

1. We build a computer, where the real numbers are represented using 5 digits as explained below:

S	A	B	C	E
-----	-----	-----	-----	-----

where

- S is the sign bit; 0 is positive and 1 is negative
- A, B, C : First three significant digits in decimal expansion with decimal point occurring between A and B
- E is the exponent in base 10 with a bias of 5
- All digits after the third significant digit are chopped off
- $+0$ is represented by setting $S = 0$ and $A = 0$; B, C, E can be anything
- -0 is represented by setting $S = 1$ and $A = 0$; B, C, E can be anything
- $+\infty$ is represented by setting $S = 0$ and $A = B = C = E = 9$
- $-\infty$ is represented by setting $S = 1$ and $A = B = C = E = 9$
- Not A Number is represented by setting S to be other than 0 and 1.

For example, the number $\pi = 3.14159\dots$ is represented as follows. Chopping off after the third significant digit, we have $\pi = +3.14 \times 10^0$. Hence, the representation of π on our machine is:

0	3	1	4	5
---	---	---	---	---

The number $-0.001259\dots$ is represented as follows. Chopping off after the third significant digit, we have -1.25×10^{-3} . Hence, the representation of on our machine is:

1	1	2	5	2
---	---	---	---	---

Now answer the following questions:

- (a) How many non-zero Floating Point Numbers (from now on abbreviated as FPN) can be represented by our machine?

Solution:

The representation of $-\infty$ on our machine is:

1	9	9	9	9
---	---	---	---	---

$$\text{True exponent} = E - \text{bias} = 9 - 5 = 4$$

With decimal point occurring between A and B , the decimal point can float upto 4 places to the right, which implies that $-\infty = -99900$

The next smallest negative floating point number represented on this machine is -99899 . Chopping off after the third significant, we have -9.98×10^4

$$E = \text{True exponent} + \text{bias} = 4 + 5 = 9$$

Hence, the representation of on our machine is:

1	9	9	8	9
---	---	---	---	---

The representation of $+\infty$ on our machine is:

0	9	9	9	9
---	---	---	---	---

$$\text{True exponent} = E - \text{bias} = 9 - 5 = 4$$

With decimal point occurring between A and B , the decimal point can float upto 4 places to the right, which implies that $+\infty = 99900$

The next largest positive floating point number represented on this machine is $+99899$. Chopping off after the third significant, we have $+9.98 \times 10^4$

$$E = \text{True exponent} + \text{bias} = 4 + 5 = 9$$

Hence, the representation of on our machine is:

0	9	9	8	9
---	---	---	---	---

(b) How many FPN are in the following intervals?

- (9,10)

Solution:

The smallest number exceeding 9 that can be represented exactly on this machine is 9.01

0	9	0	1	5
---	---	---	---	---

The largest number not exceeding 10 that can be represented exactly on this machine is 9.99

0	9	9	9	5
---	---	---	---	---

Hence, there are 99 FPN in the interval (9,10)

- (10,11)

Solution:

The smallest number exceeding 10 that can be represented exactly on this machine is 10.1

0	1	0	1	6
---	---	---	---	---

The largest number not exceeding 11 that can be represented exactly on this machine is 10.9

0	1	0	9	6
---	---	---	---	---

Hence, there are 9 FPN in the interval (10,11)

- (0,1)

Solution:

The smallest number exceeding 0 that can be represented exactly on this machine is 0.00001

0	1	0	0	0
---	---	---	---	---

The largest number not exceeding 1 that can be represented exactly on this machine is 0.999

0	9	9	9	4
---	---	---	---	---

(c) Identify the smallest positive and largest positive FPN on the machine

Solution:

Smallest positive floating point number:

$$S = 0 \quad E = 0 \quad ABC = 100$$

$$\text{True exponent} = E - \text{bias} = 0 - 5 = -5$$

$$10^{-5} = 0.00001$$

0	1	0	0	0
---	---	---	---	---

Largest positive floating point number:

$$S = 0 \quad E = 9 \quad M = 998$$

$$\text{True exponent} = E - \text{bias} = 9 - 5 = 4$$

$$9.98 \times 10^4 = 99800.0$$

0	9	9	8	9
---	---	---	---	---

(d) Identify the machine precision

Solution:

Machine precision is defined as the difference between the smallest number exceeding 1 that can be represented on the machine and 1. The smallest number exceeding 1 that can be represented on this machine is $1.01 = 1 + 10^{-2}$

Hence, the machine precision is $\epsilon_m = 10^{-2} = 0.01$

(e) What is the smallest positive integer not representable exactly on this machine?

