- $(-307.1875)_{10}$
- 000100110011.0011
- $1.001100110011*2^{8}$
- S=1; E=8,E'=8+127=135=87H=10000111
- M=001100110011

- $1.00111010011001*2^{10}$
- S=0; E=10; E'=10+1023=1033
- E'=E+1023=10+1023=1033D=409H=10000001001b
- M=0011101011001

| 0 | 10000001001 | 00111010110010000000.....00 |
|---|---|---|

- 4093AC8000000000

# Sign and exponent digits

- In a 32-bit number, suppose we allocate 24 bits to represent a fractional mantissa.
- Assume that the mantissa is represented in sign and magnitude format, and we have allocated one bit to represent the sign.
- We allocate 7 bits to represent the exponent, and assume that the exponent is represented as a 2's complement integer.
- There are no bits allocated to represent the base, we assume that the base is implied for now, that is the base is 2.
- Since a 7-bit 2's complement number can represent values in the range -64 to 63, the range of numbers that can be represented is:

$$0.0000001 \times 2^{-64} <= |x| <= 0.9999999 \times 2^{63}$$

- In decimal representation this range is:

$$0.5421 \times 10^{-20} <= |x| <= 9.2237 \times 10^{18}$$

adjusting the exponent. As computations proceed, a number that does not fall in the representable range of normal numbers might be generated.

- In single precision, it requires an exponent less than -126 (underflow) or greater than +127 (overflow). Both are exceptions that need to be considered.

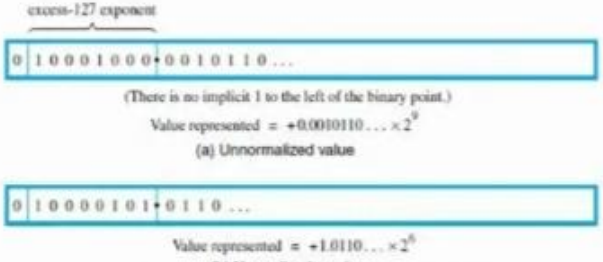


**Figure 9.27** Floating-point normalization in IEEE single-precision format.

**SPECIAL VALUES**

- The end values 0 and 255 of the excess-127 exponent E' are used to repr
- When E'=0 and the mantissa fraction m is zero, the value exact 0 is repre
- When E'=255 and M=0, the value ∞ is represented, where ∞ is the number by zero.
- when E'=0 and M!=-, denormal numbers are represented. Their value is $(-1)^S \times 0.M \times 2^{-126}$
- When E'=255 and M!=0, the value represented is called not a number(NaN). A NaN is the result of performing an invalied operation such as 0/0 or $\sqrt{0}$.

**ARITHMETIC OPERATIONS ON FLOATING-POINT NUMBERS**

**Multiply Rule**

1) Add the exponents & subtract 127.
2) Multiply the mantissas & determine sign of the result.
3) Normalize the resulting value if necessary.

**Divide Rule**

1) Subtract the exponents & add 127.
2) Divide the mantissas & determine sign of the result.