Computing 13, 299—312 (1974)

Arithmetiken für komplexe Kreise

M. Hauenschild, Bochum

Eingegangen am 4. März 1974

Zusammenfassung — Abstract

Arithmetiken für komplexe Kreise. Drei Einschließungskreise werden definiert für Mengen, die bei Addition, Subtraktion, Multiplikation und Division komplexer Kreise entstehen. Die algebraische Struktur der dadurch erzeugten Arithmetiken und die Radien der Einschließungskreise werden verglichen. Neben zwei von N. Krier [5] dargestellten Arithmetiken sind die kleinsten Einschließungskreise angegeben, durch die die algebraische Struktur von $P(\mathbb{C})$ voll erhalten bleibt („optimale Kreisarithmetik"). Bei der Berechnung einer inversen Matrix mit Arithmetiken für komplexe Kreise und achsenparallele Rechtecke werden Fehlerschranken und Rechenzeiten angegeben.

Arithmetic Procedures for Complex Circles. Three enclosing circles are defined for sets that are obtained by addition, subtraction, multiplication and division of complex circles. The algebraic structures resulting from these arithmetic procedures and the radii of the enclosing circles are compared. In addition to the two arithmetic procedures described by N. Krier [5], the smallest enclosing circles are defined which preserve the algebraic structure of $P(\mathbb{C})$ ("optimal circular arithmetic"). Calculations of an inverse matrix are made by arithmetic procedures for complex circles and axially-parallel rectangles, and the error bounds and computation times are compared.

1. Problemstellung

Die im Körper der komplexen Zahlen $\mathbb{C}$ definierten Operationen Addition, Subtraktion, Multiplikation und Division induzieren entsprechende Operationen in $P(\mathbb{C})$, der Potenzmenge von $\mathbb{C}$, die wir wie die Operationen in $\mathbb{C}$ durch die Symbole $+, -, \cdot$ und $:$ kennzeichnen.

Definition 1:

Für $A, B \in P(\mathbb{C})$ und $* \in \{+, -, \cdot, :\}$ sei $A * B := \{a * b \mid a \in A \wedge b \in B\}$.

Die Division $A : B$ ist nicht erklärt, falls $0 \in B$. Als Inversion bezeichnen wir den Übergang von B nach $B^{-1} := 1 : B$.

Die imaginäre Einheit in $\mathbb{C}$ sei i mit $i^2 = -1$, und ein Querstrich über einer komplexen Zahl kennzeichne den Übergang zur konjugiert komplexen Zahl.

Da die Teilmengen mit einem Element und die komplexen Zahlen isomorph sind, identifizieren wir $\{c\}$ mit c für alle $c \in \mathbb{C}$. (Vgl. zu Satz 1 [6], S. 55 und [1], S. 455.)

Satz 1:

Die durch Definition 1 erzeugte algebraische Struktur auf $\mathrm{P}(\mathbb{C})$ *hat folgende Eigenschaften:*

1. *Addition und Multiplikation sind kommutativ und assoziativ.*
2. *Es gilt das Subdistributivgesetz*
$$A\cdot(B+C)\subseteq A\cdot B+A\cdot C.$$
Für $A=a$ *gilt* $\qquad a\cdot(B+C)=a\cdot B+a\cdot C.$
3. *Für alle definierten Operationen gilt die Inklusionsmonotonie, d. h.*
$$A * B\subseteq C * D, \text{ falls } A\subseteq C \text{ und } B\subseteq D.$$
Damit gilt die Inklusionsmonotonie für beliebige rationale Ausdrücke.
4. *Die Neutralelemente sind* $0\in\mathbb{C}$ *für die Addition und* $1\in\mathbb{C}$ *für die Multiplikation.*
5. *Inverse Elemente existieren nur für Teilmengen mit einem Element, d. h. für* $c\in\mathbb{C}$ $(c\neq 0$ *bei der Multiplikation).*

Als Teilmenge von $\mathrm{P}(\mathbb{C})$ definieren wir die Menge der (komplexen) Kreise $\mathrm{K}(\mathbb{C})$. (Vgl. [5], S. 7.)

