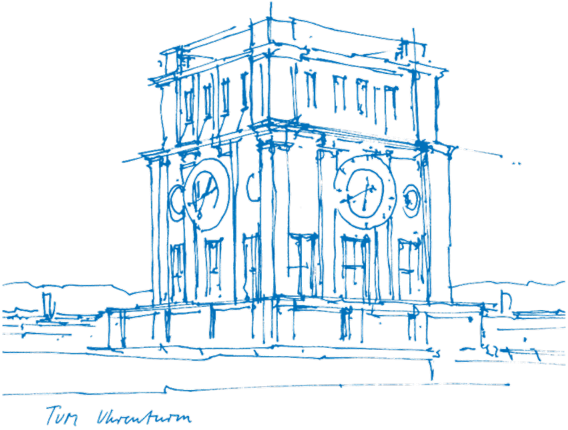
Lehrstuhl für Rechnerarchitektur und Parallelsysteme
Fakultät der Informatik
Technical University of Munich

Grundlagenpraktikum: Rechnerarchitektur

Gruppe 175 - Abgabe zu Aufgabe A316



Viktor Bayo, Hans Preinfalk, Georgy Chomakhashvili

Lehrstuhl für Rechnerarchitektur und Parallelsysteme
Fakultät der Informatik
Technische Universität München

August 31st, 2022

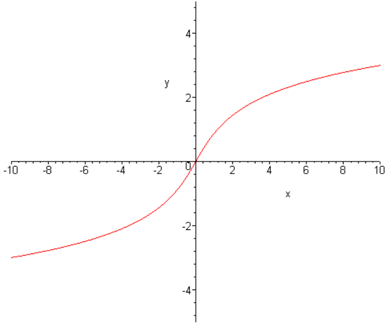
Outline

Problemstellung

- Approximation der Umkehrfunktion von sinh (also von arcsin)
- Theoretischer Teil
  - Herleitung
    - Reiehndarstellung
    - Lookup-Tabellen
  - Analyse
    - Genauigkeit
    - Laufzeit
- Praktischer Teil
  - $double\ approxArsinh\_series(double\ x)$
  - $double\ approxArsinh\_lookup(double\ x)$

Problemstellung
Eigenschaften der Funktion

- Die Funktion: $\text{arsinh}(x) = \ln(x + \sqrt{x^2 + 1})$ mit $x \in \mathbb{R}$
- Definitionsbereich $]-\infty; +\infty[$
- Wertebereich $]-\infty; +\infty[$
- punktsymmetrisch zum Ursprung
- streng monoton steigend
- Asymptote: $f(x) \to \pm\ln(2|x|)$ für $x \to \pm\infty$
- Ableitung: $\frac{d}{dx}\,\text{arsinh}(x) = \frac{1}{\sqrt{x^2+1}}$



Outline

Lösungsansatz

Rahmenprogramm

- Flags: V < int >, B < int >,< float >,h,help
- Hauptimplementierung - Reihendarstellung
  - V, V0 - Hauptimplementierung
  - V1 - Lookup-Tabellen
  - V2 - Vergleichsimplementierung
- optionale Argumente ≥ 0 und $\in N$
- Standardwerte
  - Zahl - 0
  - Hauptimplementierung
  - Ein Durchlauf
- nur eine oder keine Zahl

Lösungsansatz

Reihendarstellung(Herleitung)

- $|x| < 1$
  - arsinh$(x) = \ln(x + \sqrt{x^2 + 1})$

$= \int \frac{dx}{\sqrt{1+x^2}}$

$= \int (1 + x^2)^{-\frac{1}{2}} dx$ binomische Reihe

$= \int \sum_{k=0}^{\infty} (-1)^k \cdot \binom{-\frac{1}{2}}{k} \cdot (x^2)^k$ Binomialkoeffizienten

$= \int \sum_{k=0}^{\infty} (-1)^k \cdot \frac{(2k-1)!!}{(2k)!!} \cdot x^2$ Taylorreihe

$= \text{x} - \frac{1}{2} \cdot \frac{x^3}{3} + \frac{1 \cdot 3 \cdot x^5}{2 \cdot 4 \cdot 5} - \frac{1 \cdot 3 \cdot 5 \cdot x^5}{2 \cdot 4 \cdot 6 \cdot 7} + \ldots + (-1)^k \frac{1 \cdot 3 \cdot 5 \cdot \ldots \cdot (2k-1) \cdot x^{2k+1}}{2 \cdot 4 \cdot 6 \cdot \ldots \cdot 2k \cdot (2k+1)} + \ldots$

