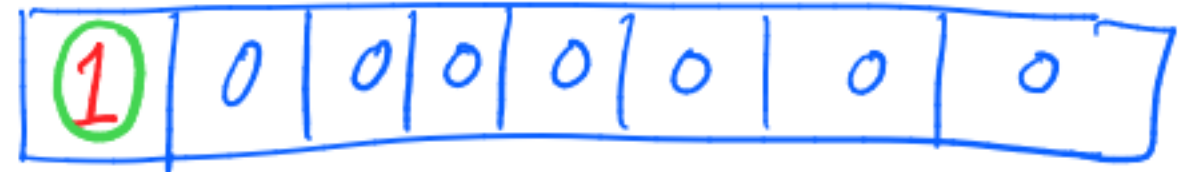


-5

$$(5)_{10} = (101)_2$$

1 → The number is negative
 0 → " " " positive.

2's complement

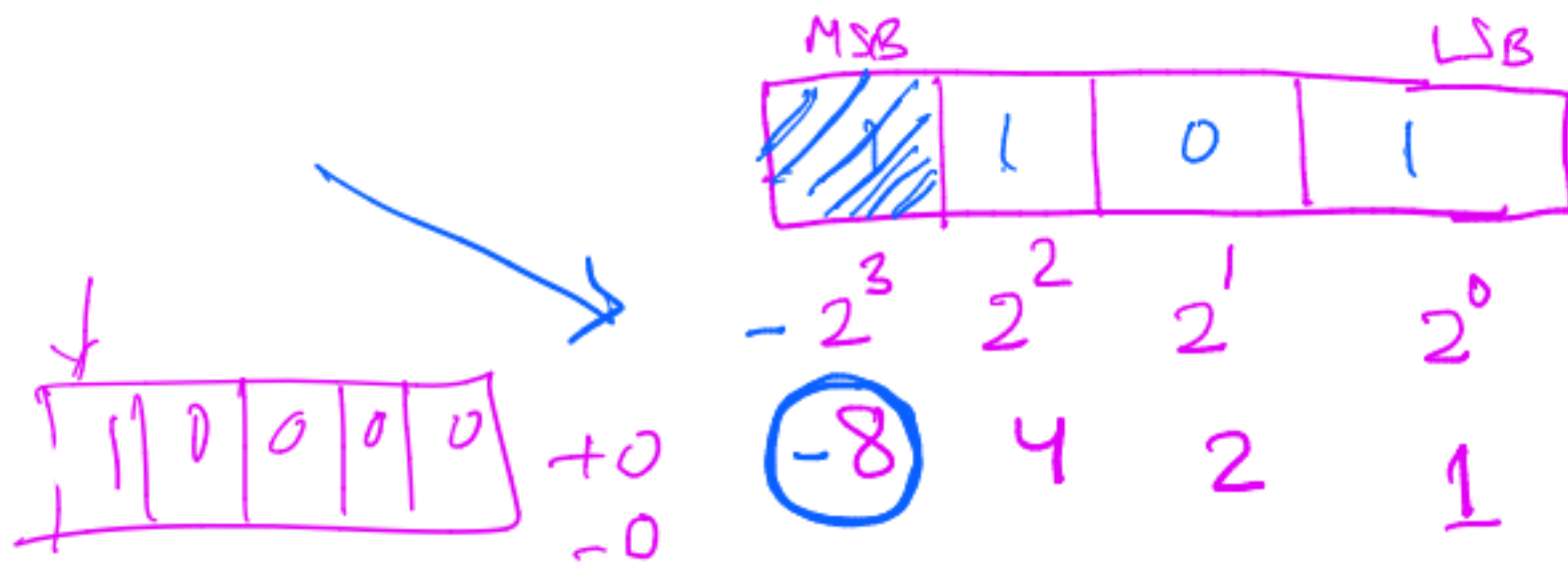


-0
 +0

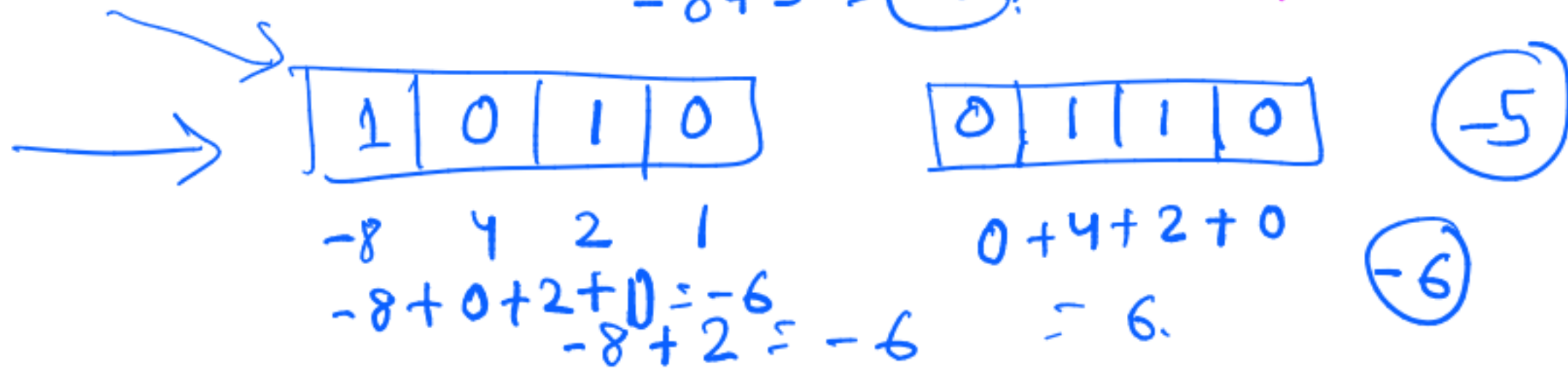
