



ISTANBUL **TECHNICAL** UNIVERSITY

BLG202E
Recitation I

19.02.2019

Outline



- MATLAB Tutorial
- Floating Point Systems
 - Decimal to Binary
 - The IEEE Standard
- Taylor's Theorem

Floating Point Systems

Example #1

Convert $(11.1875)_{10}$ to binary representation $(?.?)_2$

$$\begin{aligned}(11)_{10} &= 2^3 + 3 \\ &= 2^3 + 2^1 + 1 \\ &= 2^3 + 2^1 + 2^0 \\ &= 1 \times 2^3 + 0 \times 2^2 + 1 \times 2^1 + 1 \times 2^0 \\ &= (1011)_2\end{aligned}$$

$$\begin{aligned}2^0 &= 1 \\ 2^1 &= 2 \\ 2^2 &= 4 \\ 2^3 &= 8 \\ &\cdot \\ &\cdot \\ &\cdot\end{aligned}$$

Floating Point Systems

Example #1

Convert $(11.1875)_{10}$ to binary representation $(?.?)_2$

$$\begin{aligned}(0.1875)_{10} &= 2^{-3} + 0.0625 \\ &= 2^{-3} + 2^{-4} \\ &= 0 \times 2^{-1} + 0 \times 2^{-2} + 1 \times 2^{-3} + 1 \times 2^{-4} \\ &= (.0011)_2\end{aligned}$$

$$\begin{aligned}2^{-1} &= 0.5 \\ 2^{-2} &= 0.25 \\ 2^{-3} &= 0.125 \\ 2^{-4} &= 0.0625 \\ &\vdots\end{aligned}$$

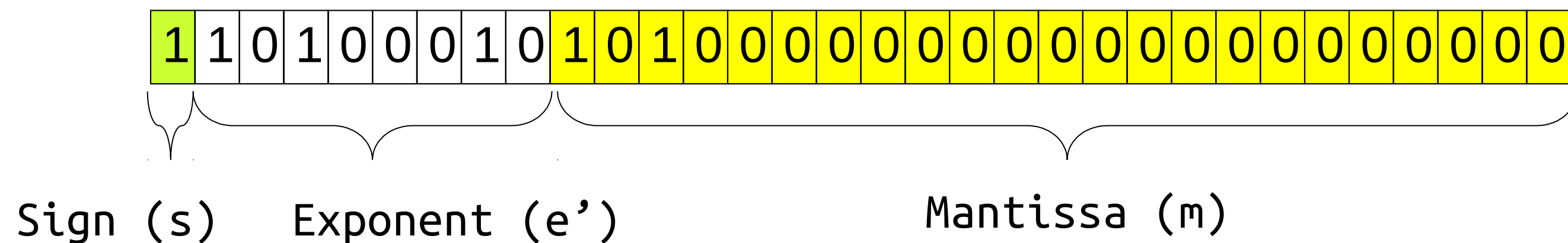
$$(11.1875)_{10} = (1011.0011)_2$$

Diagram illustrating the IEEE 754 single-precision floating-point format (32 bits):

- Sign (s):** 1 bit (highlighted in green).
- Exponent (e'): 8 bits (highlighted in light blue).**
- Mantissa (m): 23 bits (highlighted in yellow).**

Floating Point Systems

Example #2



Normalize the exponent by subtracting 127

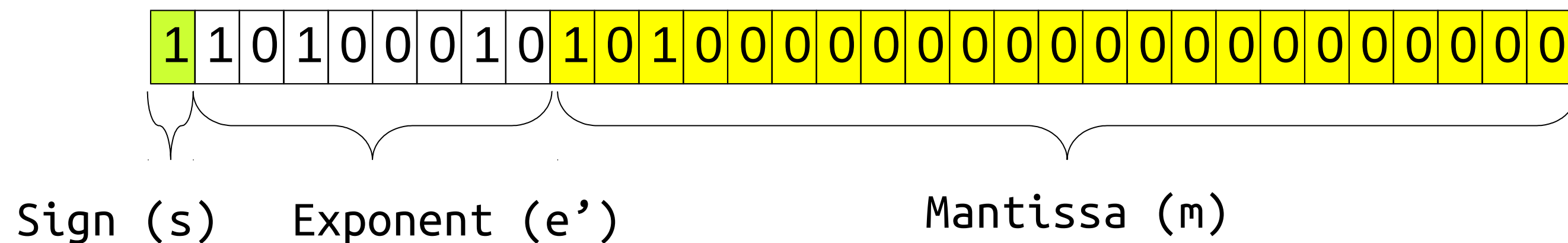
$$\text{Value} = (-1)^s \times 2^{(e' - 127)} \times (1.m)$$

Diagram illustrating the IEEE 754 single-precision floating-point format (32 bits):

- Sign (s):** 1 bit (highlighted in green).
- Exponent (e'):** 8 bits (highlighted in light yellow).
- Mantissa (m):** 23 bits (highlighted in yellow).

The bit sequence shown is: 1 1 0 1 0 0 0 1 0 1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.

