Exercici I **Norma IEEE-754**

1. Consulta la documentació sobre l'l'aritmètica de punt flotant i l'aritmètica de MATLAB® disponible al campus virtual, enllaç documentació.
2. Escriu una explicació (màxim **3 fulls** Din-A4) sobre l'aritmètica de punt flotant, la norma IEEE-754 i l'aritmètica de MATLAB® .
3. Cita **totes** les fonts bibliogràfiques consultades.

Norma IEEE-754

El año 1985, Institute for Electrical and Electronic Engineers (IEEE) va publicar el informe Binary Floating Point Arithmetic Standard 754 − 1985, en el que se especifican normas para representar numeros en punto flotante con precisión simple doble y formatos de precision extendidos. El año 2008 fue publicada el IEE Std 754-2008.

La norma IEEE-754 define:

- Formatos aritméticos: conjuntos de datos de punto flotante binarios y decimales, que consisten en números finitos, includios los ceros con singo, y los números desnormalizados, infinitos y valores especiales "no numéricos" (Nan).
- Formatos de intercambio: codificaciones (cadenas de bits) que se pueden utilizar para intercambiar datos de punto flotante de forma eficiente y compacta.
- Reglas de redondeo: propiedades que deben satisfacer al redondear los numeros.
- Operaciones aritmeticas.
- Indicaciones de excepciones: overflow, underflow, divisiones por cero...

Formatos aritméticos

Single and Double Precision:

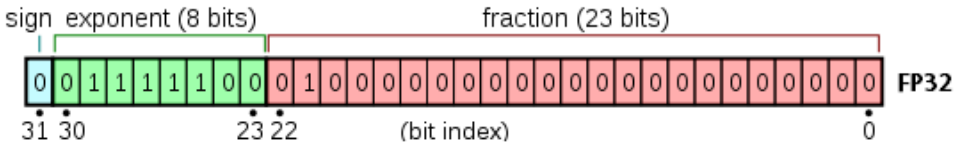
Single Precision

la represenatción en punto flotante de simple precisión ocupa 1 word(32 digitos), asignando: 1 bit para el signo del numero real, 8 bits para el exponente y 23 bits para la mantisa.

Un nombre real x en simple precisión se reprensenta de la siguiente manera:

$$fl(x) = (-1)^s \times M \times 2^{c-127}.$$

$$M = 1.m_{22}m_{21}...m_1m_0 = 1 + m_{22}\cdot 2^{-1} + m_{22}\cdot 2^{-2} + ... + m_1\cdot 2^{-20} + m_2\cdot 2^{-21}$$



Donde **s** es el valor del del signe. **c** es el valor del exponente (con signo). $m_{22}m_{21}\cdots m_1m_0$ el valor de la mantissa.

Los valores de los exponentes con todo 0's o todo 1's, estan reservados.

Denormal number (Double Precision)

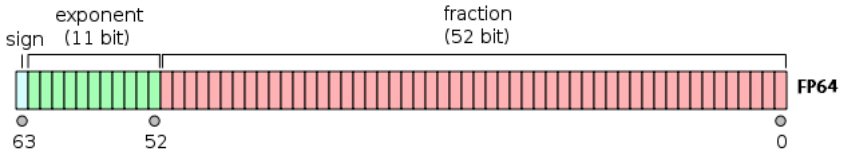
Double Precision

la representación en punto flotante con doble precision ocupa 2 words(64 digitos), asignando: 1 bit para el signo del numero real, 11 bits para el exponente y 52 bits para la mantisa.

Un nombre real x en simple precisión se reprensenta de la siguiente manera:

$$fl(x) = (-1)^s \times M \times 2^{c-1023}$$

$$M = 1.m_{51}m_{50}...m_1m_0 = 1 + m_{51}\cdot 2^{-1} + m_{50}\cdot 2^{-2} + ... + m_1\cdot 2^{-59} + m_2\cdot 2^{-50}$$



Donde **s** es el valor del del signe. **c** es el valor del exponente (con signo). $m_{51}m_{50}\cdots m_1m_0$ el valor de la mantisa.

Los valores de los exponentes con todo 0's o todo 1's, estan reservados.

Denormal number (Double Precision)

Si el exponente es 0 y la mantisa diferente de 0.

Inf

Si el exponente es todos 1 y la mantisa es 0 representa a los numeros infinitos. En caso de que el signo sea 0 respresenta el +infinito.En caso de que el signo sea 1 representa -infinito.

Nan(Not-A-Number)

Si el exponente es todos 1 y la mantisa es diferente de 0 representa Nan(Not-A-Number). Esto se utiliza para expresar un resultado imposible de calcular (raíces negativas, division por cero)...

Zero

Si el exponente es todos 0 y la mantisa 0 representa al 0. En caso de que el signo sea 0 representa el +0. En casi de que el signo sea 1 representa -0.

Denormal number (

Si el exponente es 0 y la mantisa diferente de 0.

Signo (s)	*Exponente (exp)*	*Mantisa (m)*	© carlospes.com *Valor en base 10*
0 ó 1	0 < **exp** < 2047	Indiferente	$(-1)^{s} \cdot 1,m \cdot 2^{exp-1023}$
0	**exp** = 2047	**m** = 0	$+\infty$
1	**exp** = 2047	**m** = 0	$-\infty$
0 ó 1	**exp** = 2047	**m** ≠ 0	NaN
0 ó 1	**exp** = 0	**m** = 0	0
0 ó 1	**exp** = 0	**m** ≠ 0	$(-1)^{s} \cdot 0,m \cdot 2^{-1022}$

Signo (s)	*Exponente (exp)*	*Mantisa (m)*
0 ó 1	0 < **exp** < 255	Indiferente
0	**exp** = 255	**m** = 0
1	**exp** = 255	**m** = 0
0 ó 1	**exp** = 255	**m** ≠ 0
0 ó 1	**exp** = 0	**m** = 0
0 ó 1	**exp** = 0	**m** ≠ 0

Biografia:

https://es.wikipedia.org/wiki/IEEE_coma_flotante

https://blogs.mathworks.com/cleve/2014/07/07/floating-point-numbers/

https://blogs.mathworks.com/cleve/2014/07/21/floating-point-denormals-insignificant-but-controversial-2/