

Hinweise:

- **Bearbeitungszeit:** 120 Minuten
 - **Hilfsmittel:** einfacher Taschenrechner (kein CAS, keine Grafik)
 - Schreiben Sie bitte auf jedes Blatt Ihren **Name**.
 - Geben Sie **Lösungswege** an und begründen Sie Ihre Antworten!
-

1. Bezeichnen im Folgenden mit “float8” binäre Gleitkommazahlen mit Mantissenlänge 5 Bit (einschließlich Hidden-Bit) und Exponentenlänge 3 Bit. Das erste Bit codiert das Vorzeichen, dann folgt der Exponent, dann die Mantisse. Die dargestellte Zahl ist also

$$(-1)^S (2^4 + M) 2^{E-7},$$

wobei S, M, E die in dem Bitmuster $SEEMMMM$ kodierten Ganzzahlen sind. Abgesehen von der geringen Länge sollen float8-Zahlen die üblichen Eigenschaften normalisierter Gleitkommazahlen nach IEEE 754 haben.

- (a) Welcher Dezimalzahl entspricht die binär dargestellte float8-Zahl 11100110?
(2 Punkte)
- (b) Finden Sie die Dezimaldarstellung für die betragskleinste echt negative normalisierte float8-Zahl! (2 Punkte)
2. Wir rechnen mit normalisierten dezimalen Gleitkommazahlen $m 10^e$, wobei m höchstens 3 Stellen hat und $e \geq 0$ gelten soll. Gerundet wird stets mathematisch. Berechnen Sie die folgenden Ausdrücke indem Sie vor und nach jeder Rechenoperation auf eine darstellbare Zahl runden.
- (a) $(8600 + 64) + 64$ (2 Punkte)
- (b) $8600 + (64 + 64)$ (2 Punkte)
- (c) Vergleichen Sie die berechneten Werte mit dem exakten Ergebnis und begründen Sie eventuelle Abweichungen! (2 Punkte)
3. Die Funktion $f(x) = x^2 + 14x + 49$ soll für verschiedene Eingaben x ausgewertet werden.
- (a) Bestimmen Sie die Konditionszahl in Abhängigkeit von x ! Für welche Eingaben ist die Kondition gut, für welche ist sie schlecht? (4 Punkte)
- (b) Die Funktion soll an der Stelle 5.001 ausgewertet werden. Der Einfachheit halber wird aber mit $x = 5$ gerechnet. Wie groß ist der relative Eingabefehler? Wie groß ist der zu erwartende Fehler im Funktionswert? (Auf das Zusammenfassen der entstehenden Ausdrücke zu Dezimalzahlen kann verzichtet werden.) (4 Punkte)

- (c) Die Funktion kann auch in der Form

$$f(x) = (x + 7)(x + 7)$$

geschrieben werden. Diese liefert einen alternativen Algorithmus zur Berechnung der Funktionswerte. Ändert sich dadurch die Kondition? Wenn ja: Wie? Begründen Sie Ihre Antwort! (2 Punkte)

4. Führen Sie für die Funktion

$$f(x) = 2x^3 + x^2 - 3x - 3$$

zwei Schritte des Newton-Verfahrens zur Nullstellenbestimmung manuell aus! Nutzen Sie $x_0 := 0$ als Startwert! Geben Sie die allgemeine Iterationsvorschrift sowie die beiden Iterierten an! Ist hier Konvergenz des Verfahrens zu erwarten (mit Begründung)? (6 Punkte)

5. Die Konditionzahl der Matrix A sei $\kappa_2(A) = 10$.

- (a) Wie groß darf der relative Fehler in b beim exakten Lösen von $Ax = b$ sein, damit der relative Lösungsfehler höchstens 3 Prozent beträgt? (3 Punkte)
- (b) Ist diese Fehlerschranke erfüllt, wenn statt der exakten rechten Seite $b = [6, 6, 3]^T$ die fehlerbehaftete rechte Seite $\tilde{b} = [5.9998, 6.0001, 3.0002]^T$ verwendet wird? (3 Punkte)

6. Die drei Punkte $(0, 2)$, $(1, 2)$, $(2, 3)$ sollen bestmöglich (im Sinne der Methode der kleinsten Quadrate) durch eine Parabel angenähert werden. Stellen Sie die Zielfunktion für das zu lösende Minimierungsproblem auf und geben Sie ein lineares Gleichungssystem (Systemmatrix und rechte Seite) an, welches die Lösung des Minimierungsproblems liefert! Das Gleichungssystem muss nicht (!) gelöst werden. (5 Punkte)

7. Geben Sie das Interpolationspolynom in Lagrange-Form zu den Punkten $(0, 0)$, $(1, 4)$, $(2, 1)$, $(3, 4)$ an! Die Lagrange-Basispolynome vom Grad $n - 1$ zu den Stützstellen x_1, \dots, x_n sind

$$L_{n-1,k}(x) := \prod_{l \neq k} \frac{x - x_l}{x_k - x_l}, \quad k = 1, \dots, n.$$

(3 Punkte)

8. Was bedeutet Normalisierung im Zusammenhang mit Gleitkommazahlen? Warum wird sie stets vorgenommen? Gehen Sie auch auf den Begriff des Hidden-Bit ein! (3 Punkte)
9. Wie unterscheidet sich das Konvergenzverhalten von Bisektionsverfahren und Newton-Verfahren beim Lösen nichtlinearer Gleichungen? Nennen Sie mindestens zwei wesentliche Unterschiede! (2 Punkte)