# 3940 AS 01

# Main

## Question 1

### (i) Number 925150

The four significant figure is 925200.

So, the absolute error $E_A$ is:

$$E_A = |x - \overline{x}| = |925150 - 925200| = 50 \tag{1}$$

The relative error $E_R$ is:

$$E_R = |\frac{x - \overline{x}}{x}| = |\frac{925150 - 925200}{925150}| = 0.000054 \tag{2}$$

The percantage error $E_P$ is:

$$E_P = E_R \times 100\% = 0.0054\% \tag{3}$$

### (ii) Number 26.36125

The four significant figure is 26.36

So, the absolute error $E_A$ is:

$$E_A = |x - \overline{x}| = |26.36125 - 26.36| = 0.00125 \tag{4}$$

The relative error $E_R$ is:

$$E_R = |\frac{x - \overline{x}}{x}| = \frac{E_A}{|x|} = \frac{0.00125}{26.36125} = 0.000047 \tag{5}$$

The percentage error $E_P$ is:

$$E_P = E_R \times 100\% = 0.0047\% \tag{6}$$

## Question 2

Given $x = 5.675$, $y = 4.373$, and $z = 3.373$, so values of $u_1 = x(y - z)$ and $u_2 = xy - xz$ are:

$$
\begin{aligned}
u_1 &= 5.675 \times (4.373 - 3.373) \\
&= 5.676 \times 1.000 = 5.675 \\
u_2 &= 5.675 \times 4.373 - 5.675 \times 3.373 \\
&= 24.82 - 19.14 = 5.68
\end{aligned} \tag{7}
$$

$$
\begin{aligned}
E_{R,u_1} &= |(5.675 - 5.675)/5.675| = 0 \\
E_{R,u_2} &= |(5.675 - 5.68)/5.675| = 0.00088
\end{aligned} \tag{8}
$$

Since $E_{R,u_1} < E_{R,u_2}$, $u_1$ or the former one is more accurate.

# Question 3

**(a)** $-10\pi + 6e - \frac{3}{32}$

The exact value of the expression is $-15.200$.

Suppose the expression is performed using three-digit rounding arithmetic. Rounding $\pi$ and $e$ to three digits gives $\pi^* = 3.14$ and $e^* = 2.72$, so the result of the expression is:

$$
-10 \times 3.14 + 6 \times 2.72 - 3/32 = -15.1 - 0.0938 = -15.2 \tag{9}
$$

Therefore, the absolute error $E_A$ and the relative error $E_R$ are

$$
\begin{aligned}
E_A &= 0 \\
E_R &= 0
\end{aligned} \tag{10}
$$

**(b)** $\frac{\sqrt{13}+\sqrt{11}}{\sqrt{13}-\sqrt{11}}$

The exact value of the expression is 23.958

Suppose the expression is performed using three-digit rounding arithmetic. Let $a = \sqrt{13}, b = \sqrt{11}$, so rounding $a$ and $b$ to three digits gives $a^* = 3.61$ and $b^* = 3.32$, such that the result of the expression is:

$$
\frac{\sqrt{13}+\sqrt{11}}{\sqrt{13}-\sqrt{11}} = \frac{3.61 + 3.32}{3.61 - 3.32} = 23.90 \tag{11}
$$

Therefore, the absolute error $E_A$ and the relative error $E_R$ are

$$
\begin{aligned}
E_A &= |23.958 - 23.90| = 0.058 \\
E_R &= \frac{E_A}{|23.958|} = 0.0024
\end{aligned} \tag{12}
$$

# Question 4

Given function

$$
f(x) = x^3 - 6.1x^2 + 3.2x + 1.5 \tag{13}
$$

where $x = 4.71$

Using three digit rounding arithmetic, we have

$$\begin{aligned} x^2 &= 4.71 \times 4.71 = 22.1841 = 22.2 \\ x^3 &= x^2 \times x = 22.2 \times 4.71 = 104.562 = 105 \\ 6.1x^2 &= 6.1 \times 22.2 = 135.42 = 135 \\ 3.2x &= 3.2 \times 4.71 = 15.072 = 15.1 \end{aligned} \tag{14}$$

Therefore, the value of $f(x)$ where $x = 4.71$ could be evaluated as

$$f(4.71) = 105 - 135 + 15.1 + 1.5 = -13.4 \tag{15}$$

## Question 5

Given function

$$r = 3h(h^6 - 2) = 3h^7 - 6h \tag{16}$$

For 5% error in $h$ when $h = 1$,

Therefore,

$$\begin{aligned} \frac{dr}{r} &= (21h^6 - 6)\frac{dh}{r} \\ &= (21 - 6) \times 0.05 \\ &= 0.75 \end{aligned} \tag{17}$$

So, the percentage error $E_P$ is 75%.