Definition 2:

Es sei $\mathrm{K}(\mathbb{C}):=\{Z=[z;r]\mid z\in\mathbb{C}\wedge r\in\mathbb{R},\, r\geq 0\}$. Dabei ist
$$Z=[z;r]:=\{c\in\mathbb{C}\mid |c-z|\leq r\}$$
der (komplexe) Kreis mit dem Radius $r(Z):=r$ und dem Zentrum $z(Z):=z$.

$\mathbb{C}$ ist Teilmenge von $\mathrm{K}(\mathbb{C})$. Bei der Verknüpfung zweier Kreise nach Definition 1 entstehen in den folgenden Fällen wieder Kreise. (Vgl. [3], S. 308 und [5], S. 10 ff.)

Satz 2:

Für $Z, Z_1, Z_2\in\mathrm{K}(\mathbb{C})$ *erhalten wir*

1. $Z_1\pm Z_2=[z_1\pm z_2;\, r_1+r_2]$.
2. $Z^{-1}=\left[\dfrac{\bar{z}}{z\bar{z}-r^2};\, \dfrac{r}{z\bar{z}-r^2}\right]$, *falls* $|z|>r$.
3. $Z_1\cdot Z_2=[z_1 z_2;\, |z_1|\, r_2+|z_2|\, r_1+r_1 r_2]$, *falls* $z_k=0$ *oder* $r_k=0$ *für mindestens ein k.*
4. $Z_1 : Z_2=\left[\dfrac{z_1\bar{z}_2}{z_2\bar{z}_2-r_2^2};\, \dfrac{|z_1|\, r_2+|\bar{z}_2|\, r_1+r_1 r_2}{z_2\bar{z}_2-r_2^2}\right]$, *falls* $|z_2|>r_2$ *und* $z_k=0$ *oder* $r_k=0$ *für mindestens ein k.*

Die im folgenden definierten Einschließungskreise sollen die bei der Verknüpfung komplexer Kreise nach Definition 1 entstehenden Mengen enthalten. Wenn $z_k\neq 0<r_k$ für $k=1, 2$, erhalten wir bei Multiplikation und Division zweier Kreise

kein Element aus K(ℂ). Vor der Berechnung der Einschließungskreise müssen wir diese Mengen untersuchen. Die Division werden wir stets als Komposition von Inversion und Multiplikation erklären. Da nach Satz 1 für $z_k \neq 0$

$$[z_1; r_1] \cdot [z_2; r_2] = z_1\, z_2 \cdot \left[1; \frac{r_1}{|z_1|}\right] \cdot \left[1; \frac{r_2}{|z_2|}\right] \tag{1}$$

gilt, können wir uns auf den Spezialfall $[1; a] \cdot [1; b]$ mit $0 < a, b$ beschränken.

Wir definieren nun drei Ansätze für Einschließungskreise.

Definition 3:

Sei $Z_1, Z_2 \in K(\mathbb{C})$ und $* \in \{+, -, \cdot, :\}$. Dann bestimmen die folgenden Operationen Arithmetiken für (komplexe) Kreise:

1. $Z_1 \mathbin{\triangle\!\!\!\!*}\, Z_2 := \triangle(Z_1 * Z_2)$ erzeugt die **minimale Kreisarithmetik.** Dabei ist $\triangle(Z_1 * Z_2)$ der $Z_1 * Z_2$ enthaltende Kreis mit minimalem Radius.
2. $Z_1 \circledast Z_2 := \bigcirc(Z_1 * Z_2)$ erzeugt die **optimale Kreisarithmetik.** Dabei ist $\bigcirc(Z_1 * Z_2)$ der $Z_1 * Z_2$ enthaltende Kreis mit minimalem Radius, der enthalten ist in dem Kreis $[0; \sup\{|c| \mid c \in Z_1 * Z_2\}]$.
3. $Z_1 (*) Z_2 := ()\,(Z_1 * Z_2)$ erzeugt die **zentrierte Kreisarithmetik.** Dabei ist $()\,(Z_1 * Z_2)$ der $Z_1 * Z_2$ enthaltende Kreis mit minimalem Radius, dessen Zentrum sich durch die entsprechende Verknüpfung der Zentren der Ausgangskreise ergibt, d. h. $z(Z_1 (*) Z_2) = z_1 * z_2$.