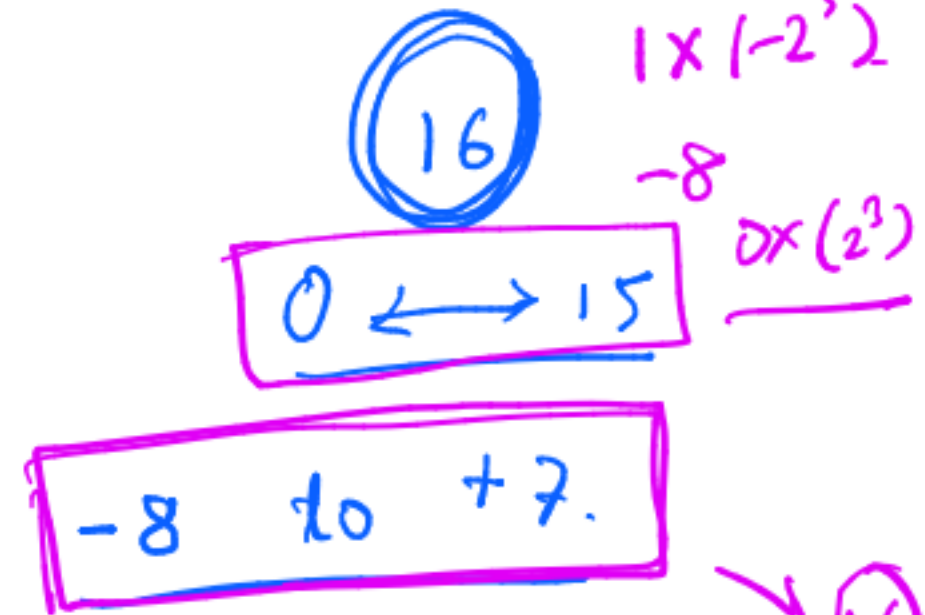
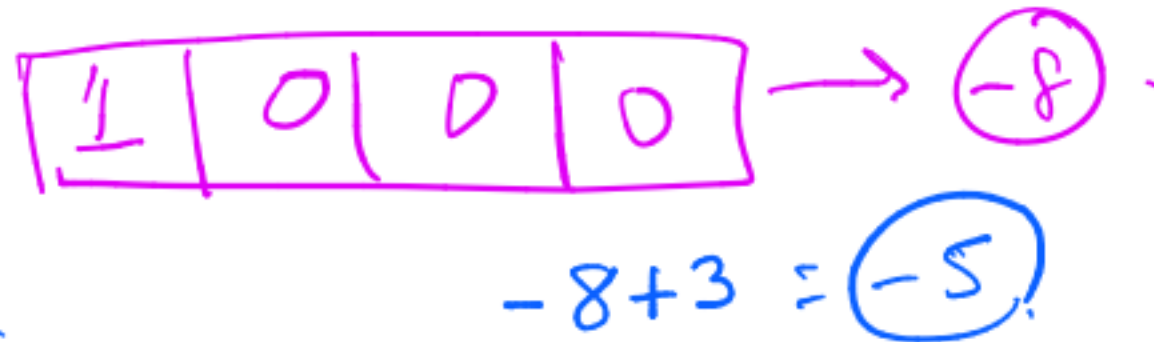
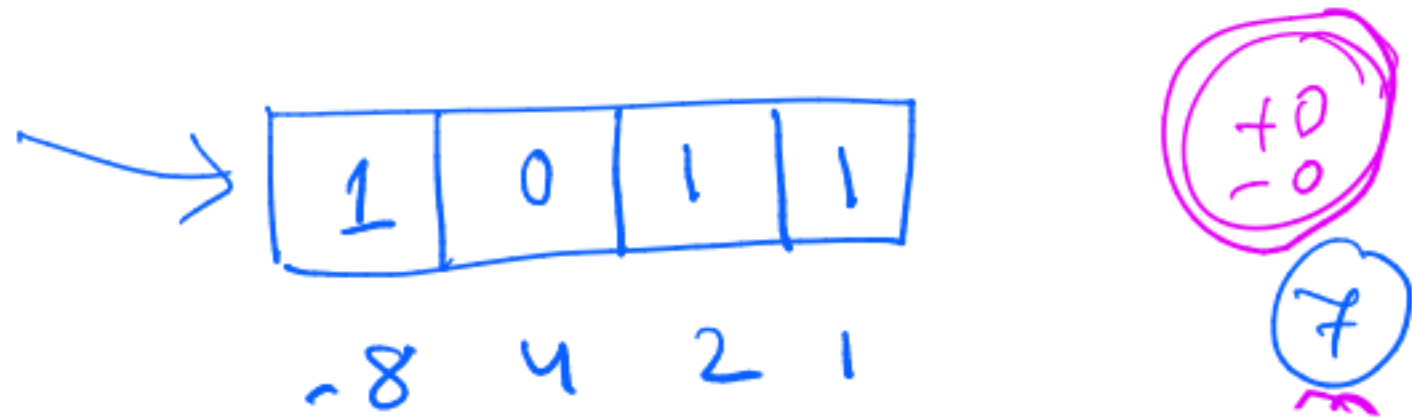
Integers