## Question 6

The derivate of the function $f(x)$,

$$f'(2) = 3.5e^{0.5x} = 3.5 \times e^{0.5\times 2} = 9.513986 \tag{18}$$

Given approximate derivative of the function $f(x)$

$$f'(x) \approx \frac{f(x+h) - f(x)}{h} \tag{19}$$

When $h = 0.3$,

$$\begin{aligned} f'(2) &\approx \frac{f(2+0.3) - f(2)}{0.3} \\ &= \frac{7e^{0.5\times 2.3} - 7e}{0.3} \\ &= 10.26459 \end{aligned} \tag{20}$$

So, the percentage error $E_P$ is

$$\begin{aligned} E_P|_{h=0.3} &= \frac{|10.26459 - 9.513986|}{9.513986} \times 100\% \\ &= 7.88948\% \end{aligned} \tag{21}$$

When $h = 0.15$,

$$\begin{aligned} f'(2) &\approx \frac{f(2+0.15) - f(2)}{0.15} \\ &= \frac{7e^{0.5\times 2.15} - 7e}{0.15} \\ &= 9.87985 \end{aligned} \tag{22}$$

So, the percentage error $E_P$ is

$$\begin{aligned} E_P|_{h=0.15} &= \frac{|9.87985 - 9.513986|}{9.513986} \times 100\% \\ &= 3.845539\% \end{aligned} \tag{23}$$

Finally, the percentage error $E_P$ from $h = 0.3$ to $h = 0.15$

*Not sure which percentage error is correct since the question desciption is vague, so I calculated all of them.* 🙂

$$E_P = |\frac{10.26459 - 9.87985}{9.87985}| = 3.894189\% \tag{24}$$

## Question 7

### (a) 0 10000001010 1001001100000000000000000000000000000000000000000000

The leftmost bit is 0, which means the number is positive.

The next 11 bits, 10000001010, gives the characteristic and are equivalent to the decimal number

$$c = 1 \cdot 2^{10} + 1 \cdot 2^3 + 1 \cdot 2^1 = 1034 \tag{25}$$

The exponential part of the number is, therefore, $2^{1034-1023} = 2^{11}$

The final 52 bits specify that the mantissa is

$$f = 0.5^1 + 0.5^4 + 0.5^7 + 0.5^8 = 0.57421875 \tag{26}$$

As a consequence, this machine number precisely represents the decimal number

$$\begin{aligned} (-1)^s 2^{c-1023}(1+f) &= (-1)^0 2^{1034-1023}(1+f) \\ &= 3224.0 \end{aligned} \tag{27}$$

### (b) 1 10000001010 1001001100000000000000000000000000000000000000000000

The leftmost bit is 1, which means the number is negative.

The other part of the number is the same as question (a).

So the value is $-3224.0$

## Question 8

The second-degree Taylor polynomial of $f(x) = (1+x)^{1/2}$ about $x = 0$ is

$$T(2) = f(0) + f'(0)x + \frac{f''(0)}{2}x^2 = 1 + \frac{1}{2}x - \frac{1}{8}x^2 \tag{28}$$

When $x = 0.05$, we have

$$f(0.05) \approx T(2)|_{x=0.05} = 1 + 0.5 \times 0.05 - 0.125 \times 0.05 = 1.01875 \tag{29}$$

Since the actual value of $f(0.05)$ is $(1+0.05)^{0.5} = 1.0246950766$

So, the truncation error of the approximate $f(0.05)$ is

$$E_T = |f(0.05) - T(2)|_{x=0.05}| = 0.005945076596 \tag{30}$$

## Question 9

Given $f(x) = x^3 - e^{-x}$, $x_0 = 0.5$

### (a)

$$\begin{aligned} T_2(x) &= f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2 + O((x-0.5)^3 \\ &= 0.5^3 - e^{-0.5} + (3 \cdot 0.5^2 + e^{-0.5})(x - 0.5) + \frac{6 \cdot 0.5 - e^{-0.5}}{2}(x-0.5)^2 + O((x-0.5)^3 \\ &= -0.481531 + 1.35653(x-0.5) + 1.19673(x-0.5)^2 + O((x-0.5)^3 \end{aligned} \tag{31}$$