Soweit die Operationen nach Definition 3 mit denen nach Definition 1 übereinstimmen, werden wir auch die einfachen Operationszeichen verwenden.

Die Arithmetiken nach Definition 3.1 und 3.3 hat N. Krier [5] dargestellt und die erste optimal-abschätzende Kreisarithmetik genannt. Wir werden uns mit der Arithmetik nach Definition 3.2 ausführlicher beschäftigen.

2. Der Radius des kleinsten [1;a] · [1;b] einschließenden Kreises mit vorgegebenem reellem Zentrum

Wir bestimmen zunächst den relevanten Teil des Randes von $[1; a] \cdot [1; b]$, nämlich den Durchschnitt des Randes mit dem Rand der konvexen Hülle. Er ist ein Teil der Einhüllenden der Kurvenschar $F(\varphi) := \{(1 + a e^{i\varphi})(1 + b e^{i\psi}) \mid 0 \le \psi \le 2\pi\}$ für $0 \le \varphi \le 2\pi$. Die Einhüllende ist Lösung der Gleichung

$$\left(|y-1|^2 - (a^2 + b^2 + a^2 b^2)\right)^2 = (2ab)^2 \left(1 + a^2 + b^2 + 2\,\mathrm{Re}\,(y-1)\right). \tag{2}$$

Die Symmetrie der Randkurve zur reellen Achse macht es sinnvoll, nur reelle Punkte als Zentren vorzugeben. Aus (2) bestimmen wir den Abstand der Randpunkte y vom vorgegebenen reellen Zentrum y_m als Funktion des Realteils der Randpunkte. Soweit die Ordnung der Größen $a, b, 1$ in ℝ zu Fallunterscheidungen führen würde, setzen wir $u := \min\{a, b, 1\}$, $v := \mathrm{med}\,\{a, b, 1\}$ und $w := \max\{a, b, 1\}$, so daß stets $0 < u \le v \le w$ gilt. Für den relevanten Teil der Randkurve erhalten wir nach der Substitution $x := y - 1$

$$r_{x_m}(\mathrm{Re}\,x) := |x - x_m|$$
$$= \sqrt{a^2+b^2+a^2 b^2+x_m^2-2\,x_m \,\mathrm{Re}\,x+2\,ab\sqrt{1+a^2+b^2+2\,\mathrm{Re}\,x}} \tag{3}$$

für den Bereich $x_1 \le \mathrm{Re}\,x \le x_3$. Dabei ist

$$x_1 := \begin{cases} -\frac{1}{2}(1+a^2+b^2), \text{ falls } w \le u+v, \\ uv-uw-vw, \text{ falls } w \ge u+v, \end{cases} \tag{4}$$
$$x_3 := a+b+ab.$$

r_{x_m} ist stetig für $x_1 \le \mathrm{Re}\,x \le x_3$ und zweimal stetig differenzierbar für $x_1 < \mathrm{Re}\,x < x_3$ mit negativer 2. Ableitung, also streng monoton fallender 1. Ableitung

$$\frac{d}{d(\mathrm{Re}\,x)} r_{x_m}(\mathrm{Re}\,x) = \frac{-x_m\sqrt{1+a^2+b^2+2\,\mathrm{Re}\,x}+ab}{r_{x_m}(\mathrm{Re}\,x)\sqrt{1+a^2+b^2+2\,\mathrm{Re}\,x}}. \tag{5}$$

