Letzte Themen: **Scilab**

- Kontrollstrukturen: Verzweigung, Schleifen
- Erzeugen von Plots
- Funktionen: Syntax, Parameter, Parameterübergabe

Heutige Themen: **Numerik**

- Darstellung von Gleitpunktzahlen
- Rundungsfehler durch Gleitpunktarithmetik
- Auslöschung

- Auf jedem Rechner sind nur endlich viele Zahlen exakt darstellbar $\Rightarrow$ **normalisierte Gleitpunktdarstellung:**

$$x = f * b^e$$

  - b: die Basis des Zahlensystems,
  - e: der Exponent (innerhalb fester Schranken: $r \leq e \leq R$),
  - f: die Mantisse mit einer festen Anzahl m von Stellen,

$$f = \pm 0.d_1 \ldots d_m, \quad 0 \leq d_i \leq b-1, \quad d_1 \neq 0 \text{ für } x \neq 0.$$

- Die Eingabedaten müssen mittels Maschinenzahlen approximiert werden:

  Standardrundung $\mathrm{fl} : \mathbb{R} \rightarrow \mathbb{M}(b, m, r, R)$.

- Relativer Rundungsfehler:

$$\left|\frac{\mathsf{fl}(x) - x}{x}\right| \quad \leq \quad \frac{b^{1-m}}{2} =: \mathbf{eps}$$

- Die Zahl $\mathbf{eps}$ wird **Maschinengenauigkeit** genannt.

- $\mathbf{eps}$ ist die kleinste positive Zahl, die zu 1 addiert von der Rundung noch wahrgenommen wird:

$$\mathsf{fl}(1 + \mathbf{eps}) > 1$$

- $\mathsf{fl}(x) = x(1 + \delta), \quad |\delta| < \mathbf{eps}.$

- Die Verknüpfung von Maschinenzahlen durch exakte elementare arithmetische Operationen liefert nicht notwendig eine Maschinenzahl: deswegen **Gleitpunktarithmetik**, z.B.:

$$x \oplus y = \mathsf{fl}(x + y)$$

- Die typischen Eigenschaften der exakten Arithmetik (z.B. Assoziativität, Distributivität...) gelten nicht.

- Bei der Addition und Subtraktion kann eine sehr große Fehlerverstärkung (schlechte Kondition) auftreten $\Rightarrow$**Auslöschung**.

  $f(x, y) = x + y$ hat die relative Kondition

$$\kappa_{rel} = \frac{\max(|x|, |y|)}{|x + y|}$$

Sei $f : \mathbb{R} \longrightarrow \mathbb{R}$ differenzierbar an der Stelle $x \in \mathbb{R}$.
Definition der **Ableitung**:

$$f'(x) \quad = \quad \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

wird approximiert durch einen **Differenzenquotient**

$$D(f, x, h) \quad = \quad \frac{f(x+h) - f(x)}{h} \quad \approx \quad f'(x)$$

Approximationsfehler:

$$D(f, x, h) - f'(x) \;=\; C \cdot h$$

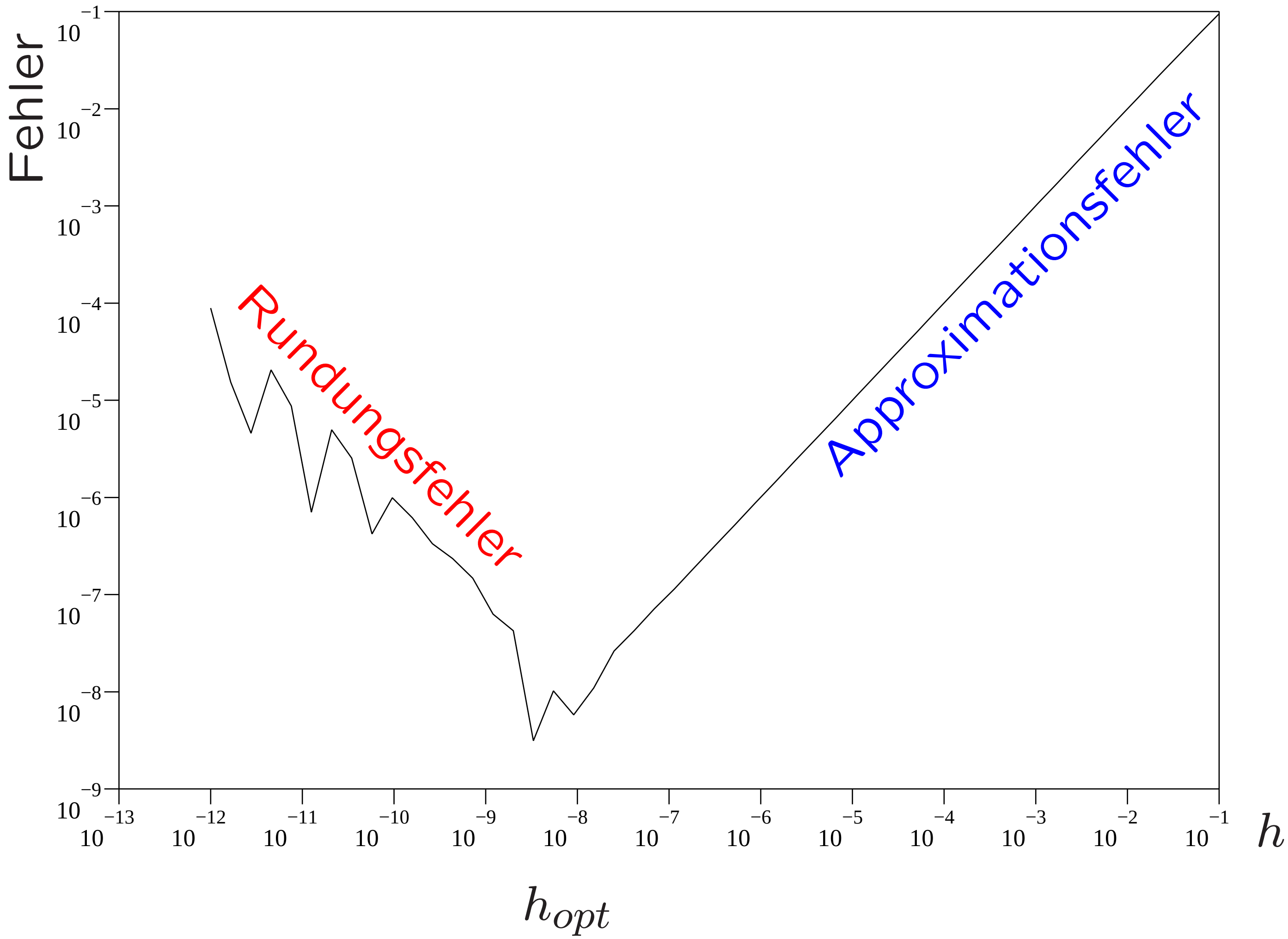
Diff.quotient mit Ordnung p : $\qquad C \cdot h^p$

Aufgrund von **Rundungsfehlern** durch GPA erhalten wir aber

$$D(f,x,h) \quad = \quad \frac{f(x+h)-f(x)+\varepsilon}{h},$$

wobei ε eine kleine Störung ist. Damit erhält man für den **Gesamtfehler**

$$\begin{aligned} Err(f,x,h) \;&:=\; D(f,x,h) \;-\; f'(x) \\ &\approx\; C \cdot h \;+\; \frac{\varepsilon}{h}. \end{aligned}$$


Fehler
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
10^{-8}
10^{-9}
10^{-13}
10^{-12}
10^{-11}
10^{-10}
10^{-9}
10^{-8}
10^{-7}
10^{-6}
10^{-5}
10^{-4}
10^{-3}
10^{-2}
10^{-1}
h
Rundungsfehler
Approximationsfehler
h_{opt}

Wir haben eine vektorwertige Funktion $f : \mathbb{R}^3 \to \mathbb{R}^3$:

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \qquad f(x) = \begin{pmatrix} f_1(x_1, x_2, x_3) \\ f_2(x_1, x_2, x_3) \\ f_3(x_1, x_2, x_3) \end{pmatrix}$$

Zur Erinnerung: Die Jacobi-Matrix von f ist

$$\begin{aligned} J_f(x) = f'(x) &= \left(\frac{\partial}{\partial x_1} f(x), \ \frac{\partial}{\partial x_2} f(x), \ \frac{\partial}{\partial x_3} f(x) \right) \\ &= \begin{pmatrix} \frac{\partial}{\partial x_1} f_1(x) & \frac{\partial}{\partial x_2} f_1(x) & \frac{\partial}{\partial x_3} f_1(x) \\ \frac{\partial}{\partial x_1} f_2(x) & \frac{\partial}{\partial x_2} f_2(x) & \frac{\partial}{\partial x_3} f_2(x) \\ \frac{\partial}{\partial x_1} f_3(x) & \frac{\partial}{\partial x_2} f_3(x) & \frac{\partial}{\partial x_3} f_3(x) \end{pmatrix}. \end{aligned}$$

Wir nähern $f'(x)$ durch Differenzenquotienten an:

$$J_k(x) = \frac{\partial}{\partial x_k} f(x) \quad \approx \quad \frac{f(x + h \cdot I_k) - f(x - h \cdot I_k)}{2h}$$

für $k = 1, 2, 3$.

Was ist zu tun?

1. Funktion `DJh(f,x,h)` schreiben, *testen* für $h = 10^{-5}$.
2. Fehler plotten (ähnlich wie in Versuch 1).

   ```
   for k=...
       Err(k) = norm( DJh(fct,x,hvec(k)) - J(x) )
   end
   ```

3. Jacobi-Matrix: `Ja`, `Jb` oder `Jc`? Warum?