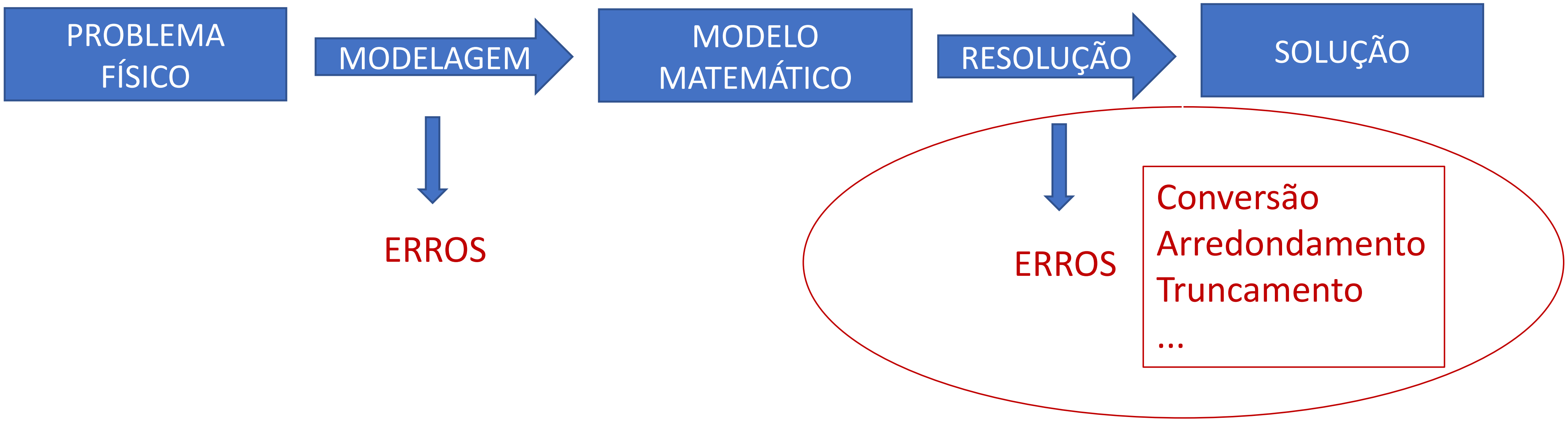
CÁLCULO NUMÉRICO

MAT 271 – CÁLCULO NUMÉRICO –2023-I

Professor Amarísio Araújo – DMA/UFV

CÁLCULO NUMÉRICO

Conjunto de métodos utilizados para resolver problemas matemáticos de forma exata ou aproximada através de procedimentos computacionais.



ERROS NA RESOLUÇÃO

- São fortemente influenciados pela precisão.

- Têm relação com a representação numérica (nós manipulamos os números na base decimal, os computadores, na base binária).

- Pode ocorrer: máquinas com precisão diferente, ainda que com o mesmo software.

TIPOS DE ERROS DE PRECISÃO

- Arredondamento: relacionados com as limitações na forma de representar um número.

- Truncamento: relacionado a um procedimento finito para um processo matemático infinito.

- De máquina: relacionado à capacidade de memória da máquina.

ERROS NA REPRESENTAÇÃO

- Na representação de um número, pode ser impossível escrever todos os dígitos que o representam de forma exata. No caso de uma máquina, que tem memória finita, esta representação depende da base, e é preciso definir a quantidade máxima de espaço para escrever os dígitos correspondentes à representação desse número.

  Por exemplo: na base decimal, o número racional $\frac{29189}{33300}$ corresponde à uma dízima periódica com período 654 (sequência de dígitos que se repete infinitamente) e antiperíodo 87 :

$$\frac{29189}{33300} = 0{,}87654654654\ldots$$

Temos, portanto uma limitação física para escrever toda a dízima (espaço e tempo). Esta é uma limitação humana e também de máquina (memória finita).

É preciso representar só uma quantidade finita da sequência de dígitos que compõem a dízima, o que pode ser feito “truncando” o número, apresentando-o, assim, de forma aproximada.