Den Radius $R(x_m)$ des kleinsten $[1;a]\cdot[1;b]$ einschließenden Kreises um $x_m \in \mathbb{R}$ erhalten wir dann als Maximum von r_{x_m}, d. h.

$$R(x_m) := \max\{r_{x_m}(\mathrm{Re}\,x) \mid x_1 \le \mathrm{Re}\,x \le x_3\}.$$

R ist eine stetig differenzierbare Funktion des vorgegebenen reellen Zentrums $x_m = y_m - 1$, die wir nur im Bereich $0 \le x_m \le uv = \min\{a, b, ab\}$ betrachten wollen.

Falls $0 \le x_m \le \frac{ab}{1+a+b}$, nimmt r_{x_m} das Maximum in $\mathrm{Re}\,x = x_3$ an, da die 1. Ableitung nicht negativ wird.

Falls $\frac{ab}{1+a+b} < x_m \le uv$, wird r_{x_m} maximal für $\mathrm{Re}\,x = x_2$, d. h. für die Nullstelle der 1. Ableitung von r_{x_m}, die bestimmt ist durch

$$2\,x_1 < 2\,x_2 := \left(\frac{ab}{x_m}\right)^2 - (1+a^2+b^2) < 2\,x_3. \tag{6}$$

Den gesuchten Radius berechnen wir nun aus (3) und (6).

Satz 3:

Der Radius des kleinsten $[1;a]\cdot[1;b]$ einschließenden Kreises um das vorgegebene reelle Zentrum x_m ist in den Bereichen A) $0 \le x_m \le \frac{ab}{1+a+b}$ und B) $\frac{ab}{1+a+b} < x_m \le uv$ bestimmt durch die stetig differenzierbare Funktion

$$R(x_m) = \begin{cases} A)\; r_{x_m}(x_3) = a+b+ab-x_m \\ B)\; r_{x_m}(x_2) = \sqrt{(1+x_m)(a^2+x_m)(b^2+x_m)\,x_m^{-1}} \end{cases} \tag{7}$$

mit der Ableitung

$$\frac{d}{d\,x_m} R(x_m) = \begin{cases} A) \quad -1 \\ B) \quad \frac{1}{2\,x_m^2\,R(x_m)}\left(2\,x_m^3+(1+a^2+b^2)\,x_m^2-a^2 b^2\right). \end{cases} \tag{8}$$

Minimalen Radius erhalten wir für die Nullstelle x_0 der 1. Ableitung mit $\frac{ab}{1+a+b}<x_0<uv$. *Für $x_m=x_0$ gilt*

$$R(x_0)=r_{x_o}(x_0)=\sqrt{3\,x_0^2+2\,(1+a^2+b^2)\,x_0+a^2+b^2+a^2\,b^2}. \tag{9}$$

3. Minimale Kreisarithmetik

Für die in Satz 2 aufgeführten Verknüpfungen zweier Kreise nach Definition 1 ist der minimale Einschließungskreis nach Definition 3.1 der in Satz 2 angegebene Kreis. Für $z_k \neq 0 < r_k$, $k=1,2$ erhalten wir bei Multiplikation und Division zweier Kreise den minimalen Einschließungskreis aus (1) und (9), da stets $Z_1 : Z_2 = Z_1 \cdot Z_2^{-1}$.

Satz 4:

Bei der Verknüpfung zweier Kreise nach Definition 3.1 (*minimale Kreisarithmetik*) *erhalten wir für $z_k \neq 0 < r_k$, $k=1,2$*

1. $Z_1 \triangle Z_2 = \big[z_1\,z_2\,(1+x_0);\ (3\,|z_1\,z_2|^2\,x_0^2+2\,(|z_1\,z_2|^2+|z_1|^2\,r_2^2$
$$+|z_2|^2\,r_1^2)\,x_0+|z_1|^2\,r_2^2+|z_2|^2\,r_1^2+r_1^2\,r_2^2)^{\frac{1}{2}}\big].$$
   Dabei ist x_0 die positive Nullstelle des Polynoms P mit
$$P(x) := 2\,|z_1\,z_2|^2\,x^3+(|z_1\,z_2|^2+|z_1|^2\,r_2^2+|z_2|^2\,r_1^2)\,x^2-r_1^2\,r_2^2.$$
2. $Z_1 \triangle\!\!\!\!: Z_2 = Z_1 \triangle Z_2^{-1}$.
3. *Bei den übrigen Verknüpfungen gelten die Formeln aus Satz* 2.

