Scientific notation

~~0.0~~1.

1 0100011 10 . . . .

~~0.~~010011

767.54

↓

7.6754×10^2

$1. \_\_\_\_ \times 2^{E-B}$ → k bit exp → $2^{k-1}-1$

$1010.10 \rightarrow 1.01010 \times 2^{11} \rightarrow 3$

$+\ 10.11 \rightarrow 1.011 \times 2^{1} \rightarrow 1$

E = 514

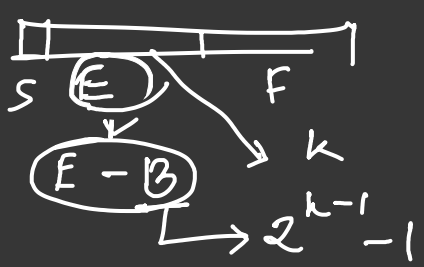
E = 512

10

↓

2^9-1

B = 511



E − B = 3 → 1

E = 3 + B

18

$1010.10 \rightarrow X =$ 0 | 10000001 0 | 01 010 0 0 0 . . . .

1.01010×2^{11}

1 | 10 bits | 21 bits

1 → . . . 0

10.11

1.011×2^1 $Y =$ 0 | 100 0 0 0 0 0 0 0 | 0 1 1 0 0 . . .