O fato de a memória de uma máquina ser finita traz também a necessidade de “arredondamento” para um número composto por uma sequência finita e “muito grande” de dígitos ou mesmo a substituição do número por um “número de máquina”.

ERROS DE ARREDONDAMENTO

Arredondar um número na casa (dígito) d_i significa desconsiderar as casas $d_{i+j}, j = 1,2,3, \ldots$, de tal forma que:

- d_i seja a última casa se $d_{i+1} < 5$;

- $d_i + 1$ seja a última casa se $d_{i+1} \geq 5$.

13,935682

5 casas

4 casas

13,93568

13,9357

ERROS DE TRUNCAMENTO

São erros que ocorrem em processos que, em princípio, são infinitos, como, por exemplo, nas séries (séries numéricas, séries de funções).

EXEMPLO:

$$\frac{1}{3} = 0{,}33333\ldots = \frac{3}{10} + \frac{3}{100} + \frac{3}{1000} \ldots + \frac{3}{10^k} + \cdots = 3\sum_{n=1}^{\infty}\left(\frac{1}{10}\right)^n$$

O dígito 3 é infinitamente repetido na representação do número racional acima. É preciso, então, nos restringir a um número finito destas repetições, ou seja, considerar um número finito de parcelas da série acima.

ERROS DE TRUNCAMENTO

OUTRO EXEMPLO:

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} \dots + \frac{x^k}{k!} + \cdots = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

EM PARTICULAR:

$$e = 2 + \frac{1}{2!} + \frac{1}{3!} \dots + \frac{1}{k!} + \cdots = \sum_{n=0}^{\infty} \frac{1}{n!}$$

Portanto, para obter uma representação finita do número irracional (e transcendente) e, devemos usar um número finito de parcelas da série acima (truncar a série).

ERROS DE MÁQUINA

Os computadores têm memória finita. Portanto, eles só podem representar uma quantidade finita de números. A representação desses números, na base binária, segue um sistema chamado de “sistema de ponto flutuante” que preestabelece quais e quantos números podem ser representados na máquina.

NÚMEROS DE MÁQUINA

REPRESENTAÇÃO NUMÉRICA

- O nosso sistema convencional de representação de um número é o sistema decimal, ou seja, o sistema de base 10 (formada pelos dígitos de 0 a 9).

- Os computadores modernos usam o sistema binário, isto é, o sistema de base 2 (formada pelos dígitos 0 e 1).

BASE DECIMAL

Todo número pode ser entendido como uma combinação de potências inteiras de 10, usando como coeficientes dígitos da base {0,1,2,3,4,5,6,7,8,9}.

$$3657 \quad = 7 \times 10^0 + 5 \times 10^1 + 6 \times 10^2 + 3 \times 10^3$$

$$3657{,}984 \ = 3 \times 10^3 + 6 \times 10^2 + 5 \times 10^1 + 7 \times 10^0 + 9 \times 10^{-1} + 8 \times 10^{-2} + 4 \times 10^{-3}$$

BASE BINÁRIA

Todo número pode ser entendido como uma combinação de potências inteiras de 2, usando como coeficientes dígitos da base {0,1}

$101 = 1 \times 2^0 + 0 \times 2^1 + 1 \times 2^2$ ⟶ 5 na base decimal

$10001 = 1 \times 2^0 + 0 \times 2^1 + 0 \times 2^2 + 0 \times 2^3 + 1 \times 2^4$ ⟶ 17 na base decimal

CONVERSÃO DA BASE DECIMAL PARA A BINÁRIA

Como converter a representação de um número na base 10 para a sua representação na base 2? Vejamos com os exemplos ilustrados a seguir:

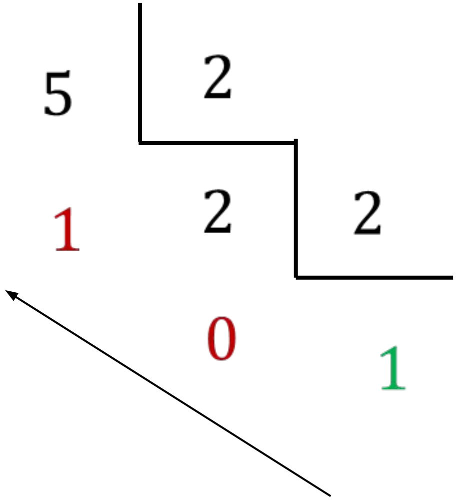
$(5)_{10} = (?)_2$

Divisões sucessivas por 2:

$5/2 = 2\ resto\ 1$

$2/2 = 1\ resto\ 0$

(não é possível mais divisão inteira)



$\boxed{\text{Logo: } (5)_{10} = (101)_2}$

CONVERSÃO DA BASE DECIMAL PARA A BINÁRIA

$(17)_{10} = (?)_2$

Divisões sucessivas por 2:

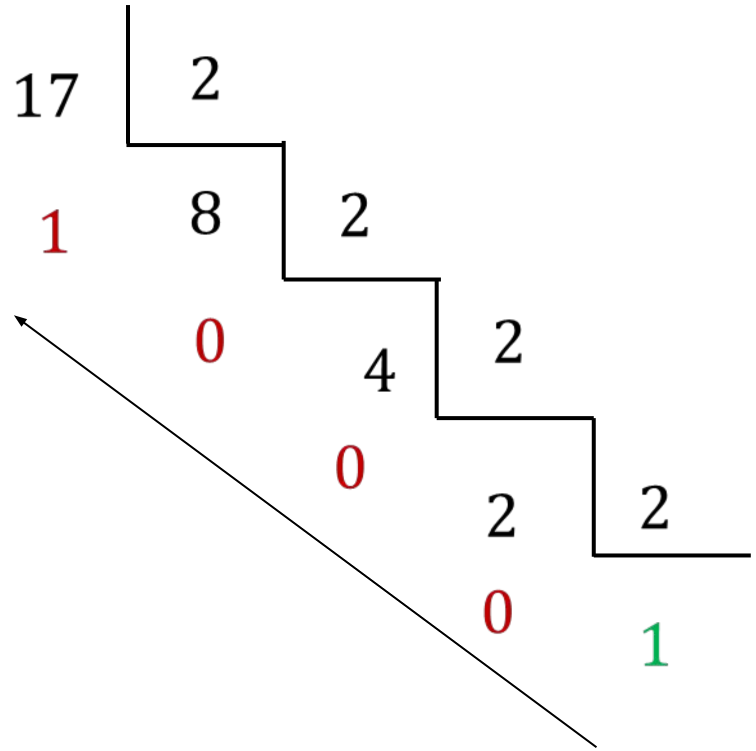
$17/2 = 8\ resto\ 1$

$8/2 = 4\ resto\ 0$

$4/2 = 2\ resto\ 0$

$2/2 = 1\ resto\ 0$

(não é possível mais divisão inteira)



Logo: $(17)_{10} = (10001)_2$

CONVERSÃO DA BASE BINÁRIA PARA A DECIMAL

$(5{,}25)_{10} = (?)_2$

A parte inteira, 5, já sabemos converter. $\boxed{(5)_{10} = (101)_2}$

Como converter a parte decimal 0,25?