Bei der minimalen Kreisarithmetik ist der hohe Rechenaufwand bei der Multiplikation und die Verletzung wesentlicher Strukturaussagen von Satz 1 unbefriedigend.

Es gelten nicht:
1. das Assoziativgesetz der Multiplikation,
2. das Subdistributivgesetz und
3. die Inklusionsmonotonie bei Multiplikation und Division.

Zu 3. bringen wir ein Beispiel. Sei $Z_1 := [i;1]$, $Z_2 := [2\,i;1.5]$ und $Z_3 := [0;2]$. Dann gilt $Z_1 \subseteq Z_3$, aber

$$Z_1 \triangle\!\!\!\!: Z_2 = Z_1 \triangle Z_2^{-1} = [i;1] \triangle [-\tfrac{8}{7}\,i;\tfrac{6}{7}] =$$
$$[1.6091803524;\ 2.4817743384] \not\subseteq [0;4] =$$
$$[0;2] \triangle [-\tfrac{8}{7}\,i;\tfrac{6}{7}] = Z_3 \triangle Z_2^{-1} = Z_3 \triangle\!\!\!\!: Z_2,$$

da

$$r(Z_3 \triangle\!\!\!\!: Z_2)-r(Z_1 \triangle\!\!\!\!: Z_2)<1.519<1.609<|z(Z_3 \triangle\!\!\!\!: Z_2)-z(Z_1 \triangle\!\!\!\!: Z_2)|.$$

Nach der Definition von Fischer ([2], S. T190) ist die Inklusionsmonotonie notwendig für komplexe Intervall-Arithmetiken. Bei der Untersuchung von Gegenbeispielen erkennen wir als wesentlichen Grund für die Verletzung der genannten Strukturaussagen, daß für $z_k \neq 0 < r_k$, $k=1,2$

$$\sup\{|c| \mid c \in Z_1 \triangle Z_2\} > \sup\{|c| \mid c \in Z_1 \cdot Z_2\} \tag{10}$$

gilt. Deshalb betrachten wir nun den Einschließungskreis nach Definition 3.2, durch den das Betragssupremum nicht vergrößert wird.

4. Optimale Kreisarithmetik

Die in Satz 2 angegebenen Formeln können wir auch für die optimalen Einschließungskreise nach Definition 3.2 übernehmen. Bei dem in Satz 3 behandelten Spezialfall der Multiplikation wird das Betragssupremum von $[1;a]\cdot[1;b]$ in $y_3 = x_3 + 1 = (1+a)(1+b)$ angenommen. Wählen wir das Zentrum x_m im Bereich $\frac{ab}{1+a+b} < x_m \leq uv$, so wird das Betragssupremum $1 + r_{x_m}(x_2) > 1 + r_{x_m}(x_3) = (1+a)(1+b)$, da nach (5) $\frac{d}{d(\operatorname{Re} x)} r_{x_m}(\operatorname{Re} x) < 0$ für $x_2 < \operatorname{Re} x < x_3$. Den Einschließungskreis nach Definition 3.2 finden wir also im Bereich $0 \leq x_m \leq \frac{ab}{1+a+b}$ für $x_m = \frac{ab}{1+a+b}$. Wir erhalten

$$[1;a] \odot [1;b] = \left[\frac{(1+a)(1+b)}{1+a+b};\ (a+b)\,\frac{(1+a)(1+b)}{1+a+b}\right]. \tag{11}$$

Der allgemeine Fall folgt nun aus (1).