$= \text{x} \cdot \sum_{k=0}^{\infty} \frac{(2k-1)!!(-x^2)^k}{(2k)!!(2k+1)}$

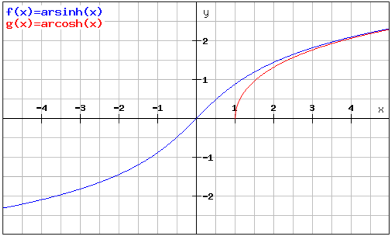
Lösungsansatz
Reihendarstellung(Herleitung)

- $|x| \geq 1$
  - Nach der Relation zwischen zwei hyperbolischen Funktionen Areasinus und Areakosinus

arsinh(x) = (sign(x)) arcosh($\sqrt{x^2+1}$)

wobei sign(x))$-$Signumfunktion

wobei arcosh(x) = $\ln(2x) - \sum_{k=0}^{\infty} \frac{(2k-1)!!}{(2k)\cdot(2k)!!} \cdot x^{-2k}$



Lösungsansatz
Reihendarstellung (Implementierung)

- $|x| < 1$
  - Taylorreihe
- $|x| >= 1$
  - Taylorreihe der Areakosinus Hyperbolicus
- $|x| < 12$
  - Reihendarstellung des Logarithmus
- $|x| >= 12$
  - Logarithmus rekursiv berechnen

Lösungsansatz
Lookup Tabellen(Herleitung)

- lineare Interpolation
  - Die Werte der Geraden als Approximation

$$f(x) = m \cdot x + b \qquad \text{mit} \qquad m = \tfrac{f(x_1)-f(x_0)}{x_1-x_0} \qquad \text{und} \qquad b = f(x_0) - \tfrac{f(x_1)-f(x_0)}{x_1-x_0} \cdot x_0$$

$$f(x) = \tfrac{f(x_1)-f(x_0)}{x_1-x_0} \cdot x + f(x_0) - \tfrac{f(x_1)-f(x_0)}{x_1-x_0} \cdot x_0$$

Einsetzen

$$f(x) = f(x_0) + \tfrac{(x-x_0)}{x_1-x_0} \cdot (f(x_1) - f(x_0))$$

lineare Interpolation

- Punkte proportional zueinander wählen

Lösungsansatz
Lookup Tabellen(Implementierung)

1. Mit dem Betrag arbeiten, da punktsymmetrisch
2. Mit aufsteigenden Standartwerten vergleichen
2. Passende Tabelle finden (insgesamt 93)
3. Passendes Intervall finden ($x_{i-1} < x <= x_i$ mit $x_{i-1} \cdot 1.08 = x_i$)
4. Wenn $x == x_i \rightarrow$ arsinh-Wert von der Tabelle ablesen
5. Wenn $x < x_i$ und $x > x_{i-1}$, dann lineare Interpolation mit x_{i-1} und x_i berechnen
6. Berechneten arsinh-Wert je mit 1 oder -1 multiplizieren
7. Wert zurückgeben

Outline

Genauigkeit

- Präzision von Fließkommazahlen ist hardwarebedingt!
- Die Variation der Genauigkeit abhängig von der Eingabegröße
  - $|x| < 1$
  - $|x| == 1$ (oder nahe 1)
  - $1 \leq |x| < 12$
  - $|x| > 12$
- V1
  - Zwischen 0 und 3 Nachkommastellen
  - Je kleiner der Abstand zwischen Punkten, desto genauer die Approximation
  - Verbesserungsvorschlag: Proportionalitätsfaktor c verkleinern

Genauigkeit

- Beispielausgaben (rot markierte Bereiche sind fehlerhaft)

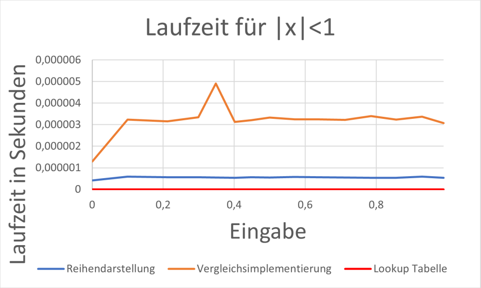
x	V0	V1	V2	Erwartet
0.5	0.48121182505960371	0.48114817845353913	0.48121182505960371	0.481211825059603447
1.0	0.881373587019542493	0.881122391422050288	0.881373587019543270	0.881373587019543025
9.37545	2.934073864818847799	2.933432601961815767	2.934073864818852684	2.934073864818852671
4324356.456	15.972921071536200444	15.972582216140555289	15.972921071536228865	15.972921071536229602
98888888888.5	26.010409902892391187	26.009677936276474952	26.010409902892391187	26.010409902892390020