Multiplicações sucessivas por 2:

$0{,}25 \times 2 = 0{,}5$

$0{,}5 \times 2 = 1{,}0$

(não tem mais parte decimal, é inteiro)

$$\begin{array}{r} 0{,}25 \\ \times\ 2 \\ \hline 0{,}50 \\ \times\ 2 \\ \hline 1{,}00 \end{array}$$

$\boxed{\text{Logo}: (5{,}25)_{10} = (101{,}01)_2}$

CONVERSÃO DA BASE BINÁRIA PARA A DECIMAL

- $(100)_2 = 1 \times 2^2 + 0 \times 2^1 + 0 \times 2^0 = (4)_{10}$
- $(100{,}1)_2 = 1 \times 2^2 + 0 \times 2^1 + 0 \times 2^0 + 1 \times 2^{-1} = (4{,}5)_{10}$
- $(101)_2 = 1 \times 2^2 + 0 \times 2^1 + 1 \times 2^0 = 4 + 0 + 1 = (5)_{10}$

REPRESENTAÇÃO NUMÉRICA EM MÁQUINA

- Computadores usam o Sistema de Ponto Flutuante Normalizado (SPFN) para representar um número:

$$\pm 0, c_1 c_2 c_3 \dots c_n \times b^E$$

Onde:

- c_n são dígitos entre 0 e $b-1$ (constituem a mantissa);
- c_1 é diferente de 0;
- b é um número natural (base);
- E é um número inteiro (expoente).

REPRESENTAÇÃO NUMÉRICA EM MÁQUINA

- Em função da memória finita dos computadores, o SPFN estabelece parâmetros bem definidos em seu projeto, que são os seguintes:

  - O valor da base: b;
  - O número n de caracteres da mantissa;
  - O expoente mínimo E_1 e o expoente máximo E_2
  - $E_1 < 0$ e $E_2 > 0$.

$$\boxed{SPFN(b, n, E_1, E_2)}$$

IEEE: INSTITUTE OF ELECTRICAL ENGINEERING AND ELECTRONICS ENGINEERS

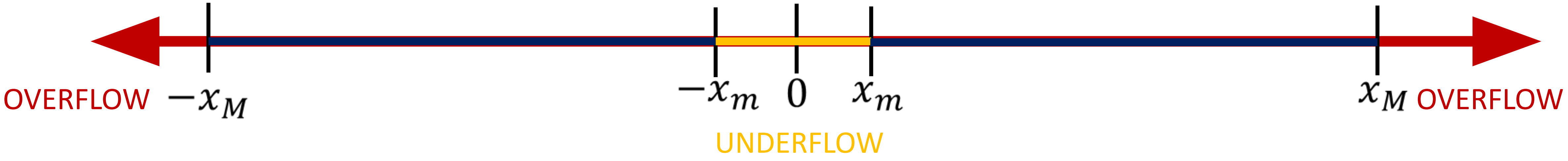
REPRESENTAÇÃO NUMÉRICA EM MÁQUINA

- Com os parâmetros definidos, $SPFN(b, n, E_1, E_2)$, temos:

  - O menor número positivo da máquina: $x_m = (0{,}10\ldots0) \times b^{E_1}$;

  - O maior número positivo da máquina: $x_M = (0, [b-1][b-1]\ldots[b-1]) \times b^{E_2}$;

  - Quantidade de números representáveis: $2 \times (b-1) \times b^{n-1} \times (E_2 - E_1 + 1) + 1$



EXEMPLO SIMPLES

Seja uma máquina com: $SPFN(2,3,-1,2)$, ou seja: $b = 2$, $n = 3$, $E_1 = -1$ e $E_2 = 2$.

O menor número positivo da máquina: $x_m = (0{,}100) \times 2^{-1}$;

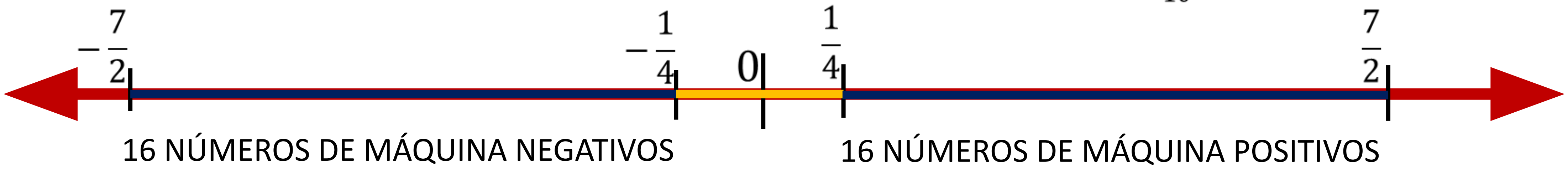
O maior número positivo da máquina: $x_M = (0{,}111) \times 2^{2}$;