Floating-point

→ Bias



$$0 \times 1 + 0 \times 4 + 1 \times 2 + 0 \times 1$$



Augmented

short, long, signed and unsigned.

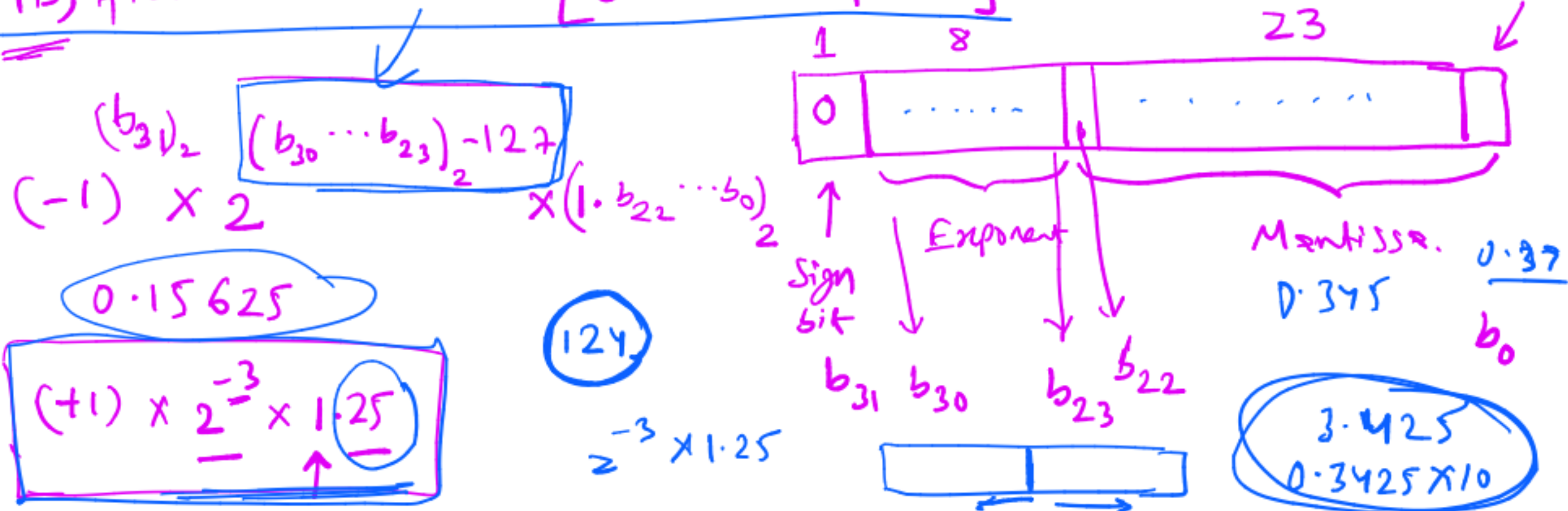
Short int \leq int (2 bytes).

Long int \geq int 4 bytes.

- (2B) Unsigned int — 0 to 65535 $[0 \text{ to } 2^{16}-1]$ ←
 (2B) Signed int — -32768 to +32767 $[-2^{15} \text{ to } 2^{15}-1]$ ←
 (1B) char — 0 to 255 $B \rightarrow \text{byte}$
 — -128 to 127 $b \rightarrow \text{bit}$
 (4B) Long int — 0 to $2^{32}-1$ $4 \times 8 = 32 \text{ bits}$
 — -2^{31} to $2^{31}-1$

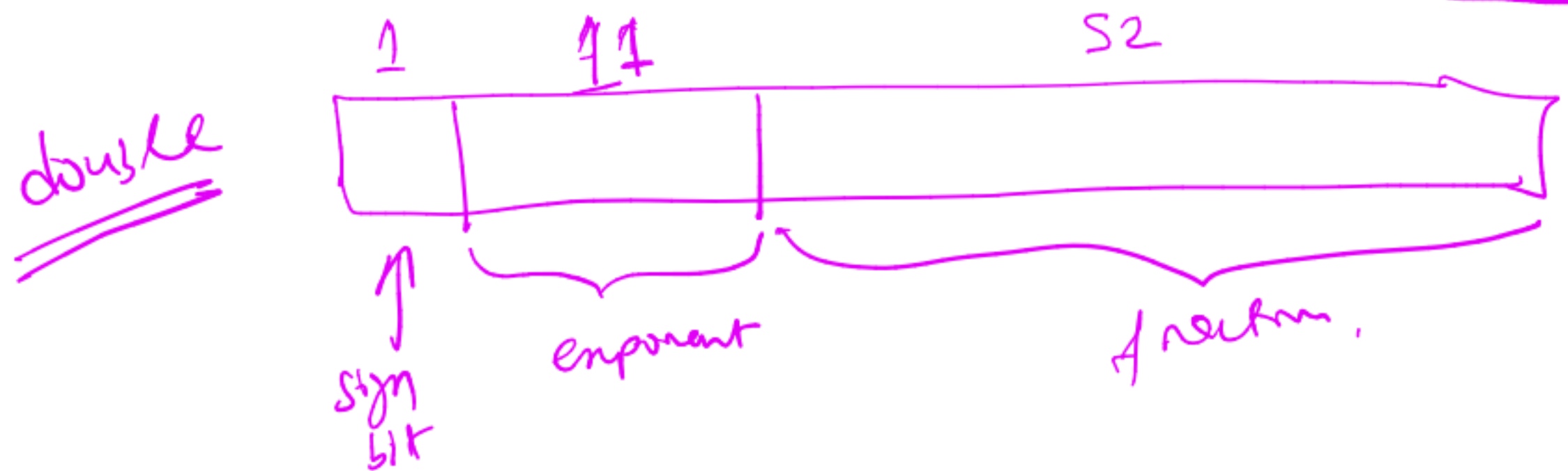
signed $\left[-2^{n-1} \text{ to } 2^{n-1}-1 \right]$
 unsigned $\left[0 \text{ to } 2^n-1 \right]$