### (b)

When $x = 0.8$, the value of the Taylor Polynomial $T_2(x)$ is

$$T_2(0.8) = 0.0331337 \tag{32}$$

while the value of $f(0.8)$ is $f(0.8) = 0.8^3 - e^{-0.8} = 0.062671$

Therefore, the actual error $|f(0.8) - T_2(0.8)| = |0.062671 - 0.0331337| = 0.0295373$

## Question 10

### (a) $\{\frac{2^n+3}{2^n+7}\}$

$$\begin{aligned} \lim_{n\to\infty} \frac{2^n+3}{2^n+7} &= \lim_{n\to\infty} \frac{1+3/2^n}{1+7/2^n} \\ &= 1 \end{aligned} \tag{33}$$

The rate of convergence could be calculated as below

$$
\begin{aligned}
\mu &= \lim_{n\to\infty} |\frac{a_{n+1}-1}{a_n-1}| \\
&= \lim_{n\to\infty} |\frac{(2^{n+1}+3)/(2^{n+1}+7)-1}{(2^n+3)/(2^n+7)-1}| \\
&= \lim_{n\to\infty} |\frac{(2^{n+1}+3-2^{n+1}-7)/(2^{n+1}+7)}{(2^n+3-2^n-7)/(2^n+7)}| \\
&= \lim_{n\to\infty} |\frac{2^n+7}{2^{n+1}+7}| \\
&= \frac{1}{2}
\end{aligned} \tag{34}
$$

**(b)** $\{\frac{1-2n^2}{3n^2+n-1}\}$

$$
\lim_{n\to\infty} \frac{1-2n^2}{3n^2+n-1} = \lim_{n\to\infty} \frac{1/n^2-2}{3+1/n-1/n^2} = -\frac{2}{3} \tag{35}
$$

The rate of convergence could be calculated as below

$$
\begin{aligned}
\mu &= \lim_{n\to\infty} |\frac{a_{n+1}}{a_n}| \\
&= \lim_{n\to\infty} |\frac{\frac{1-2(n+1)^2}{3(n+1)^2+n}+\frac{2}{3}}{\frac{1-2n^2}{3n^2+n-1}+\frac{2}{3}}| \\
&= \lim_{n\to\infty} |\frac{\frac{3+2n}{9(n+1)^2+3n}}{\frac{1+2n}{9n^2+3n-3}}| \\
&= \lim_{n\to\infty} |\frac{(3+2n)(9n^2+3n-3)}{(1+2n)(9(n+1)^2+3n)}| \\
&= 1
\end{aligned} \tag{36}
$$

**(c)** $\{\ln\frac{2n-1}{2n+1}\}$

$$
\lim_{n\to\infty} \ln\frac{2n-1}{2n+1} = \lim_{n\to\infty} \ln\left(1-\frac{2}{2n+1}\right) = 0 \tag{37}
$$

The rate of convergence could be calculated as below

$$
\begin{aligned}
\mu &= \lim_{n\to\infty} |\frac{a_{n+1}}{a_n}| \\
&= \lim_{n\to\infty} |\frac{\ln\frac{2n+1}{2n+3}}{\ln\frac{2n-1}{2n+1}}| \\
&= -\frac{2}{3}/-\frac{2}{3} \\
&= 1
\end{aligned} \tag{38}
$$

**(d)** $\{\sin\frac{1}{n}\}$

$$
\lim_{n\to\infty} \sin\frac{1}{n} = 0 \tag{39}
$$

The rate of convergence could be calculated as below

$$
\begin{aligned}
\mu &= \lim_{n\to\infty} |\frac{a_{n+1}}{a_n}| \\
&= \lim_{n\to\infty} |\frac{\sin\frac{1}{n+1}}{\sin\frac{1}{n}}| \\
&= \lim_{n\to\infty} |\frac{\frac{1}{n+1}}{\frac{1}{n}}| \\
&= 1
\end{aligned} \tag{40}
$$