WS 25/26 Numerik 0 Loesungen

Igor Dimitrov Malte Herzog

2025-10-10

Table of contents

Preface

WS 25/26 Numerics 0 Solutions

1 Aufgabe 1

```
- Igor Dimitrov
- Malte Herzog
- Raphael Stumm
```

a)

$$\begin{aligned}
[0.5731E5]_8 &= (5 \times 8^{-1} + 7 \times 8^{-2} + 3 \times 8^{-3} + 1 \times 8^{-4}) \times 8^5 \\
&= 20480 + 3584 + 192 + 8 \\
&= 24264 \\
&= [0.24264E5]_{10} \in \mathbb{F}(10, 5, 1)
\end{aligned}$$

b)

$$\begin{aligned}
&0.3 \times 2 = 0.6 && m_1 = 0 \\
&0.6 \times 2 = 1.2 && m_2 = 1 \\
&0.2 \times 2 = 0.4 && m_3 = 0 \\
&0.4 \times 2 = 0.8 && m_4 = 0 \\
&0.8 \times 2 = 1.6 && m_5 = 1 \\
&0.6 \times 2 = 1.2 && m_6 = 1 && \text{(pattern repeats)} \\
&...
\end{aligned}$$

Somit:

$$\begin{aligned}
0.3 &= [0.0100110011001\,...]_2 \\
&= [0.100110011001\,...\,E-1]_2 \\
&= [0.10011001100E-1]_2 && \text{(abrunden r = 11)} \\
&\in \mathbb{F}(2, 11, 2)
\end{aligned}$$

$[0.10011001100E-1]_2 = 0.2998046875\ldots$

Es gilt:

$r=1$	$0.\underline{2}998046875 = 0.3 = 0.3$
$r=2$	$0.\underline{29}98046875 = 0.30 = 0.3$
$r=3$	$0.\underline{299}8046875 = 0.300 = 0.3$
$r=4$	$0.\underline{2998}046875 = 0.2998 \neq 0.3$

D.h der Fehler erst ab r = 4 bemerkbar

2 Aufgabe 3

b) Allgemein gilt:

$$\frac{|x \circledast y - x * y|}{|x * y|} \leq \frac{1}{2}\beta^{1-r}$$

Für $1 \oplus x$, und $\beta = 2$:

$$\begin{aligned}
\frac{|1 \oplus x - (1 + x)|}{|1 + x|} &= \frac{|1 - 1 - x|}{|1 + x} \\
&= \frac{|x|}{|1 + x|} \\
&= \frac{1}{2}\beta^{1-r} \\
&= 2^{-r} && (\beta = 2) \\
\iff \quad x &\leq \frac{2^{-r}}{1 - 2^{-r}}
\end{aligned}$$

da $2^{-r} \ll 1$, kann annährend $x \leq 2^{-r}$ als Abschätzung genommen werden. Dann

- Für double-precision: $r = 53 \Rightarrow x \leq 2^{-53} \approx 1.1102230246251565e{-}16$ gilt $1{+}x = 1$
- Für single-precision: $r = 24 \Rightarrow x \leq 2^{-24} \approx 5.960464477539063e - 8$ gilt $1 + x = 1$

c) Nein, das sind nicht die kleinsten darstellbaren Zahlen, sondern die Zahlen, bei denen die Abschätzung des relativen Fehlers erhalten bleibt. Die kleinsten (normalen) Zahlen erhält man dadurch, dass man die Exponenten auf die kleinsten Werte setzt:

- single precision: $2^{-126} = 1.17549435 \times 10^{-38}$
- double precision: $2^{-1022} = 2.2250738585072014 \times 10^{-308}$