18 ISTANBUL TECHNICAL UNIVERSITY



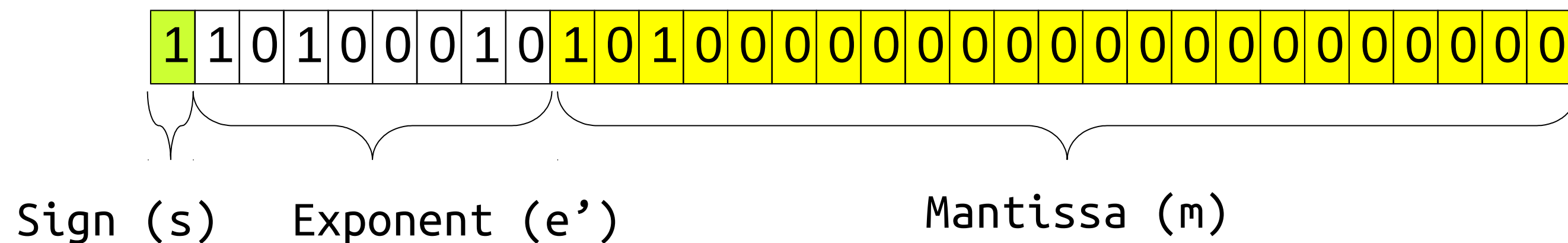
$$\begin{aligned}\text{Value} &= (-1)^s \times (1.m)_2 \times 2^{e'-127} \\ &= (-1)^1 \times (1.10100000)_2 \times 2^{(10100010)_2 - 127}\end{aligned}$$

Diagram illustrating the IEEE 754 single-precision floating-point format (32 bits):

- Sign (s):** 1 bit (highlighted in green).
- Exponent (e'):** 8 bits (highlighted in light blue).
- Mantissa (m):** 23 bits (highlighted in yellow).

The bit sequence shown is: 1 1 0 1 0 0 0 1 0 1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.

$$\begin{aligned}\text{Value} &= (-1)^s \times (1.m)_2 \times 2^{e'-127} \\ &= (-1)^1 \times (1.10100000)_2 \times 2^{(10100010)_2 - 127} \\ &= (-1) \times (1.625) \times 2^{162-127}\end{aligned}$$



$$\begin{aligned}\text{Value} &= (-1)^s \times (1.m)_2 \times 2^{e'-127} \\ &= (-1)^1 \times (1.10100000)_2 \times 2^{(10100010)_2 - 127} \\ &= (-1) \times (1.625) \times 2^{162-127} \\ &= (-1) \times (1.625) \times 2^{35} = -5.5834 \times 10^{10}\end{aligned}$$

Taylor's Theorem

Importance of Taylor's Theorem: This theorem can be used to approximate functions like \sin , \exp and \log by polynomials (which are easy to compute) and provides an estimate of the error involved in the approximation.

Taylor's Theorem. Let f be an $(n + 1)$ times differentiable function on an open interval containing the points a and x . Then

$$f(x) = f(a) + f'(a)(x - a) + \frac{f''(a)}{2!}(x - a)^2 + \dots + \frac{f^{(n)}(a)}{n!}(x - a)^n + R_n(x)$$

where

$$R_n(x) = \frac{f^{(n+1)}(c)}{(n + 1)!}(x - a)^{n+1}$$

for some number c between a and x .

Error

Taylor's Theorem



The remainder term is given by

$$R_n(x) = f(x) - f(a) - f'(a)(x-a) - \frac{f''(a)}{2!}(x-a)^2 - \dots - \frac{f^{(n)}(a)}{n!}(x-a)^n.$$

Do not confuse $f(x)$ with the approximation. It is the original function that we want to find its estimation !

Taylor's Theorem



The remainder term is given by

$$R_n(x) = f(x) - f(a) - f'(a)(x-a) - \frac{f''(a)}{2!}(x-a)^2 - \dots - \frac{f^{(n)}(a)}{n!}(x-a)^n.$$

Fix x and a . For t between x and a set

$$F(t) = f(x) - f(t) - f'(t)(x-t) - \frac{f''(t)}{2!}(x-t)^2 - \dots - \frac{f^{(n)}(t)}{n!}(x-t)^n,$$

so that $F(a) = R_n(x)$. Then

$F(t)$ can be seen as a general function and $R_n(x)$ is a sample from this function

Taylor's Theorem



The remainder term is given by

$$R_n(x) = f(x) - f(a) - f'(a)(x-a) - \frac{f''(a)}{2!}(x-a)^2 - \dots - \frac{f^{(n)}(a)}{n!}(x-a)^n.$$

Fix x and a . For t between x and a set

$$F(t) = f(x) - f(t) - f'(t)(x-t) - \frac{f''(t)}{2!}(x-t)^2 - \dots - \frac{f^{(n)}(t)}{n!}(x-t)^n,$$

so that $F(a) = R_n(x)$. Then

The diagram shows the differentiation of the function $F(t)$. Three red arrows point from the terms in the definition of $F(t)$ to the corresponding terms in the derivative $F'(t)$. The first arrow points from $-f(t)$ to $-f'(t)$. The second arrow points from $-f'(t)(x-t)$ to $+f'(t)$. The third arrow points from $-\frac{f''(t)}{2!}(x-t)^2$ to $+\frac{f''(t)}{2!}(x-t)$. A bracket groups the remaining terms in $F(t)$ and the corresponding terms in $F'(t)$.

