperature sensor uses a thermistor as the sensing device, as shown in Figure 2.

A thermistor is a resistor whose resistance changes significantly with change of temperature. Thermistors that have a resistance that falls with increasing temperature are called ntc thermistors because they have a negative temperature coefficient. Figure 3 shows how the resistance of an ntc thermistor decreases as the temperature increases.

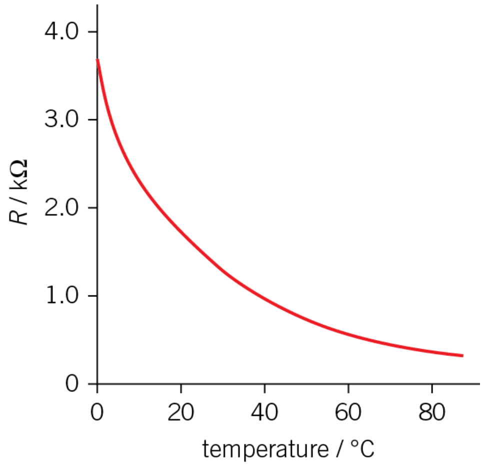


Figure 3 *Resistance variation with temperature for an ntc thermistor*

Link

Thermistors were looked at in Topic 12.4, Components and their characteristics, in Year 1 of the *AQA Physics* student book.

Note that:

- the variation of resistance with temperature is non-linear
- the rate of change of resistance with temperature decreases as the temperature increases.

Assuming that the thermistor in Figure 2 is an ntc thermistor, when its temperature increases, its resistance decreases. As a result, the current in the circuit increases, so the pd across the fixed resistance (= current × resistance of the fixed resistor) must therefore increase. The pd across the thermistor therefore decreases because it is equal to the source pd minus the pd across the fixed resistor. Because the output pd in Figure 2 is across the thermistor, the output pd therefore decreases.

Note that if the output pd had been across the fixed resistor, the output pd would increase.

Worked example

The thermistor in Figure 3 is connected in series with a 4.0 kΩ resistor *R* and a 5.0 V battery. Calculate the pd across *R* when the thermistor temperature is 40°C.

Solution

At 40°C, the thermistor resistance is 1.0 kΩ.

The circuit resistance is therefore 5.0 kΩ.

The pd across $R = \frac{\text{resistance of } R}{\text{circuit resistance}} \times \text{battery pd} = \frac{1.0}{5.0} \times 5.0\text{ V} = 1.0\text{ V}$.

Light sensor

A light sensor uses a light-dependent resistor (LDR) as the sensing device. The resistance of an LDR decreases non-linearly when the intensity of light incident on it increases. Typically, the resistance of an LDR decreases from about 1 MΩ in darkness to about 100 Ω in sunlight. Figure 4 shows an LDR in a potential divider.

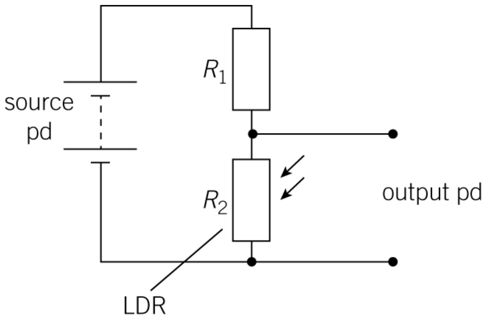


Figure 4 *Using a light-dependent resistor in a potential divider*

- When the light intensity is increased, the resistance of the LDR decreases, so the current in the circuit increases. As a result, the pd across the fixed resistor increases. This means that the pd across the LDR decreases and so the output pd decreases.
- When the light intensity is decreased, the current in the circuit decreases, so the pd across the fixed resistor decreases. Hence the pd across the LDR increases, so the output pd increases.

Summary questions

1 An ntc thermistor *T* and a 10 kΩ resistor *R* are connected in series with each other and a 12 V battery, as shown in Figure 5.

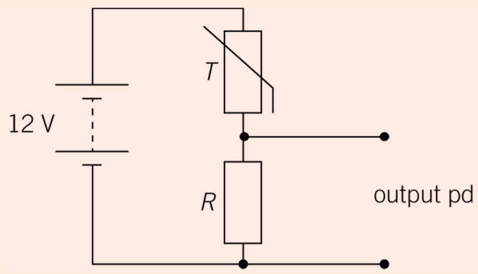


Figure 5

a At a particular temperature, the output pd is 4.0 V. Calculate the resistance of the thermistor at this temperature.

b State and explain how the output pd changes as the temperature of the thermistor increases.

2 The resistor in the above circuit is replaced by a variable resistor *VR*. When the thermistor is at a particular temperature, the variable resistor is adjusted until the output pd is 6.0 V. The thermistor is then cooled to a lower temperature. State and explain how the variable resistor must be adjusted to keep the output voltage at 6.0 V.

3 Draw a potential divider that gives an output pd between 0 and 5.0 V using a 0–100 kΩ variable resistor, a resistor *R*, and a 9.0 V battery.

4 In the circuit in Figure 5, the resistor is replaced by a 100 kΩ resistor, and the thermistor is replaced by an LDR that has a dark resistance of 500 kΩ and a resistance of 50 kΩ in a darkroom when the room lights are on. Calculate the output pd from this circuit in:

a darkness

b in the darkroom when the room lights are on.