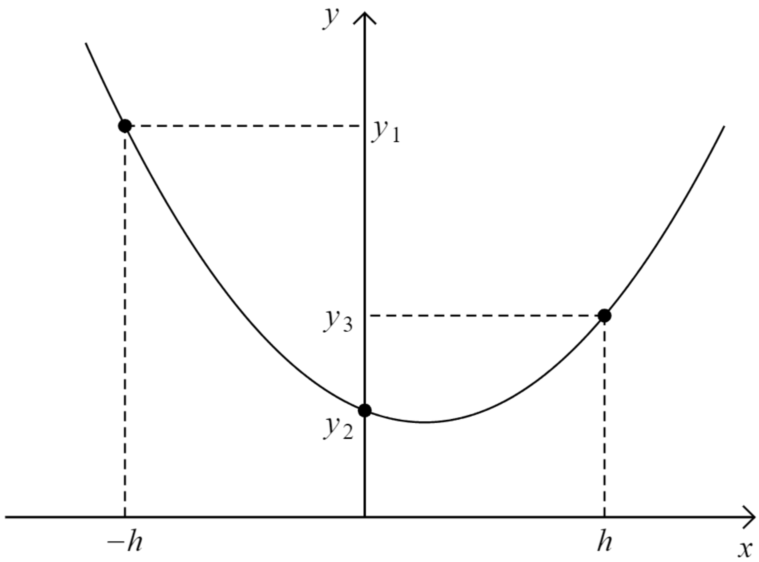
1. [Maximum points: 30]

In this problem you will investigate a method for approximating the area between a curve and the x-axis studied by German mathematician Johannes Kepler in the 17th Century and British mathematician Thomas Simpson in the 18th Century.

Let $f(x) = ax^2 + bx + c$ where $a, b, c \in \mathbb{R}$ and $f(x) \geq 0$. The parabola $y = f(x)$ passes through the points $(-h, y_1)$, $(0, y_2)$ and (h, y_3). This is shown in the diagram below.



Let $A = \int_{-h}^{h} f(x)\, dx$.

(a) Determine an expression for A in terms of h, a and c. [3]

(b) In terms of a, b, c and/or h write down an expression for [3]

(i) y_1

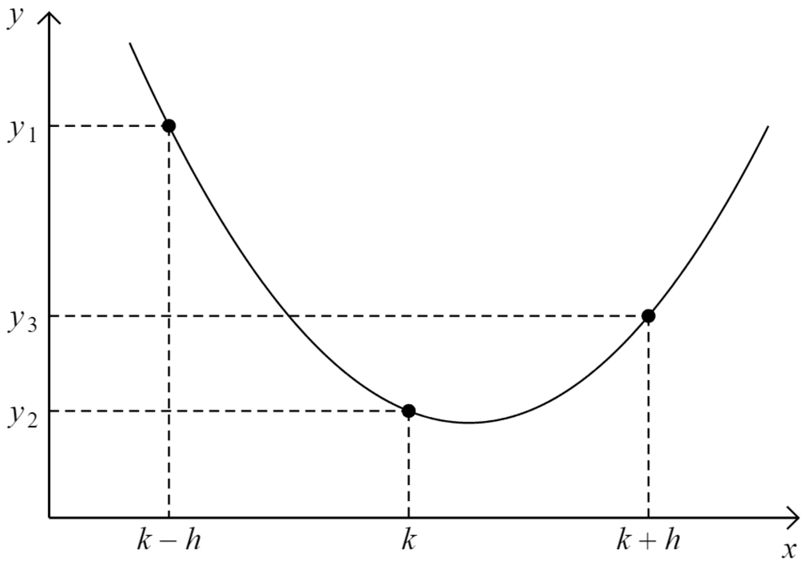
(ii) y_2

(iii) y_3

(c) Hence show that $A = \dfrac{h(y_1 + 4y_2 + y_3)}{3}$. [2]

The problem continues on the next page

The diagram below shows the graph of $y = f(x)$ translated by the vector $\begin{pmatrix} k \\ 0 \end{pmatrix}$ where $k \in \mathbb{R}$.