Quantidade de números de máquina: $2 \times (2-1) \times 2^{3-1} \times (2-(-1)+1) + 1 = 33$

Convertendo para a base decimal, temos os correspondentes na reta dos reais:

$$x_m = (0{,}100) \times 2^{-1} = ((1 \times 2^{-1} + 0 \times 2^{-2} + 0 \times 2^{-3}) \times 2^{-1})_{10} = \left(\frac{1}{4}\right)_{10}$$

$$x_M = (0{,}111) \times 2^{2} = ((1 \times 2^{-1} + 1 \times 2^{-2} + 1 \times 2^{-3}) \times 2^{2})_{10} = \left(\frac{7}{2}\right)_{10}$$

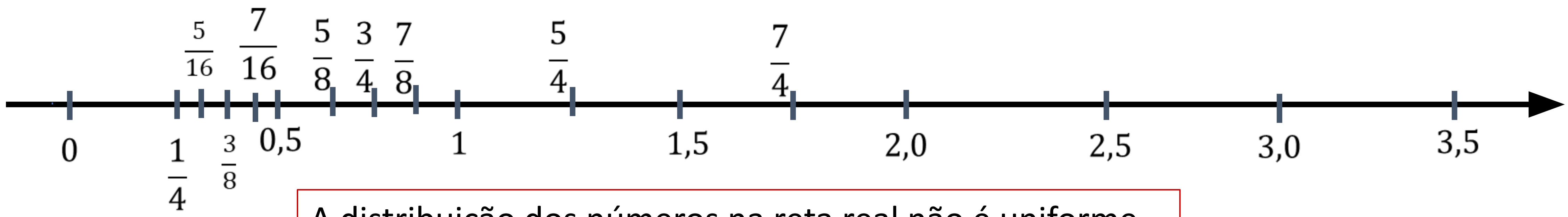


REPRESENTAÇÃO NUMÉRICA EM MÁQUINA

Vejamos quais são os números de máquina positivos ($b = 2, n = 3, E_1 = -1$ e $E_2 = 2$):

E	b^E	100	101	110	111
-1	1/2	1/4	5/16	3/8	7/16
0	1	1/2	5/8	3/4	7/8
1	2	1	5/4	3/2	7/4
2	4	2	5/2	3	7/2

NÚMEROS NA RETA



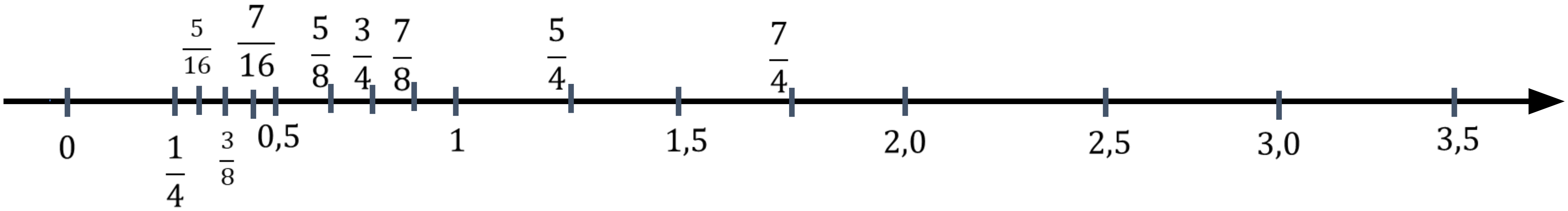
A distribuição dos números na reta real não é uniforme.

ERRO DE APROXIMAÇÃO EM MÁQUINA

Se $x := \frac{7}{8}$ e $y := \frac{7}{4}$, então $x, y \in SPFN(2{,}3, -1{,}2)$.

