

$$\pm M \times 2^{\pm E}$$

1 bit for  $\pm$

$k$  bits for  $\pm E$

$n$  bits for  $M$

$k = 8$  or  $11$

$n = 23$  or  $52$



**Normalized:**  $\pm E$  is not its maximum or minimum value

$$1 \leq M < 2$$



$$\pm E = e + 1 - 2^{k-1}$$

$$M = 1 + f/2^n$$



**Denormalized:**  $\pm E$  is its minimum value (which is negative)

$$0 \leq M < 1$$



$$\pm E = 2 - 2^{k-1}$$

$$M = f/2^n$$



**Infinity:**  $\pm E$  is its maximum value



**Not-a-Number:**  $\pm E$  is its maximum value (many representations!)



$$\pm M \times 2^{\pm E}$$

1 bit for  $\pm$

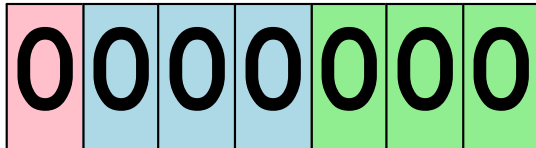
$k$  bits for  $\pm E$

$n$  bits for  $M$

$k = 8$  or  $11$

$n = 23$  or  $52$

$$k = 3 \quad n = 3$$



**Normalized:**  $\pm E$  is not its maximum or minimum value

$$1 \leq M < 2$$

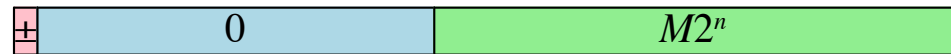


$$\pm E = e + 1 - 2^{k-1}$$

$$M = 1 + f/2^n$$

**Denormalized:**  $\pm E$  is its minimum value (which is negative)

$$0 \leq M < 1$$



$$\pm E = 2 - 2^{k-1}$$

$$M = f/2^n$$

**Infinity:**  $\pm E$  is its maximum value



$\pm =$

**Not-a-Number:**  $\pm E$  is its maximum value

(many representations!)



$\pm E =$

$M =$

$$\pm M \times 2^{\pm E}$$

1 bit for  $\pm$

$k$  bits for  $\pm E$

$n$  bits for  $M$

$k = 8$  or  $11$

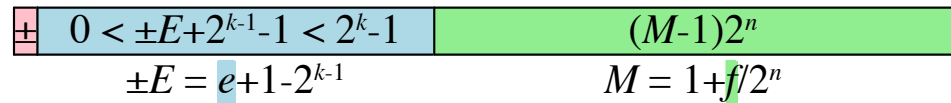
$n = 23$  or  $52$

$$k = 3 \quad n = 3$$



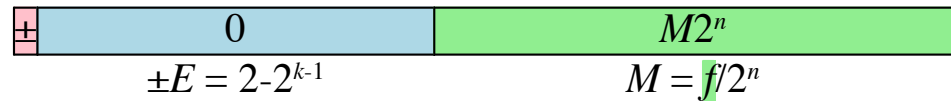
**Normalized:**  $\pm E$  is not its maximum or minimum value

$$1 \leq M < 2$$



**Denormalized:**  $\pm E$  is its minimum value (which is negative)

$$0 \leq M < 1$$



**Infinity:**  $\pm E$  is its maximum value



**Not-a-Number:**  $\pm E$  is its maximum value (many representations!)



$\pm =$

$\pm E =$

$M =$

$$\pm M \times 2^{\pm E}$$

1 bit for  $\pm$

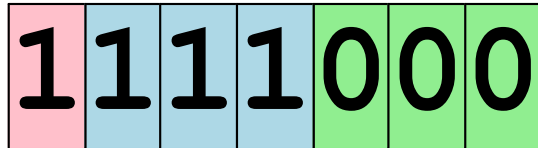
$k$  bits for  $\pm E$

$n$  bits for  $M$

$k = 8$  or  $11$

$n = 23$  or  $52$

$$k = 3 \quad n = 3$$



**Normalized:**  $\pm E$  is not its maximum or minimum value

$$1 \leq M < 2$$



$$0 < \pm E + 2^{k-1} - 1 < 2^k - 1$$

$$(M-1)2^n$$

$$\pm E = e + 1 - 2^{k-1}$$

$$M = 1 + f/2^n$$

**Denormalized:**  $\pm E$  is its minimum value (which is negative)

$$0 \leq M < 1$$



$$0$$

$$M2^n$$

$$\pm E = 2 - 2^{k-1}$$

$$M = f/2^n$$

**Infinity:**  $\pm E$  is its maximum value



$$2^k - 1$$

$$0$$

**Not-a-Number:**  $\pm E$  is its maximum value

(many representations!)



$$2^k - 1$$

$$\text{non-0}$$

$\pm =$

$\pm E =$

$M =$

$$\pm M \times 2^{\pm E}$$

1 bit for  $\pm$

$k$  bits for  $\pm E$

$n$  bits for  $M$

$k = 8$  or  $11$

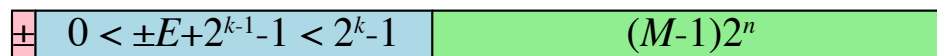
$n = 23$  or  $52$

$$k = 3 \quad n = 3$$



**Normalized:**  $\pm E$  is not its maximum or minimum value

$$1 \leq M < 2$$



**Denormalized:**  $\pm E$  is its minimum value (which is negative)

$$0 \leq M < 1$$



**Infinity:**  $\pm E$  is its maximum value



**Not-a-Number:**  $\pm E$  is its maximum value

(many representations!)



$\pm =$

$\pm E =$

$M =$

$$\pm M \times 2^{\pm E}$$

1 bit for  $\pm$

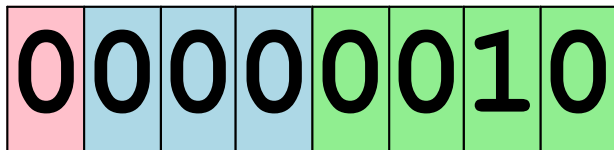
$k$  bits for  $\pm E$

$n$  bits for  $M$

$k = 8$  or  $11$

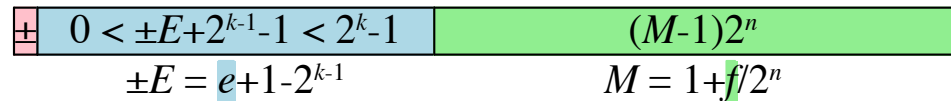
$n = 23$  or  $52$

$$k = 3 \quad n = 4$$



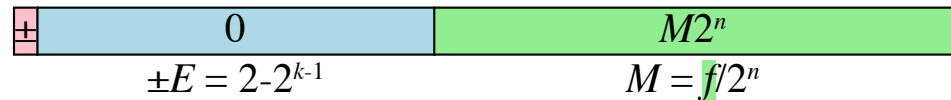
**Normalized:**  $\pm E$  is not its maximum or minimum value

$$1 \leq M < 2$$



**Denormalized:**  $\pm E$  is its minimum value (which is negative)

$$0 \leq M < 1$$



**Infinity:**  $\pm E$  is its maximum value



**Not-a-Number:**  $\pm E$  is its maximum value

(many representations!)



$\pm =$

$\pm E =$

$M =$