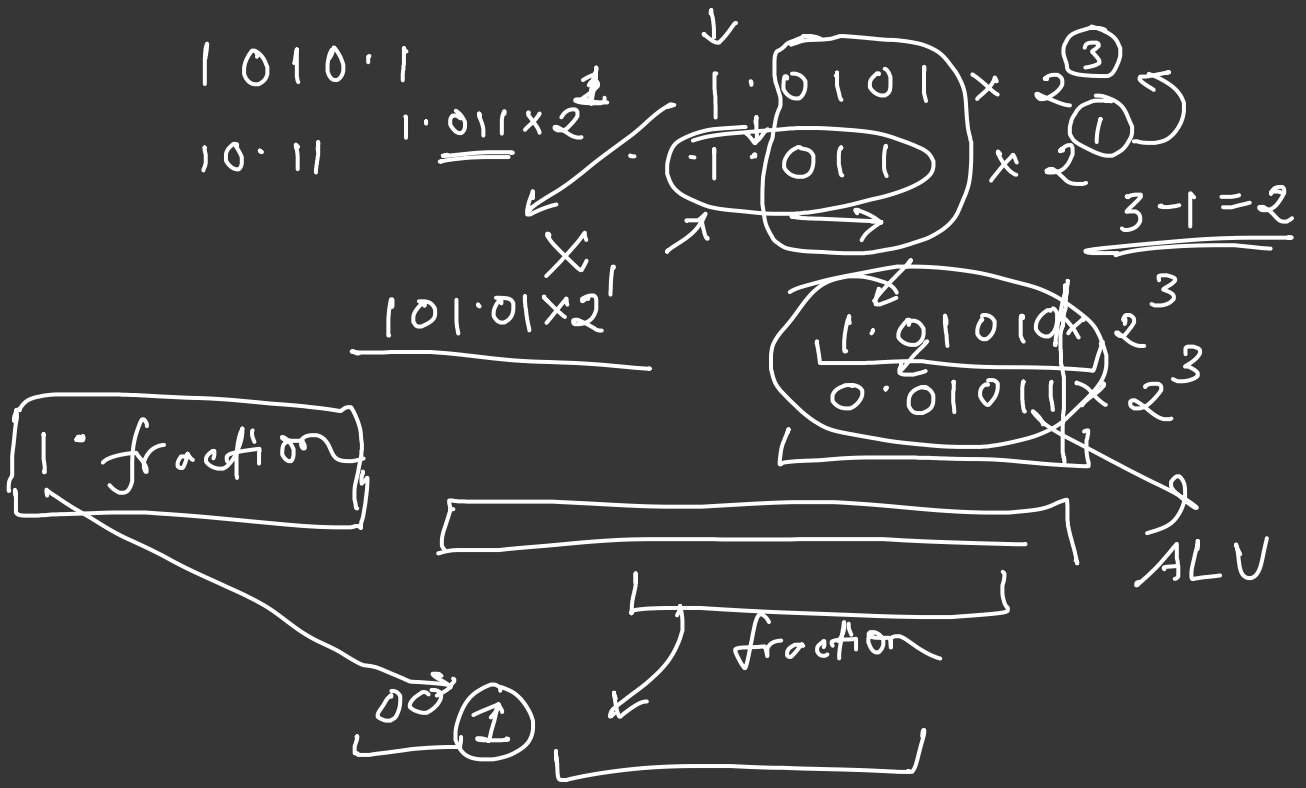
1010.10

10.11

1101.01 + 1.10101×2^{11} → 1000.0010

0

F → 10101



Steps ⓪ sign check → sign bit 1 → negate

① extract fraction, append 001
(hidden bit before fraction)

