# Floating-Point Formats

Version 1.0 © 2013-06-13 Adam Sawicki, adam\_\_DELETE\_\_@asawicki.info, <http://asawicki.info>
Source: <http://en.wikipedia.org>

## General Parameters

|  |  |  |  |
| --- | --- | --- | --- |
| **Name** | Half | Single | Double |
| **Bytes** | 2 | 4 | 8 |
| **Bits = sign + exponent + significand** | 16 = 1 + 5 + 10 | 32 = 1 + 8 + 23 | 64 = 1 + 11 + 52 |
| **Exponent bias, range** | 15, −14 … 15 | 127, −126 … 127 | 1023, −1022 … 1023 |
| **Significant decimal digits** | ≈ 3.311 | ≈ 7.225 (6 … 9) | ≈ 15.955 (15 … 17) |
| **Minimum positive subnormal** | 2−24 ≈ 5.96 × 10−8 | 2−149 ≈ 1.4 × 10−45 | 2−1022-52 ≈ 5 × 10−324 |
| **Minimum positive normal** | 2−14 ≈ 6.10 × 10−5 | 2−126 ≈ 1.18 × 10−38 | 2−1022 ≈ 2.2250738585072014 × 10−308 |
| **Next value after 1** | 1 + 2−10 = 1.0009765625 | 1 + 2−23 ≈ 1.0000001 | 1 + 2−52 ≈ 1.0000000000000002 |
| **Integers represented exactly** | 0 … 211 = 2048 | 0 … 224 = 16,777,216 | 0 … 253 = 9,007,199,254,740,992 |
| **Maximum** | (2−2−10) × 215 = 65504 | (2−2−23) × 2127 ≈ 3.4 × 1038 | (1 + (1 − 2−52)) × 21023 ≈ 1.7976931348623157 × 10308 |

## Example Values (hexadecimal)

|  |  |  |  |
| --- | --- | --- | --- |
| **0.0** | 0000 | 00000000 | 00000000 00000000 |
| **1.0** | 3C00 | 3F800000 | 3FF00000 00000000 |
| **-1.0** | BC00 | BF800000 | BFF00000 00000000 |
| **0.5** | 3800 | 3F000000 | 3FE00000 00000000 |
| **2.0** | 4000 | 40000000 | 40000000 00000000 |
| **Example NaN** | FFFF | FFFFFFFF | FFFFFFFF FFFFFFFF |

## Special Values

|  |  |  |
| --- | --- | --- |
| **Exponent (binary)** | **Significand (binary)** | **Meaning** |
| 000…0 | 000…0 | −0 / +0 |
| 000…0 | Non-zero | Subnormal |
| 111…1 | 000…0 | −Inf / +Inf (Infinity) |
| 111…1 | Non-zero | NaN (Not a Number) |
| Other value | Any value | Normalized value |