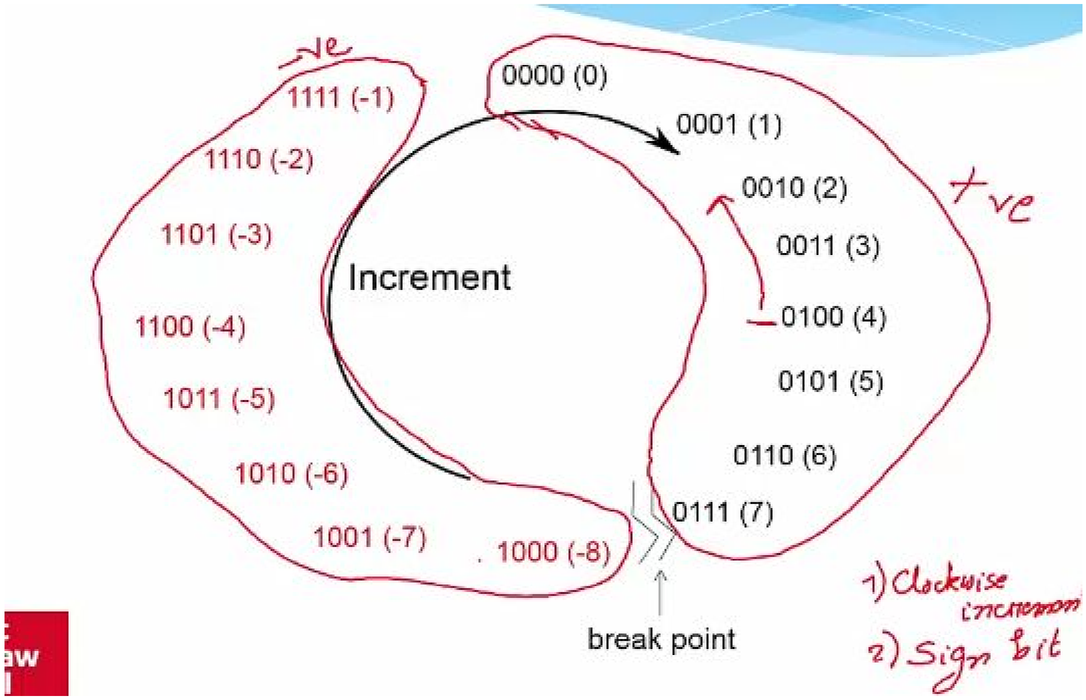
Assignment 1

Question 1:

2’s complement method can be understood by Number circle with negative numbers If we write 4 bits numbers. All the numbers with leading 0 are taken positive and leading with 1 are negative.



Like if we want to represent $(-2)_{10}$ in binary we will first represent |-2| = 2 so the binary representation will be $(0010)_2$ then we take its complement $(1101)_2$ then we add $(0001)_2$ to that which makes it $(1110)_2$. we can see that we represent [-8,7] integers value in using only 4 bits.

Why to have this much complicated method if we can just have msd to represent sign. Because when we do that we have

i) two representation of zeros

a) 0000
b) 1000

ii) Addition/Subtraction will be difficult lets say 5 - 3

$(0101)_2$ - $(1011)_{2,}$ will be difficult because we don't know how to deal with the most significant 1 in representation of 3.

Question 2:

$(2.25)_{10}$ can be broken into 2 parts $(2+0.25)_{10}$. So, binary representation for respective terms are $(10.01)_2$ is equivalent to $(1.001 * 2^1)$

Now we now for ieee 754 format storage is done by 1 sign bit 8 bits for exponent and 23 bits for mantissa.

Sign bit is 1 as number is negative
Exponent adjusted = exponent unadjusted + $2^{(8-1)} - 1$
So exponent adjusted = $(128)_{10} = (1000\ 0000)_2$
Now mantissa = $(001\ 0000\ 0000\ 0000\ 0000\ 0000)_2$
So full representation will be 1 10000000 00100000000000000000000 in binary
<u>1100</u> <u>0000</u> <u>0001</u> <u>0000</u> <u>0000</u> <u>0000</u> <u>0000</u> <u>0000</u> = C 0 1 0 0 0 0 0
Hexadecimal representation will be 0x C0100000.

Question 3:
When a number is less than one. Normal representation will lead to loss of precision. So keep the precision of the number and cause underflow error.So, Denormal numbers were introduced to fill this gap around zero.
Let's take f = 2^{-126}
g=f/2;
if(g==0)
print('not expected') ←If we don't have a representation for 2^{-127} we can see g==0 going to be true. we don't want this so to avoid this kind of situation we introduced denormal numbers.