(f) Consider solving the following recurrence on our machine:

$$a_{n+1} = 5a_n - 4a_{n-1}$$

with $a_1 = a_2 = 2.932$. Compute a_n for $n \in 3, 4, 5, 6, 7$ on our machine (work out what the machine would do by hand). Note a_1, a_2 would be chopped to three significant digits to begin with. Next note that at each step in the recurrence $5a_n$ and $4a_{n-1}$ would be chopped down to the first three significant digits before the subtraction is performed.

Solution:

$$a_{n+1} = 5a_n - 4a_{n-1}$$

0	2	9	3	5
---	---	---	---	---

Compute a_3 : $a_1 = a_2 = 2.93$

$$\begin{aligned} a_3 &= 5a_2 - 4a_1 \\ &= 5(2.93) - 4(2.93) \\ &= 14.65 - 11.72 \\ &= 1.46 \times 10^1 - 1.17 \times 10^1 \\ &= 0.29 \times 10^1 = 2.90 \end{aligned}$$

0	2	9	0	5
---	---	---	---	---

Compute a_4 : $a_2 = 2.93$, $a_3 = 2.90$

$$\begin{aligned} a_4 &= 5a_3 - 4a_2 \\ &= 5(2.90) - 4(2.93) \\ &= 14.50 - 11.72 \\ &= 1.45 \times 10^1 - 1.17 \times 10^1 \\ &= 0.28 \times 10^1 = 2.80 \end{aligned}$$

0	2	8	0	5
---	---	---	---	---

Compute a_5 : $a_3 = 2.90$, $a_4 = 2.80$

$$\begin{aligned} a_5 &= 5a_4 - 4a_3 \\ &= 5(2.80) - 4(2.90) \\ &= 14.0 - 11.6 \\ &= 1.40 \times 10^1 - 1.16 \times 10^1 \\ &= 0.24 \times 10^1 = 2.40 \end{aligned}$$

0	2	4	0	5
---	---	---	---	---

Compute a_6 : $a_4 = 2.80$, $a_5 = 2.40$

$$\begin{aligned} a_6 &= 5a_5 - 4a_4 \\ &= 5(2.40) - 4(2.80) \\ &= 12.0 - 11.2 \\ &= 1.20 \times 10^1 - 1.12 \times 10^1 \\ &= 0.08 \times 10^1 = 8.0 \times 10^{-1} \end{aligned}$$

0	8	0	0	4
---	---	---	---	---

Compute a_7 : $a_5 = 2.40$, $a_6 = 0.80$

$$\begin{aligned} a_7 &= 5a_6 - 4a_5 \\ &= 5(0.8) - 4(2.40) \\ &= 4.0 - 9.6 = -5.6 \end{aligned}$$

1	5	6	0	5
---	---	---	---	---

2. Consider the following integral:

$$I_n = \int_0^1 x^{2n} \sin(\pi x) dx$$

(a) Obtain a recurrence for I_n in terms of I_{n-1} . (HINT: Integration by parts)

Solution:

$$I_n = \int_0^1 x^{2n} \sin(\pi x) dx$$

$$I_{n-1} = \int_0^1 x^{2(n-1)} \sin(\pi x) dx$$

Integration by parts:

$$\int u dv = uv - \int v du$$

$$u = x^{2n} \quad dv = \sin(\pi x) dx$$

$$du = 2nx^{2n-1} dx \quad v = -\frac{1}{\pi} \cos(\pi x)$$

$$\begin{aligned} I_n &= \int_0^1 x^{2n} \sin(\pi x) dx = uv \Big|_0^1 - \int_0^1 v du \\ &= -\frac{1}{\pi} x^{2n} \cos(\pi x) \Big|_0^1 + \frac{2n}{\pi} \int_0^1 x^{2n-1} \cos(\pi x) dx \\ &= -\frac{1}{\pi} 1^{2n} \cos(\pi) + \frac{2n}{\pi} \int_0^1 x^{2n-1} \cos(\pi x) dx \end{aligned}$$

$$u = x^{2n-1} \quad dv = \cos(\pi x) dx$$

$$du = (2n-1)x^{2n-2} dx \quad v = \frac{1}{\pi} \sin(\pi x)$$

$$\begin{aligned} &= -\frac{1}{\pi} 1^{2n} \cos(\pi) + \frac{2n}{\pi} \left[\frac{1}{\pi} x^{2n-1} \sin(\pi x) \Big|_0^1 - \frac{2n-1}{\pi} \int_0^1 x^{2n-2} \sin(\pi x) dx \right] \\ &= -\frac{1}{\pi} 1^{2n} \cos(\pi) + \frac{2n}{\pi} \left[\frac{1}{\pi} 1^{2n-1} \sin(\pi) - \frac{2n-1}{\pi} \int_0^1 x^{2n-2} \sin(\pi x) dx \right] \\ I_n &= -\frac{1}{\pi} 1^{2n} \cos(\pi) + \frac{2n}{\pi^2} 1^{2n-1} \sin(\pi) - \frac{2n(2n-1)}{\pi^2} I_{n-1} \end{aligned}$$