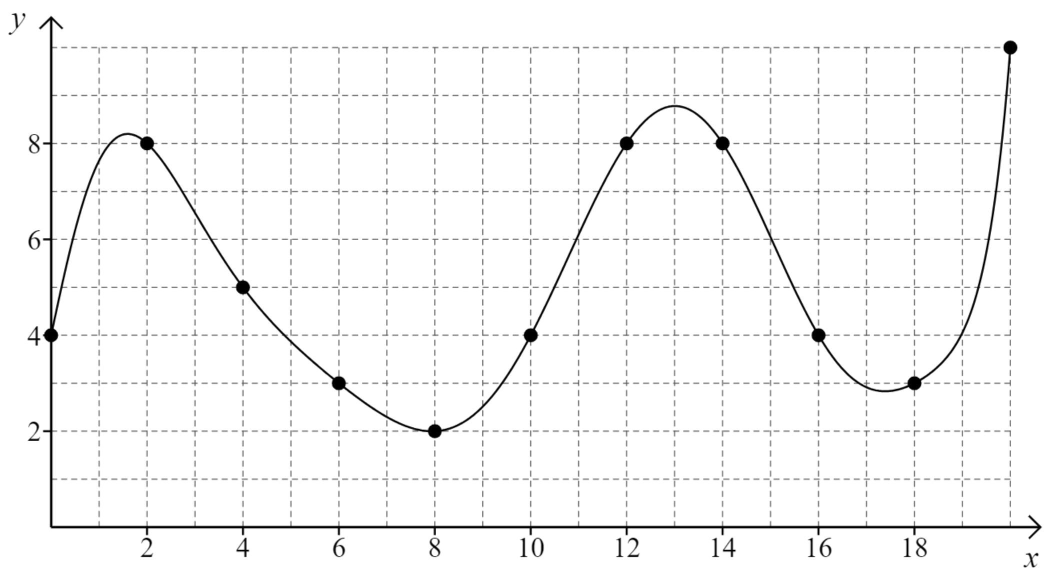
Let the equation of the graph be $y = g(x)$.

(d) Write down $g(x)$ in terms of $f(x)$. [1]

(e) Explain why the result from part (c) still holds if $A = \int_{k-h}^{k+h} g(x)\,dx$. [1]

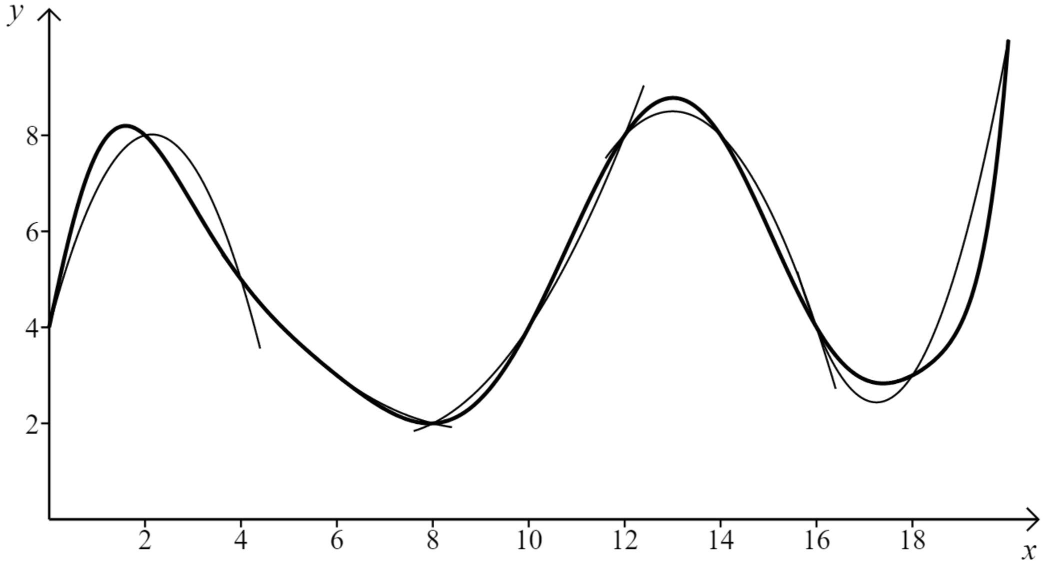
(f) Hence calculate $\int_{2}^{10} x^2 - 4x + 5\,dx$. [3]

