Technische Universität Berlin **Wintersemester 2014/2015**
Fakultät II – Mathematik und Naturwissenschaften
Institut für Mathematik
Günter Bärwolff, Jeroen Stolwijk

Numerische Mathematik I

1. Übungsblatt: Rechnerarithmetik und Rundungsfehler

Hausaufgaben: (Abgabe vor der Vorlesung am 5. November 2014)

Aufgabe 1: **(10 Punkte)**

Es gelte $x \oplus y = (x+y)(1+\epsilon)$ mit $|\epsilon| \leq k \cdot \text{eps} =: \epsilon^*$ für alle Maschinenzahlen x und y. Seien nun Maschinenzahlen $x_1, \ldots, x_n$ gegeben und sei durch

$$\tilde{s}_1 = x_1, \quad \tilde{s}_k = \tilde{s}_{k-1} \oplus x_k$$

ein Algorithmus zur Berechnung der Summe $s_k = \sum_{i=1}^{k} x_i$, $k \leq n$ definiert.

(a) Zeigen Sie: zerlegt man $\tilde{s}_k = s_k + f_k$, so gilt

$$|f_n| \leq [(1+\epsilon^*)^n - 1] \sum_{i=1}^{n} |x_i|.$$

Diskutieren Sie den relativen Fehler von $\tilde{s}_n$.

(b) Zeigen Sie, dass

$$\tilde{s}_n = \sum_{i=1}^{n} x_i(1+\delta_i) \quad \text{mit} \quad (1-\epsilon^*)^n - 1 \leq \delta_i \leq (1+\epsilon^*)^n - 1$$

und falls $n\epsilon^* < 1$ gilt:

$$|\delta_i| \leq \frac{n\epsilon^*}{1-n\epsilon^*}, \quad i = 1, \ldots, n.$$

Aufgabe 2: **(3 Punkte)**

Man betrachte die folgende Tabelle mit jeweils mathematisch äquivalenten Ausdrücken unter a) und b)

	a)	b)	
1.)	$(1+x)^2 - 1$	$x^2 + 2x$	$\|x\| \approx eps$
2.)	$\frac{\ln(1+x)}{x}$	$1 - \frac{x}{2} + \frac{x^2}{3}$	$0 < x \leq eps$

Bei welchen Ausdrücken tritt Auslöschung auf? Unterziehen sie sämtliche Ausdrücke einer Fehleranalyse.

Aufgabe 3: **(5 Punkte)**

Zu lösen sei die quadratische Gleichung

$$x^2 - 2px - q = 0 \quad \text{für} \quad p = 2,\ q = 0.0005$$

in vier- und fünfstelliger Gleitkommaarithmetik im Dezimalsystem. Dabei sollen die folgenden Algorithmen untersucht werden:

(i) $d = p^2 + q$, $x_1 = p + \sqrt{d}$, $x_2 = p - \sqrt{d}$, ($p - q$ Formel)

(ii) $d = p^2 + q$, $x_1 = p + \sqrt{d}$, $x_2 = -q/x_1$. (Vietascher Wurzelsatz)

Vergleichen Sie die Ergebnisse mit der exakten Lösung und erklären Sie die unterschiedlichen Resultate.

Aufgabe 4: **(2 Punkte)**
Bei exakter Arithmetik gilt für das arithmetische Mittel $\bar{x} = \frac{x_1+x_2}{2}$ zweier reeller Zahlen x_1 und x_2 die Ungleichung

$$\min\{x_1, x_2\} \leq \bar{x} \leq \max\{x_1, x_2\}. \tag{1}$$

Für Gleitpunktzahlen und das gemäß $\bar{x} = \mathrm{rd}(\mathrm{rd}(x_1 + x_2)/2)$ berechnete Mittel gilt die Ungleichung (1) im Allgemeinen nicht.
Finde Gleitpunktzahlen $x_1, x_2 \in \mathcal{F}(10, 2, -3, 3)$ derart, dass (1) verletzt ist. Wie ist die Berechnungsvorschrift für das arithmetische Mittel zu modifizieren, damit die Ungleichung (1) auch für das berechnete Mittel gilt?

Programmieraufgabe 1: (Abgabe in den Rechnersprechstunden bis zum 6. November 2014)
Schreiben Sie Matlab-Programme

```
e=meps(),    m=minimum(),    m=maximum()
```

die das Maschienenepsilon `eps`, die kleinste darstellbare positive Zahl x_{min} bzw. die grösste darstellbare Zahl x_{max} berechnen. Dabei ist `eps` als die kleinste Zahl definiert für die $1 \oplus$ `eps` > 1 gilt.
In den Programmen soll nur benutzt werden, dass intern eine Gleitkommadarstellung basierend auf dem Dualsystem verwendet wird. Vergleichen Sie ihre Ergebnisse mit denen der entsprechenden MATLAB Funktionen (`eps,realmin,realmax`). Interpretieren Sie die Unterschiede.

Übungsaufgabe für die Tutorien (27.10-31.10.2014):

Aufgabe 1:
Transformiere die folgenden Dezimalzahlen in ihre Binär- bzw. Hexadezimaldarstellung (d.h. $b = 2, 16$)

a) $x = 1.3$

b) $x = 12.875$

Aufgabe 2:
Es beschreibe $g(x) = ax + b$ mit $a \neq 0$ eine Gerade mit eindeutiger Nullstelle x_0. Die Funktion $f : \mathbb{R}\backslash\{0\} \times \mathbb{R} \to \mathbb{R}$ mit $f(a, b) = x_0$ gebe eine Methode an, welches aus den Koeffizienten a und b die Nullstelle berechnet. Berechnen Sie die relative und absolute Kondition des Problems.

Aufgabe 3:
Man berechne die absolute und relative Kondition der folgenden Probleme (jeweils für $x \geq 0$) beschrieben durch

a) $f(x) = x^3$ für kleine und große x,

b) $f(x) = \ln(x)$ für kleine und große x.

Aufgabe 4:
Bei welchem der folgenden (analytisch gleichen) Ausdrücken tritt Auslöschung auf?

a) $\frac{1}{n} - \frac{1}{n+2}$

b) $\frac{2}{n(n+2)}$

Man unterziehe beide Ausdrücke einer Fehleranalyse.