(b) Evaluate I_0 by hand

Solution:

$$I_n = \int_0^1 x^{2n} \sin(\pi x) dx$$

$$I_0 = \int_0^1 x^0 \sin(\pi x) dx$$

$$I_0 = \int_0^1 \sin(\pi x) dx$$

$$= \left[\frac{1}{\pi} (-\cos \pi x) \right]_0^1$$

$$= \frac{1}{\pi} (-\cos \pi) - \frac{1}{\pi} (-\cos \pi)$$

$$= \frac{1}{\pi} (1 + 1) = \frac{2}{\pi} \approx 0.63662$$

(c) Use the recurrence to obtain I_n for $n \in 1, 2, \dots, 15$ in Octave

```
#!/usr/bin/env octave
% File: recurrence.m
% Script to obtain I(n) by recurrence relation

N = 15;           % number of terms in the recurrence
I = zeros(N,1);  % initialize vector of N elements
I(1) = 0.63662;  % initial condition I(1) = 0.63662

for n = 1:N
    I(n+1) = (-1/pi) * 1^(2*n) * cos(pi) \
              + (2*n)/(pi^2) * 1^(2*n - 1) * sin(pi) \
              - ((2*n)*(2*n - 1)/(pi^2)) * I(n);
endfor

disp(I)
```

Output:

```
6.3662e-01
1.8930e-01
8.8144e-02
5.0384e-02
3.2433e-02
2.2552e-02
1.6692e-02
1.0503e-02
6.2906e-02
-1.6320e+00
6.3155e+01
-2.9560e+03
1.6533e+05
-1.0888e+07
8.3402e+08
-7.3519e+10
```

(d) Use wolframalpha to obtain I_n by directly performing the integral for $n \in 1, 2, \dots, 15$

Solution:

$$I_n = \int_0^1 x^{2n} \sin(\pi x) dx$$

$$n = 1 \quad I_1 = \int_0^1 x^{2(1)} \sin(\pi x) dx \approx 0.18930$$

$$n = 2 \quad I_2 = \int_0^1 x^{2(2)} \sin(\pi x) dx \approx 0.088144$$

$$n = 3 \quad I_3 = \int_0^1 x^{2(3)} \sin(\pi x) dx \approx 0.050384$$

$$n = 4 \quad I_4 = \int_0^1 x^{2(4)} \sin(\pi x) dx \approx 0.032433$$

$$n = 5 \quad I_5 = \int_0^1 x^{2(5)} \sin(\pi x) dx \approx 0.022561$$

$$\begin{aligned}n = 6 & \quad I_6 = \int_0^1 x^{2(6)} \sin(\pi x) dx \approx 0.016574 \\n = 7 & \quad I_7 = \int_0^1 x^{2(7)} \sin(\pi x) dx \approx 0.012679 \\n = 8 & \quad I_8 = \int_0^1 x^{2(8)} \sin(\pi x) dx \approx 0.010006 \\n = 9 & \quad I_9 = \int_0^1 x^{2(9)} \sin(\pi x) dx \approx 0.0080938 \\n = 10 & \quad I_{10} = \int_0^1 x^{2(10)} \sin(\pi x) dx \approx 0.0066802 \\n = 11 & \quad I_{11} = \int_0^1 x^{2(11)} \sin(\pi x) dx \approx 0.0056060 \\n = 12 & \quad I_{12} = \int_0^1 x^{2(12)} \sin(\pi x) dx \approx 0.0047708 \\n = 13 & \quad I_{13} = \int_0^1 x^{2(13)} \sin(\pi x) dx \approx 0.0041089 \\n = 14 & \quad I_{14} = \int_0^1 x^{2(14)} \sin(\pi x) dx \approx 0.0035754 \\n = 15 & \quad I_{15} = \int_0^1 x^{2(15)} \sin(\pi x) dx \approx 0.0031393\end{aligned}$$

(e) Explain your observation

The values obtained from the octave script recurrence relation and the integral calculated are same from $n = 0$ to $n = 4$ and differs for other values of n .