1 Floating points

1.1 Floating points

$2^n - 1$ number of digits represented in binary $2^{n-1} - 1$ for signed digits represented in binary

1.2 Fractional Numbers in Binary

binary fractions use negative powers of 2 $2^{-1} = 0.5$

1.3 Fixed-point Numbers in Binary

1.4 Floating-point Numbers/scientific notations

1234×10^4 $M \times 2^E$ Mantissa (significant) Exponent(biased) IEEE-754 standard for float 64 representation 1 bit overall sign 52 bits for Mantissa, 8 for exponent float 64 =¿ 2^{1024}