(4B) float — [6 decimal places]



- (4B) float - single precision - $1.2 \text{ E } -38$ to $3.4 \text{ E } +38$ [6 decimal places]
- (8B) double - double precision - $2.3 \text{ E } -308$ to $1.7 \text{ E } +308$ [15 decimal places]
- (12B) long double - $3.4 \text{ E } -4932$ to $1.1 \text{ E } +4932$ [18 decimal places]

4.356724



IEEE 754

(4B) long int — signed — -2^{31} to $2^{31}-1$
 unsigned — 0 to $2^{32}-1$

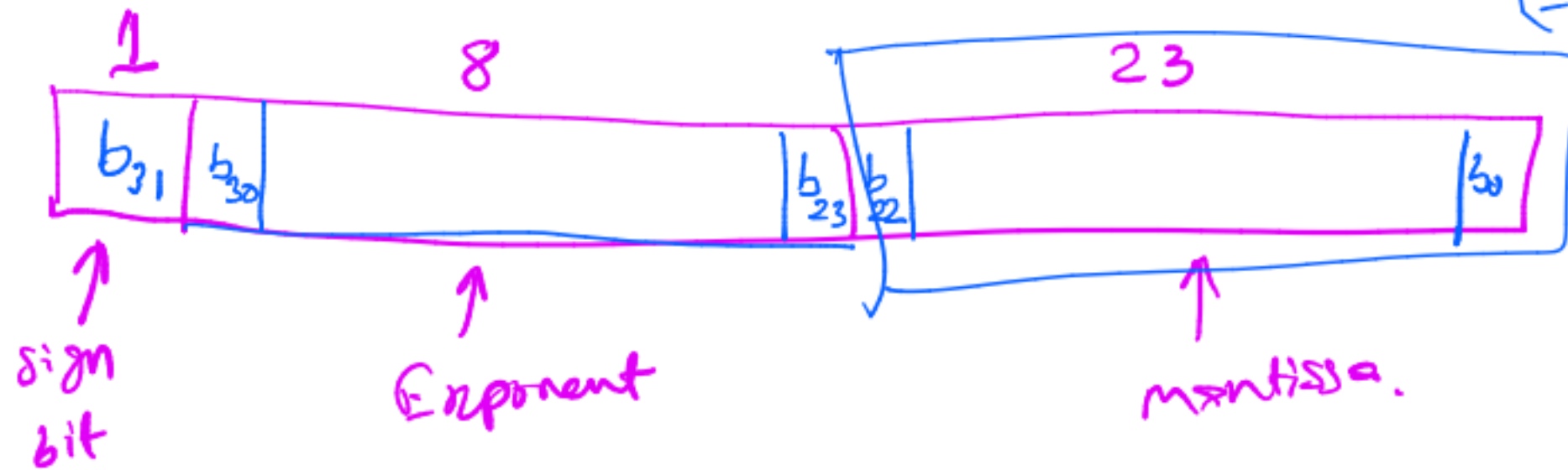
$$0.625$$

$$(-1)^0 \times 2^{-3} \times 1.5$$

(4B) float —

$$0.125$$

$$(-1)^0 \times 2^{-4} \times 1.25$$



$$0.625$$

$$(-1)^0 \times 2^{-1} \times 1.25$$

IEEE

$$(-1)^{(b_{31})_2} \times 2^{(b_{30} \dots b_{23})_2 - 127} \times 1.(b_{22} \dots b_0)_2$$