② find smaller → Right Shift

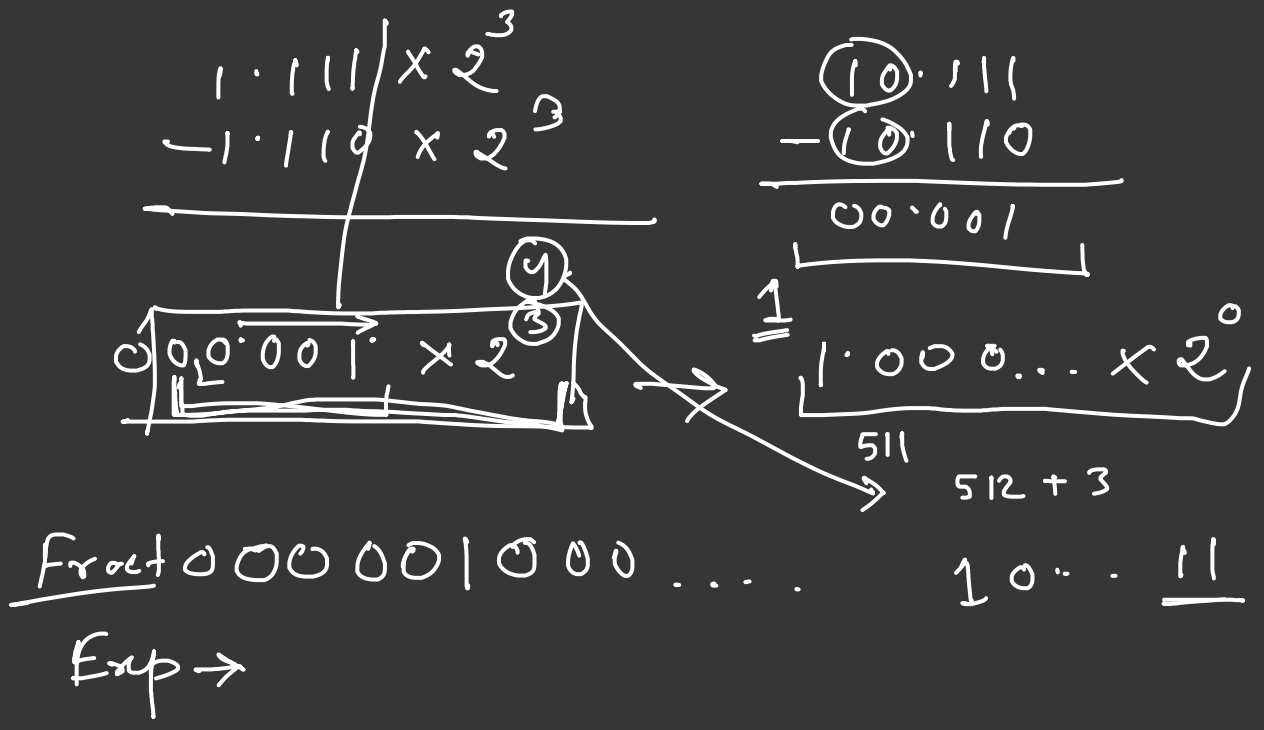
③ Add with ALU

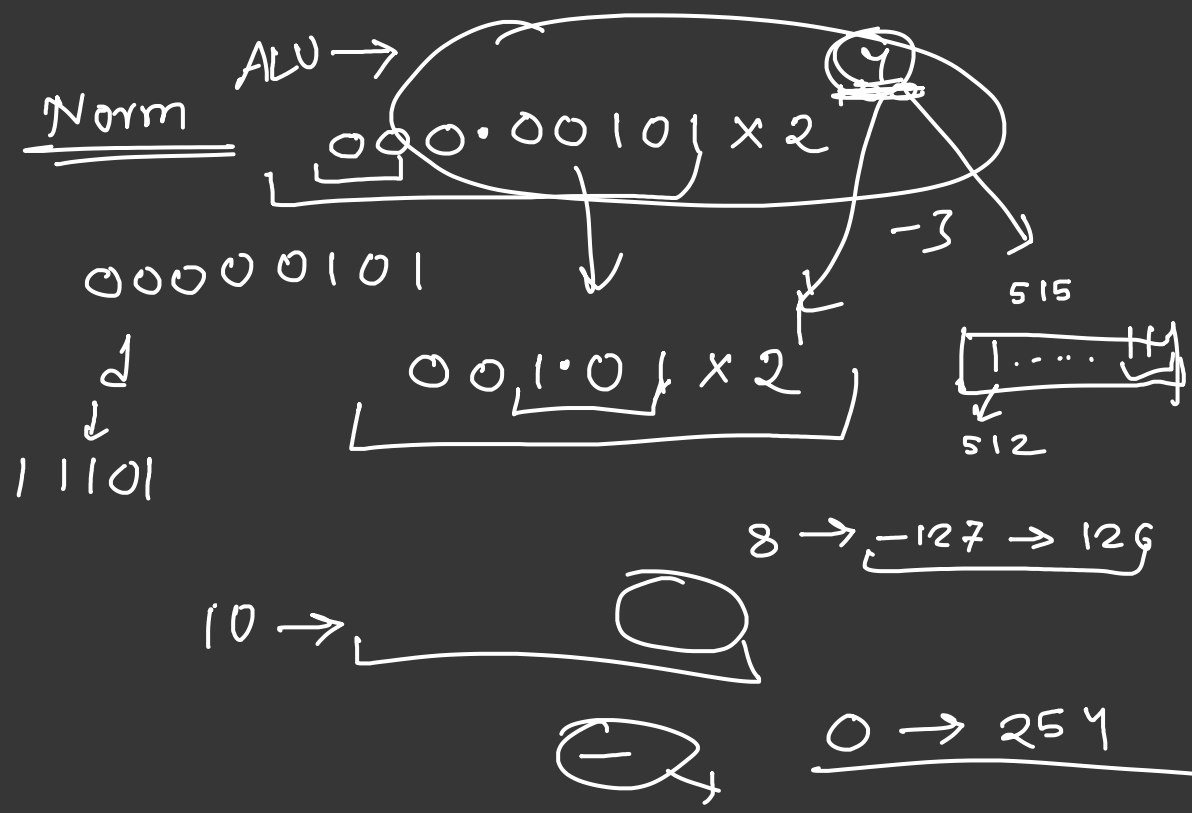
④ Normalize ✓

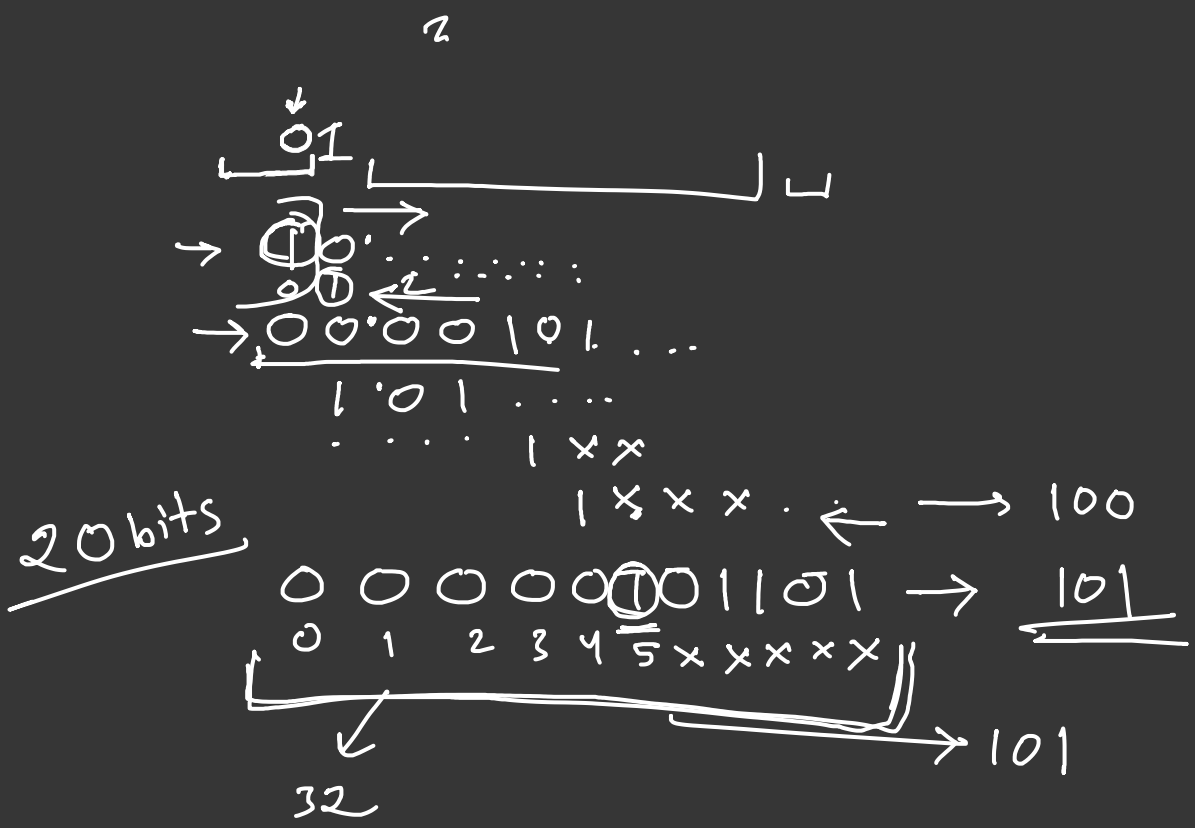
⑤ Round X

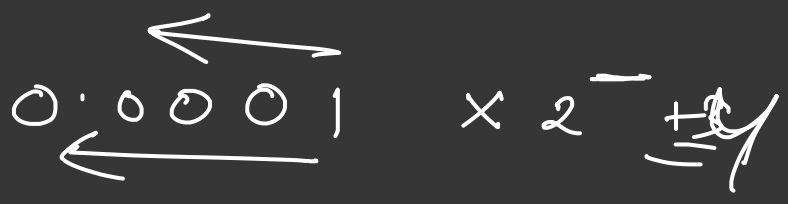
1.10101 × 2^3

0 01 01010 × 2^3
0 00 01011 × 2^3
0 01 10101 × 2^3


1.111 x 2^3
-1.110 x 2^3
00.001 x 2^3
00.001
1.000... x 2^0
511
512 + 3
Fract 000 00 | 000 ....
10.. 11
Exp →


Norm
ALU →
000.00101 x 2
00101
-3
00001 x 2
515
512
11101
8 → -127 → 126
10 →
0 → 254


20 bits
101
32



8 bit allowed E → XOR → 1 → 254 255 X

11

$2^{11} - 2$

$2^n - 2$