Outline

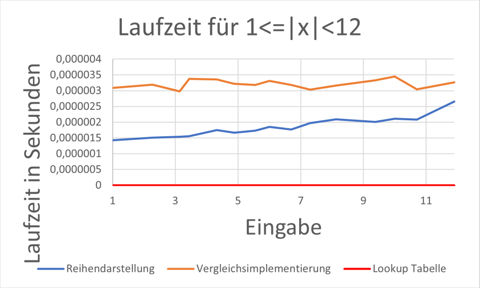
Performanz

- Wall Time
  - Während der Messung verstrichene Gesamtzeit
- CPU Time
  - Zeit für Verarbeitung des Programmes in CPU
- Monotonic Time vs Real Time
- Die steigende bzw. sinkende Laufzeit von V0 und V2 abhängig von der Eingabegröße
  - $|x| < 1$
  - $|x| < 12$
  - $|x| >= 12$
- V1 - Gewinner!

Performanz

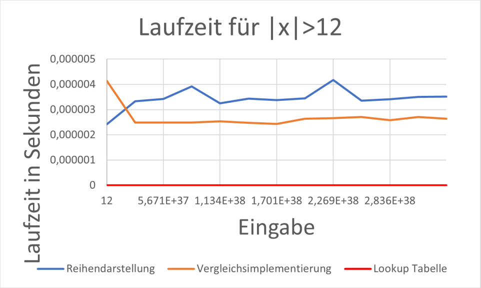


Laufzeit für |x|<1
0,000006
0,000005
0,000004
0,000003
0,000002
0,000001
0
0
0,2
0,4
0,6
0,8
Laufzeit in Sekunden
Eingabe
Reihendarstellung
Vergleichsimplementierung
Lookup Tabelle

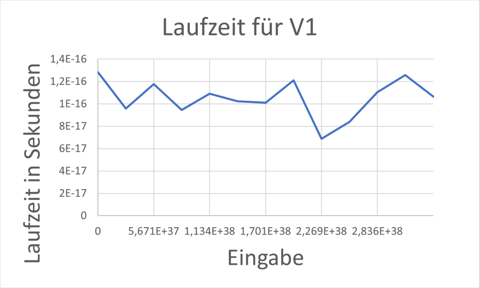


Laufzeit für 1<=|x|<12
0,000004
0,0000035
0,000003
0,0000025
0,000002
0,0000015
0,000001
0,0000005
0
1
3
5
7
9
11
Laufzeit in Sekunden
Eingabe
Reihendarstellung
Vergleichsimplementierung
Lookup Tabelle

Performanz



Laufzeit für |x|>12
0,000005
0,000004
0,000003
0,000002
0,000001
0
12
5,671E+37
1,134E+38
1,701E+38
2,269E+38
2,836E+38
Eingabe
Laufzeit in Sekunden
Reihendarstellung
Vergleichsimplementierung
Lookup Tabelle

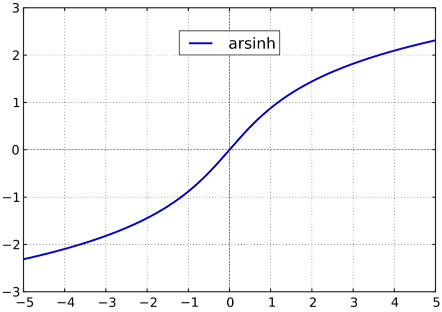


Laufzeit für V1
1,4E-16
1,2E-16
1E-16
8E-17
6E-17
4E-17
2E-17
0
0
5,671E+37
1,134E+38
1,701E+38
2,269E+38
2,836E+38
Eingabe
Laufzeit in Sekunden

Outline

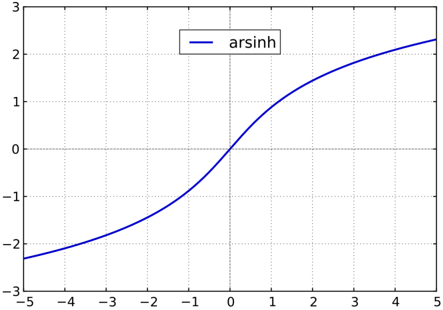
Zusammenfassung

- Rahmenprogramm
  - Entgegennahme bestimmter Optionen
  - Eingabeprüfung
- Reihendarstellung
  - mathematische Umformungen
  - höhere Genauigkeit
- Lookup-Tabellen
  - die Wahl zwischen Interpolationen
  - höhere Performanz



Zusammenfassung

- Genauigkeit
  - abhängig von Eingabe
  - V0: 4-16 Nachkommastellen
  - V1: 0-3 Nachkommastellen
  - V2: 4-16 Nachkommastellen
- Performanz
  - abhängig vom Definitionsbereich
  - V0 schneller als V1 für kleine Eingabewerte
  - V1 um ein Vielfaches schneller als V0 und V2



Durch dieses Programm erfolgt eine genaue und performante Berechnung