$$\begin{aligned} F'(t) &= -f'(t) - f''(t)(x-t) + f'(t) - \frac{f'''(t)}{2!}(x-t)^2 + f''(t)(x-t) \\ &\quad - \dots - \frac{f^{(n+1)}(t)}{n!}(x-t)^n + \frac{f^{(n)}(t)}{(n-1)!}(x-t)^{n-1} \\ &= -\frac{f^{(n+1)}(t)}{n!}(x-t)^n. \end{aligned}$$

Taylor's Theorem

The remainder term is given by

$$R_n(x) = f(x) - f(a) - f'(a)(x-a) - \frac{f''(a)}{2!}(x-a)^2 - \dots - \frac{f^{(n)}(a)}{n!}(x-a)^n.$$

Fix x and a . For t between x and a set

$$F(t) = f(x) - f(t) - f'(t)(x-t) - \frac{f''(t)}{2!}(x-t)^2 - \dots - \frac{f^{(n)}(t)}{n!}(x-t)^n,$$

so that $F(a) = R_n(x)$. Then

$$\begin{aligned} F'(t) &= -f'(t) - f''(t)(x-t) + f'(t) - \frac{f'''(t)}{2!}(x-t)^2 + f''(t)(x-t) \\ &\quad - \dots - \frac{f^{(n+1)}(t)}{n!}(x-t)^n + \frac{f^{(n)}(t)}{(n-1)!}(x-t)^{n-1} \\ &= -\frac{f^{(n+1)}(t)}{n!}(x-t)^n. \end{aligned}$$

Now defining

$$G(t) = F(t) - \left(\frac{x-t}{x-a} \right)^{n+1} F(a),$$

Taylor's Theorem



Rolle's Theorem: Let f be continuous on a closed interval $[a, b]$ and differentiable on the open interval (a, b) . If $f(a) = f(b)$, then there is at least one point c in (a, b) where $f'(c) = 0$.

Taylor's Theorem



Rolle's Theorem: Let f be continuous on a closed interval $[a, b]$ and differentiable on the open interval (a, b) . If $f(a) = f(b)$, then there is at least one point c in (a, b) where $f'(c) = 0$.

Now defining

$$G(t) = F(t) - \left(\frac{x - t}{x - a} \right)^{n+1} F(a),$$

Taylor's Theorem



Rolle's Theorem: Let f be continuous on a closed interval $[a, b]$ and differentiable on the open interval (a, b) . If $f(a) = f(b)$, then there is at least one point c in (a, b) where $f'(c) = 0$.

Now defining

$$G(t) = F(t) - \left(\frac{x-t}{x-a} \right)^{n+1} F(a),$$

we have $G(a) = 0$ and $G(x) = F(x) = 0$. Applying Rolle's theorem to the function G shows that there is a c between a and x with $G'(c) = 0$. Now

$$\begin{aligned} 0 = G'(c) &= F'(c) + (n+1) \frac{(x-c)^n}{(x-a)^{n+1}} F(a) \\ &= -\frac{f^{(n+1)}(c)}{n!} (x-c)^n + (n+1) \frac{(x-c)^n}{(x-a)^{n+1}} F(a). \end{aligned}$$

Taylor's Theorem



Rolle's Theorem: Let f be continuous on a closed interval $[a, b]$ and differentiable on the open interval (a, b) . If $f(a) = f(b)$, then there is at least one point c in (a, b) where $f'(c) = 0$.

Now defining

$$G(t) = F(t) - \left(\frac{x-t}{x-a} \right)^{n+1} F(a),$$

we have $G(a) = 0$ and $G(x) = F(x) = 0$. Applying Rolle's theorem to the function G shows that there is a c between a and x with $G'(c) = 0$. Now

$$\begin{aligned} 0 &= G'(c) = F'(c) + (n+1) \frac{(x-c)^n}{(x-a)^{n+1}} F(a) \\ &= -\frac{f^{(n+1)}(c)}{n!} (x-c)^n + (n+1) \frac{(x-c)^n}{(x-a)^{n+1}} F(a). \end{aligned}$$

But $F(a) = R_n(x)$ and rearranging the last equation gives

$$\underline{R_n(x) = F(a) = \frac{f^{(n+1)}(c)}{(n+1)!} (x-a)^{n+1}.$$

→ Bounding Taylor Error

→ <https://www.khanacademy.org/math/ap-calculus-bc/bc-series-new/bc-10-12/v/error-or-remainder-of-a-taylor-polynomial-approximation>

→ Taylor & Maclaurin Series Intuition

→ <https://www.khanacademy.org/math/ap-calculus-bc/bc-series-new/bc-10-11/v/maclaurin-and-taylor-series-intuition>