The diagram below shows the graph of $y = h(x)$. Eleven points on the curve have been identified.



The shape of the curve can be approximated by dividing the curve into five subintervals of equal width and graphing the parabolas that pass through the three points in each interval.

This is shown in the diagram below where each parabola has been slightly extended to make them more visible.



(g) Use the result of part (c) to estimate $\int_0^{20} h(x)\,dx$. [4]

Now let $f(x) = 2^x$ and $A = \int_0^1 f(x)\,dx$.

(h) By using 100 parabolas over subintervals of equal width show that [6]

$$A \approx \frac{1}{600}\left(3 + 2^{2.995} + (2^{1.995} + 2)\sum_{n=1}^{99} 2^{0.01n}\right)$$

(i) Describe the type of series formed by the part of the expression that uses sigma notation in part (h). [1]

(j) Hence use an appropriate formula to determine the approximate value of A to five decimal places. [3]

(k) Determine the exact value of A. [3]

2. [Maximum points: 28]

In this problem you will use linear regression techniques to find the equation of the curve of best fit for a non-linear set of data.

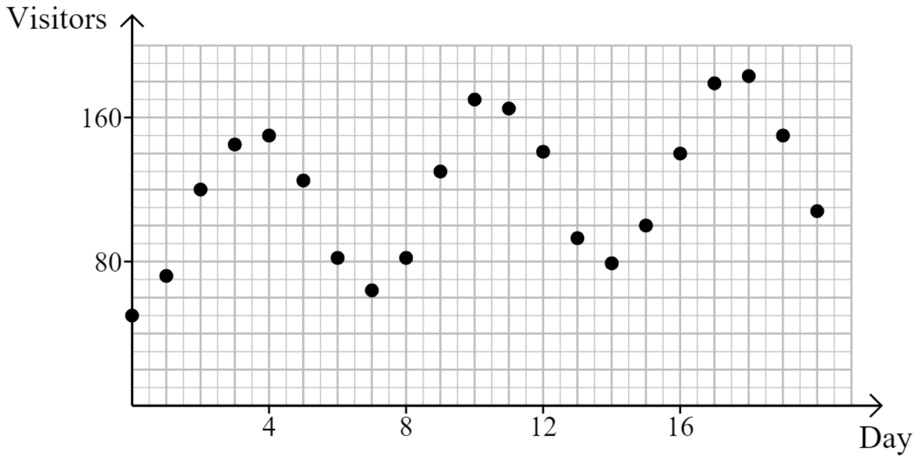
Let $f(x) = ax + b\cos(cx) + d$ and $g(x) = ax + d$ where $a, b, c, d \in \mathbb{R}$ and $c > 0$.

(a) Sketch the graphs of $y = f(x)$ and $y = g(x)$ on the same axes when $a = 1$, $b = 2$, $c = 3$ and $d = 3$ for $0 \leq x \leq 2\pi$. [4]

Paul creates a new website containing mathematics review problems for Diploma Programme students. The table below shows the number of visitors to the website for the first three weeks after the website is published. Day 0 represents a Saturday.

Day (t)	0	1	2	3	4	5	6	7	8	9	10
Visitors (v)	50	72	120	145	150	125	82	64	82	130	170
Day (t)	11	12	13	14	15	16	17	18	19	20	21
Visitors (v)	165	141	93	79	100	140	179	183	150	108	93

The scatter diagram below shows this data.



