Arbeitsblatt:
Maschinenzahlen und Gleitpunktarithmetik

Prof. Dr. B. Naumer

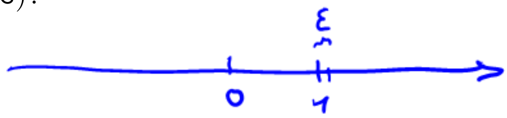
Lernziele:

- Sie verstehen, welche Probleme durch die Binärzahldarstellung entstehen und dadurch, dass Maschinenzahlen endlich viele Stellen haben.
- Sie sind sich der Unterschiede zwischen exakter Arithmetik und Gleitpunktarithmetik bewusst und können dieses Wissen auf konkrete Probleme anwenden.
- Sie berücksichtigen die Effekte der Gleitpunktarithmetik bei der Formulierung von Lösungsverfahren.

Ex 1. Warum ist $0.1 + 0.1 + 0.1 \neq 0.3$ in `R`?
Arbeitsschritte:

1. Mit welcher Genauigkeit rechnet `R` (s. Variable `.Machine`)?

   Double precision $\varepsilon = 2^{-52}$

2. Berechnen Sie in `R` die Differenz zwischen der Summe $0.1 + 0.1 + 0.1$ und 0.3.

3. Können Sie daraus eine Erklärung für obiges Problem ableiten?

4. Überlegen Sie sich Alternativen für die (numerische) Überprüfung auf Gleichheit!

   Wenn der rel. Fehler $\leq \varepsilon$, dann kann man 2 Zahlen in Gleitpunktarithmetik als identisch bezeichnen!

Ex 2. Welcher Effekt tritt bei der Gleitpunkt-Addition von einer großen mit einer kleinen Zahl auf?
Beispiel: $1.000 \cdot 10^{-2} + 4.000 \cdot 10^{3}$ (Genauigkeit: 4 signifikante Ziffern)

E_1 E_2

Arbeitsschritte:

1. Für eine Erklärung zur Durchführung der Gleitpunkt-Addition s. z. B. www.youtube.com/watch?v=IiHK18n0pm4
2. Führen Sie die Gleitpunktaddition mit einer Genauigkeit von 4 signifikanten Ziffern durch und vergleichen Sie das berechnete Ergebnis mit dem exakten.
3. Geben Sie eine Regel dafür an, wann der *Verlust signifikanter Stellen* auftritt. Wie hängt dieses Problem mit der Genauigkeit des Gleitpunktsystems zusammen?

Alle Stellen der kleineren Zahl gehen verloren, wenn die Differenz der Exponenten $> n$ (hier: 4)

$|E_1 - E_2|$ (hier: 5)

Ex 3. Welchen Effekt beobachten Sie bei der Auswertung von $f(x) = x(\underbrace{\sqrt{x+1} - \sqrt{x}}_{\text{Auslöschungseffekt}})$ mit einer Genauigkeit von 6 signifikanten Ziffern?

x	$f(x)$ berechnet	$f(x)$ exakt	$\sqrt{x+1} - \sqrt{x}$
1	0.414210	0.414214	$4 \cdot 10^{-1}$
10	1.54340	1.54347	$1.5 \cdot 10^{-4}$
100	4.99000	4.98756	$5 \cdot 10^{-2}$
1000	15.8000	15.8074	$1.5 \cdot 10^{-2}$
10000	50.0000	49.9988	$5 \cdot 10^{-3}$
100000	100.000	158.113	$1.6 \cdot 10^{-3}$

Arbeitsschritte:

1. Vergleichen Sie für $x = 100$ die exakte Auswertung mit der Auswertung in Gleitpunktarithmetik.

   In (6-stelliger) Gleitpunktarithmetik: $100 \cdot \left(rd\left(\sqrt{101.000}\right) - 10.0000\right)$

   $= 100 \cdot \left(10.0499 - 10.0000\right) = 4.99000$

2. Was ist mit *Auslöschung signifikanter Ziffern* gemeint?

   Wenn man kleine Zahlen als Differenz von zwei im Verhältnis dazu großen Zahlen berechnet, die sich nur in den hinteren Stellen unterscheiden, dann werden die signifikanten Ziffern "ausgelöscht".

3. Unter welchen Umständen ist dieser Effekt gut- bzw. bösartig?

   Wenn z.B. wie hier das Ergebnis von $\sqrt{x+1} - \sqrt{x}$ mit einer großen Zahl (für große x) mult. wird,

4. Finden Sie eine andere Formulierung für $f(x)$, die dieses Problem vermeidet.

$$f(x) = x\left(\sqrt{x+1} - \sqrt{x}\right) = x \, \frac{\left(\sqrt{x+1} - \sqrt{x}\right)\left(\sqrt{x+1} + \sqrt{x}\right)}{\sqrt{x+1} + \sqrt{x}} = x \cdot \frac{(x+1) - x}{\sqrt{x+1} + \sqrt{x}}$$

$$= \frac{x}{\sqrt{x+1} + \sqrt{x}}$$

Verständnistest:

- Gibt es Zahlen, deren Darstellung als Dezimalzahl endlich, aber als Dualzahl unendlich ist? Wenn ja, dann nennen Sie ein Beispiel.

  Ja, z.B. 0.1 oder 0.3 (s. Folien).

  Gibt es Zahlen, deren Darstellung als Dualzahl endlich, aber als Dezimalzahl unendlich ist? Wenn ja, dann nennen Sie auch dafür ein Beispiel.

  Nein, da alle ganzzahligen Potenzen von 2 Dezimalzahlen mit endlich vielen Nachkommastellen sind.

- Macht es einen Unterschied in welcher Reihenfolge die folgende Summe berechnet wird? Begründen Sie Ihre Entscheidung!

$$1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \cdots + \frac{1}{10^6}$$

  Ja, das Kommutativgesetz gilt nicht.
  Es ist immer besser in aufsteigender Reihenfolge zu addieren.

- Gesucht sind die Lösungen von $\overbrace{1}^{a} \cdot x^2 + \overbrace{200}^{b} x \overbrace{- 0.000015}^{c} = 0$.

  Mit welcher Formel würden Sie die Lösungen numerisch berechnen, wenn Sie nur mit geringer Genauigkeit rechnen könnten? Begründen Sie Ihre Entscheidung!

$$(1) \quad x_{1,2} = \frac{-b \pm \overbrace{\sqrt{b^2 - 4ac}}^{\approx b}}{2a} \qquad \text{oder} \qquad (2) \quad x_{1,2} = \frac{2c}{-b \mp \sqrt{b^2 - 4ac}}$$

Für $b > 0$: Auslöschung in (1) für x_1, dann Verwendung von (2)

im Beispiel $b = 200$

Für x_2 kann Formel (1) verwendet werden.

Für $b < 0$: genau andersherum

$$\frac{(-b \pm \sqrt{b^2-4ac}) \cdot (-b \mp \sqrt{b^2-4ac})}{2a \quad (-b \mp \sqrt{b^2-4ac})} \underset{\uparrow}{=} \frac{b^2 - (b^2 - 4ac)}{2a \cdot (-b \mp \sqrt{b^2-4ac})} = \frac{2c}{-b \mp \sqrt{b^2-4ac}}$$

3. binom. Formel