$$0.15625$$

$$124 - 127$$

$$(-1)^0 \times 2^{-3} \times 1.25 \rightarrow 1 \times 0.125 \times 1.25$$

$$2^{-2} \rightarrow 0.25$$

0.625



↓
sign
bit
0

↓
 2^{-1}

$\frac{[b_{30} \dots b_{23}] - \text{bias}}{2}$

↓
1.25

↑
[bias]

(FFFF)

$(-1)^0 \times 2^{\frac{126-127}{2}} \times 1.25$

2^{-1}

$(-1)^{\text{sign bit}} \times 2^{\frac{[b_{30} \dots b_{23}] - 127}{2}} \times 1.[b_{22} \dots b_0]$

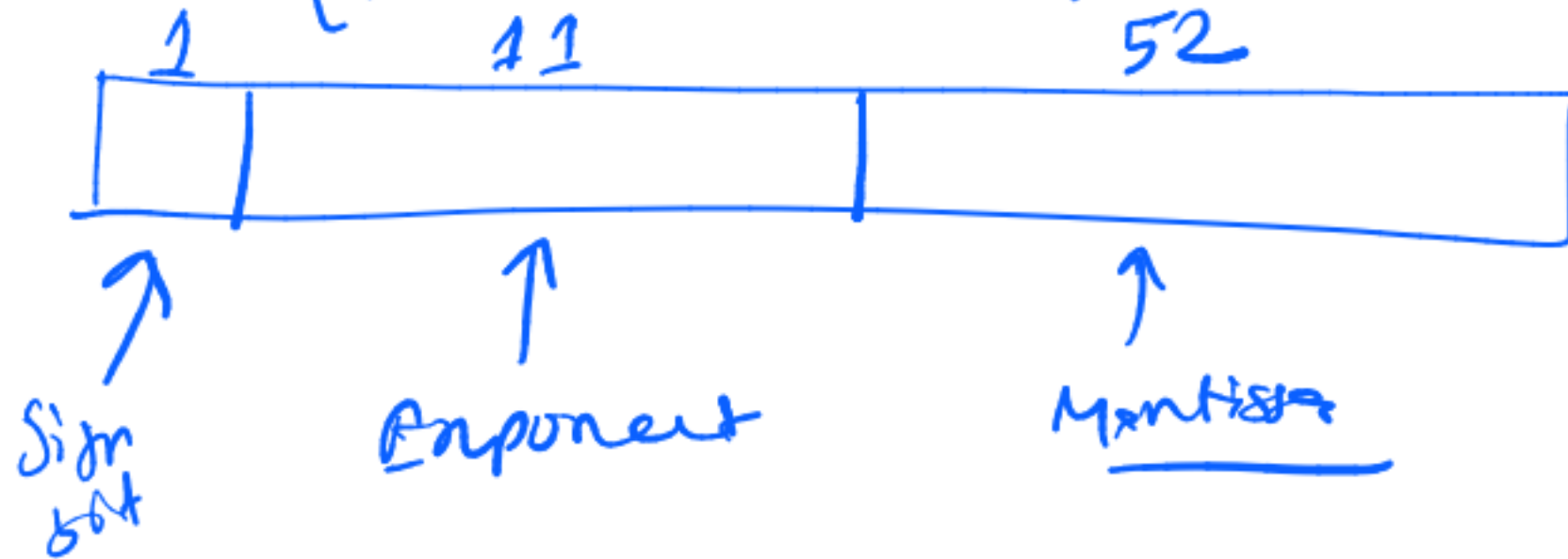
$(-1)^0 \times 2^{-1} \times 1.25$

(4B) floating-point number - [6 decimal places]

$1.2 \text{E}-38$ to $3.4 \text{E}+38$

1.2×10^{-38} to $3.4 \times 10^{+38}$

(8B) double - [15 decimal places]



(64)

length 64 bits
chap 11