(b) Explain two possible reasons for the shape of the graph. [4]

Paul wishes to find the equation in the form $v = f(t)$ where f takes the form described above. To find the values of a and d he finds the equation of the least squares regression line for the data in the form $v = g(t)$ where g takes the form described above.

(c) Find the values of a and d. [2]

(d) Explain why $c = \frac{2\pi}{7}$. [3]

Paul decides to use least squares regression techniques to find the value of b manually. Let D_i represent the difference between the actual value of each v_i and the value of $f(t_i)$. Then let

$$D=\sum_{i=0}^{21}(D_i)^2=\sum_{i=0}^{21}(v_i-at_i-b\cos(ct_i)-d)^2$$

where a, c and d take the values calculated in parts (c) and (d).

(e) By finding $\frac{dD}{db}$ show that to minimise the value of D we need [5]

$$-\sum_{i=0}^{21}v_i\cos(ct_i)+a\sum_{i=0}^{21}t_i\cos(ct_i)+b\sum_{i=0}^{21}\cos^2(ct_i)+d\sum_{i=0}^{21}\cos(ct_i)=0$$

(f) Find the value of $\sum_{i=0}^{21}v_i\cos(ct_i)$. [4]

(g) Given that [2]

$$\sum_{i=0}^{21}t_i\cos(ct_i)=10.5 \quad \sum_{i=0}^{21}\cos^2(ct_i)=11.5 \quad \sum_{i=0}^{21}\cos(ct_i)=1.00$$

find the value of b that minimises the value of D.

(h) Use your equation to estimate the amount of visitors the website will receive on day 60. [2]

(i) Comment on the accuracy of your answer to part (h). [2]

1. (a) We have

$$A = \int_{-h}^{h} ax^2 + bx + c\,dx = \left[\frac{ax^3}{3} + \frac{bx^2}{2} + cx\right]_{-h}^{h} \qquad \text{M1}$$

This is equal to

$$\frac{2ah^3}{3} + 2ch \qquad \text{A1A1}$$

(b)

(i) $ah^2 - bh + c$ A1

(ii) c A1

(iii) $ah^2 + bh + c$ A1

(c) We have $\dfrac{h(y_1 + 4y_2 + y_3)}{3} = \dfrac{h(2ah^2 + 2c + 4c)}{3} = \dfrac{2ah^3}{3} + 2ch = A$ M1A1

(d) $g(x) = f(x-k)$ A1

(e) A horizontal translation does not affect the area. A1

(f) The midpoint is $\dfrac{2+10}{2} = 6$ and $h = \dfrac{10-2}{2} = 4$. A1

So we have

$$\frac{4(1 + 4\times 17 + 65)}{3} = \frac{536}{3} \qquad \text{M1A1}$$

(g) Use part (c) with $h = 2$. This gives M1

$$\frac{2(4 + 4\times 8 + 2\times 5 + 4\times 3 + 2\times 2 + 4\times 4 + 2\times 8 + 4\times 8 + 2\times 4 + 4\times 3 + 1\times 10)}{3} \qquad \text{A1A1}$$

Which is equal to 104. A1

(h) For each subinterval we have $h = \dfrac{1}{200}$. A1

Use the results from part (c). So the value of A is approximately M1

$$1 + 4\times 2^{1/200} + 2\times 2^{2/200} + 4\times 2^{3/200} + 2\times 2^{4/200} + \cdots + 2\times 2^{198/200} + 4\times 2^{199/200} + 2)/600$$

A1A1

This is equal to

$$\frac{1}{600}\left(3 + 4\sum_{n=1}^{100} 2^{(2n-1)/200} + 2\sum_{n=1}^{99} 2^{n/100}\right) = \frac{1}{600}\left(3 + 4\cdot 2^{199/200} + (4\cdot 2^{-1/200} + 2)\sum_{n=1}^{99} 2^{n/100}\right) \qquad \text{A1A1}$$

Which simplifies to the given expression.