Mas $x + y := \frac{21}{8} = 2{,}625 \notin SPFN(2{,}3, -1{,}2)$.

Como $\frac{21}{8}$ está no intervalo $[\frac{1}{4}, \frac{7}{2}]$, ele será substituído pelo número de máquina mais próximo, que corresponde ao $\frac{5}{2} = 2.5$.



REPRESENTAÇÃO DO NÚMERO EM MÁQUINA (CONSIDERANDO BASE BINÁRIA)

UMA ILUSTRAÇÃO

0	1	1	1	0	0	0	0	0	0	0	0	0	1	0	0

SINAL MANTISSA SINAL EXPOENTE

SINAL:

0, SE O NÚMERO É POSITIVO

1, SE O NÚMERO É NEGATIVO

14

UM CASO INTERESSANTE:DÍZIMA PERIÓDICA BINÁRIA

Vamos converter o número 0,1 da base decimal para a base binária:

$$\begin{array}{r} 0{,}1 \\ \times\ 2 \\ \hline 0{,}2 \\ \times\ 2 \\ \hline 0{,}4 \\ \times\ 2 \\ \hline 0{,}8 \\ \times\ 2 \\ \hline 1{,}6 \end{array}$$

(A)

$$\begin{array}{r} 0{,}6 \\ \times\ 2 \\ \hline 1{,}2 \end{array}$$

(B)

$$\begin{array}{r} 0{,}2 \\ \times\ 2 \\ \hline 0{,}4 \\ \times\ 2 \\ \hline 0{,}8 \\ \times\ 2 \\ \hline 1{,}6 \end{array}$$

(C)

O processo se repete infinitamente

$$\boxed{(0{,}1)_{10} = (0{,}0001100110011\ldots)_2}$$

INEVITABILIDADE DOS ERROS

- Diante do que foi exposto, deve-se ter consciência de que os erros são, praticamente, inevitáveis e vão se acumulando durante o processo de solução numérica de um problema.

- É preciso, portanto, minimizar a influência de cada um dos tipos de erros, o que exige, quase sempre, um bom conhecimento analítico matemático.

- É essencial avaliar tais erros. Para isto, eles são quantificados através de medidas bem definidas, sendo as duas principais o ERRO ABSOLUTO e o ERRO RELATIVO.

AVALIANDO ERROS

Ao longo do nosso curso, usaremos duas medidas de avaliação do erro entre o valor exato $\bar{x}$ de uma solução de um dado problema e o valor aproximado $\tilde{x}$ desta solução, obtido por algum método numérico:

- Erro absoluto: $|\bar{x} - \tilde{x}|$;
- Erro relativo: $\frac{|\bar{x}-\tilde{x}|}{|\bar{x}|}$.
- O erro absoluto dá uma resultado com a unidade de medida da variável x, enquanto o erro relativo não depende da unidade de medida (é adimensional), dando uma medida percentual do erro.

Considere, por exemplo que a variável x do problema seja medida em metros. Suponha que a solução exata do problema seja $\bar{x} = 152\ m$ e que, ao aplicar um método numérico para resolver o problema de forma aproximada, seja obtida a solução aproximada $\tilde{x} = 157\ m$.

Erro absoluto: $|\bar{x} - \tilde{x}| = 5\ m$;

Erro relativo: $\frac{|\bar{x}-\tilde{x}|}{|\bar{x}|} = \frac{|152-157|}{|152|} = 0{,}032894 \approx 3{,}3\ \%$

PROBLEMAS A SEREM TRABALHADOS EM MAT 271

Ao longo do nosso curso, apresentaremos métodos numéricos para resolver os seguintes problemas:

- ❑ Equações não-lineares de uma única variável;
- ❑ Aproximações de funções por interpolação polinomial;
- ❑ Sistemas de equações lineares;
- ❑ Integração definida;
- ❑ Problemas de Valor Inicial