

Single

$$S \Rightarrow 1$$

$$f \Rightarrow 23$$

$$C \Rightarrow 8$$

$$C = m + 127 \quad (m = C - 127)$$

$$-126 < m \leq 127 \quad (*)$$

$$P = n+1 = 24$$

$$E_m = 2^{-23}$$

Double $X = (-1)^S (1.f) \times 2^M$

$$S \Rightarrow 1$$

$$f \Rightarrow 52$$

$$C \Rightarrow 11$$

$$C = m + 1023$$

$$-1022 < m \leq 1023 \quad (*)$$

$$P = 53$$

$$E_m = 2^{-52}$$

shift

$$\left. \begin{array}{l} c: \text{all zeros} \\ f: \text{all zeros} \end{array} \right\} \underline{\text{ZERO}}$$

$$c: \text{all zeros}$$

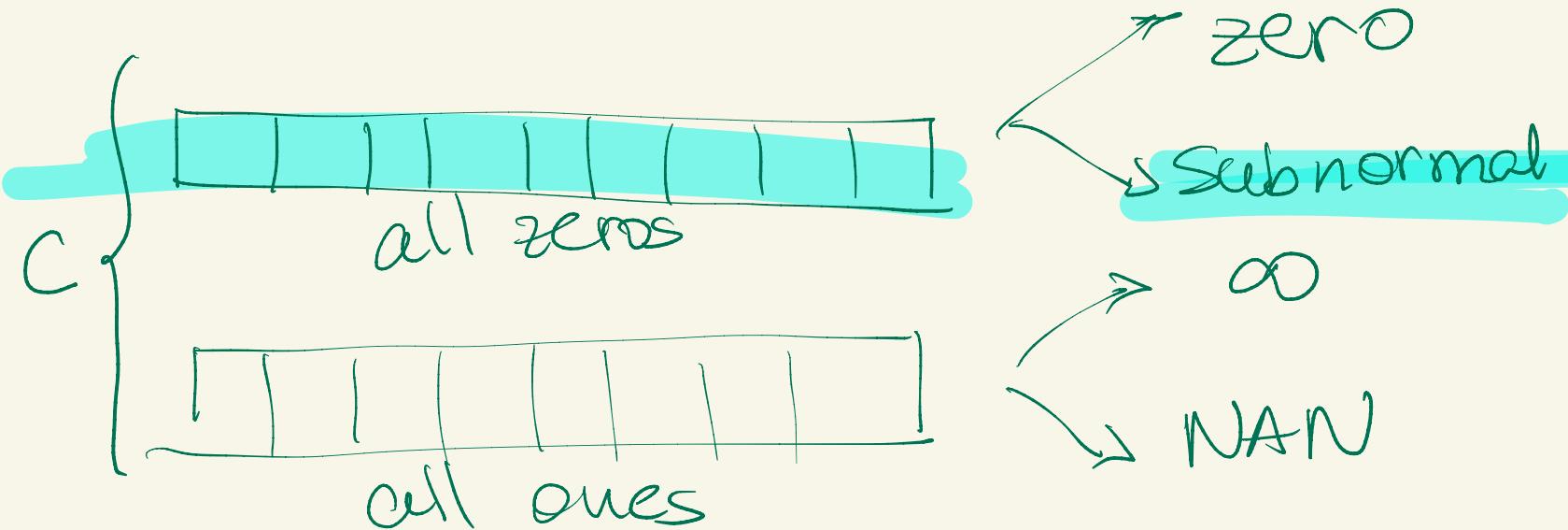
$$\left. \begin{array}{l} f: \text{anything but zero} \end{array} \right\} \text{subnormal}$$

$$\left. \begin{array}{l} c: \text{all ones} \\ f: \text{all zeros} \end{array} \right\} \infty$$

$$c: \text{all ones}$$

$$\left. \begin{array}{l} f: \text{anything but zeros} \end{array} \right\} \text{NAN}$$

Special cases:



C 1 0 1 0 | ... | 1 → smallest C

C 1 1 1 | ... | 0 → largest C

$$UFL = 2^L$$

$$OFL = 2^{(l+1)}(1-2^{-p})$$

$$(1-f) \times 2^m$$

$$2^{-126}$$

2^{-1022}

$$2^{(127+1)}(1-2^{-24})$$

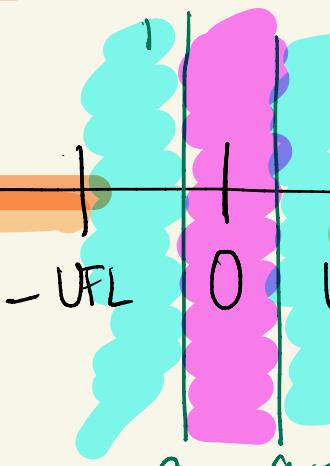
$$2^{(1023+1)}(1-2^{-53})$$

$$single \quad 2^{-23} \times 2^{-126}$$

$$double \quad 2^{-52} \times 2^{-1022}$$

OFL

-OFL



$-a \quad a$

$$x = (-1)^s (0.s) \times 2^L$$

52 bits
 $0.010010 \dots 1010 \times 2^L$

23, 52
 $0.0000 \dots 001 \times 2^L$

0.00000001.00

24 53 28 20
0.000 [100 .. 100] x 2
Single

$$t_n = \underline{\hspace{10em}}$$

$$e = f - t_n$$

$h_1 \rightarrow e_1$ given
 0.5

$$h_2 \rightarrow e_2$$
 estimate
 0.25

$$t_n = \sum_{i=0}^n \frac{f^{(i)}(x_0)}{i!} (x - x_0)^i$$

$$e = O(h^{n+1})$$

$$\left. \begin{array}{l} e_1 = \alpha h_1^{n+1} \\ e_2 = \alpha h_2^{n+1} \end{array} \right\} \frac{e_1}{e_2} = \left(\frac{h_1}{h_2} \right)^{n+1}$$

$$e_2 = e_1 \left(\frac{h_2}{h_1} \right)^{n+1}$$