(i) Geometric A1

(j) Use the geometric series formula. M1

$$\frac{1}{600}\left(3+2^{2.995}+(2^{1.995}+2)\cdot\frac{2^{0.01}(1-2^{0.99})}{1-2^{0.01}}\right)=1.44270$$ A1A1

(k) We have $c=\ln 2$. A1

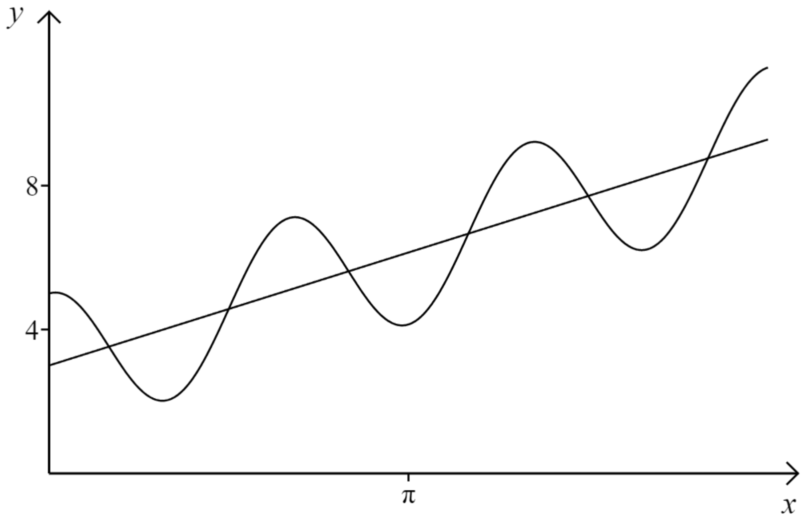
So

$$A=\int_0^1 e^{x\ln 2}\,dx=\frac{e^{\ln 2}-1}{\ln 2}=\frac{1}{\ln 2}$$ M1A1

2. (a) The domain is $0 \le x \le 2\pi$. A1

The y-axis scale is appropriate. A1

The shape and position of each graph is approximately correct. A1A1



(b) The values of v are periodic because perhaps less people will use the website on a weekend when they dont have school. A1 R1

The values of v are gradually increasing perhaps because the popularity of the website is increasing through, for example, word of mouth or advertising. A1 R1

(c) $a = 2.01$ A1

$d = 98.0$ A1

(d) There are seven days in a week so the period of the cosine function should be 7. So R1

$$\frac{2\pi}{c} = 7$$ M1

Giving

$$c = \frac{2\pi}{7}$$ A1

(e) Use the chain rule M1

$$\frac{dD}{db} = \sum_{i=0}^{21} 2(v_i - at_i - b\cos(ct_i) - d)(-\cos(ct_i))$$ A1

Expand

$$\frac{dD}{db} = -2\sum_{i=0}^{21} v_i \cos(ct_i) + 2a\sum_{i=0}^{21} t_i \cos(ct_i) + 2b\sum_{i=0}^{21} \cos^2(ct_i) + 2d\sum_{i=0}^{21} \cos(ct_i)$$ A1

Set equal to 0 and simplify M1

$$-\sum_{i=0}^{21} v_i \cos(ct_i) + a\sum_{i=0}^{21} t_i \cos(ct_i) + b\sum_{i=0}^{21} \cos^2(ct_i) + d\sum_{i=0}^{21} \cos(ct_i) = 0$$ A1

(f) The values of each $v_i \cos(ct_i)$ are A1A1A1

50	44.9	-26.7	-131	-135	-27.8	51.1	64	51.1	-28.9	-153
-149	-31.4	58.0	79	62.3	-31.2	-161	-165	-33.4	67.3	93

The sum is equal to -452. A1

(g) We have

$$452 + 2.01 \times 10.5 + 11.5b + 98.0 \times 1.00 = 0$$ M1

So

$$b = -49.7$$ A1

(h) $2.01 \times 60 - 49.7\cos(2\pi \times 60/7) + 98.0 = 189$ M1A1

(i) This is not accurate since the value in (h) is an extrapolation. A1R1