Satz 5:

Bei der Verknüpfung zweier Kreise nach Definition 3.2 (optimale Kreisarithmetik) erhalten wir für $z_k \neq 0 < r_k$, $k = 1, 2$

1. $Z_1 \odot Z_2 = [z_1 z_2 (1+x_m);\ (|z_1| r_2 + |z_2| r_1)(1+x_m)]$

$$\textit{mit } x_m = \frac{r_1 r_2}{|z_1 z_2| + |z_1| r_2 + |z_2| r_1}.$$

2. $Z_1 \odot\!\!\!\!\cdot\; Z_2 = Z_1 \odot Z_2^{-1}$

$$= \left[\frac{z_1 \bar{z}_2 (1+x_m)}{z_2 \bar{z}_2 - r_2^2};\ \frac{(|z_1| r_2 + |\bar{z}_2| r_1)(1+x_m)}{z_2 \bar{z}_2 - r_2^2}\right],$$

$$\textit{falls } |z_2| > r_2 \textit{ mit } x_m = \frac{r_1 r_2}{|z_1 \bar{z}_2| + |z_1| r_2 + |\bar{z}_2| r_1}.$$

3. *Bei den übrigen Verknüpfungen gelten die Formeln aus Satz 2.*

Die Untersuchung der von der optimalen Kreisarithmetik auf $K(\mathbb{C})$ induzierten algebraischen Struktur ergibt, daß alle Merkmale von Satz 1 erfüllt sind.

Satz 6:

Bei der optimalen Kreisarithmetik nach Definition 3.2 *erhalten wir:*

1. *Die Multiplikation ist assoziativ.*
2. *Es gilt das Subdistributivgesetz.*
3. *Die Inklusionsmonotonie gilt auch für die Multiplikation und Division.*

Von dem umfangreichen Beweis bringen wir nur die wichtigsten Teile.

zu 1: Das Assoziativgesetz ergibt sich im Spezialfall aus

$$([1;a]\odot[1;b])\odot[1;c]=[1;a]\odot([1;b]\odot[1;c])$$
$$=\left[\frac{(1+a)(1+b)(1+c)}{1+a+b+c};\,(a+b+c)\,\frac{(1+a)(1+b)(1+c)}{1+a+b+c}\right].$$

Für die folgenden Beweisteile benötigen wir zwei Abschätzungen. Seien c_1, $c_2\in\mathbb{C}-\{0\}$, p, q positive reelle Zahlen und φ der Winkel zwischen c_1 und c_2 mit $0\le\varphi\le\pi$.

a) Da $\displaystyle\int_0^{\varphi}\frac{pq\sin x}{\sqrt{p^2+q^2-2pq\cos x}}\,dx\le\int_0^{\varphi}\frac{pq}{p-q}\sin x\,dx$ für $p>q$,

$$\text{gilt } |c_1-c_2|-(|c_1|-|c_2|)\le\frac{|c_1c_2|}{|c_1|-|c_2|}(1-\cos\varphi) \text{ für } |c_1|>|c_2|. \tag{12}$$

b) Da $\displaystyle\int_0^{\varphi}\frac{pq\sin x}{\sqrt{p^2+q^2+2pq\cos x}}\,dx\ge\int_0^{\varphi}\frac{pq}{p+q}\sin x\,dx$,

$$\text{gilt } |c_1|+|c_2|-|c_1+c_2|\ge\frac{|c_1c_2|}{|c_1|+|c_2|}(1-\cos\varphi). \tag{13}$$

Wir vereinbaren noch folgende Abkürzung:

$$c_{i,k}:=|z_iz_k|+|z_i|r_k+|z_k|r_i.$$

zu 2: Wir beweisen das Subdistributivgesetz für $z_k\neq 0<r_k$, $k=1,2,3$ und $z_1+z_2\neq 0$. Mit $Z_5:=Z_2+Z_3$, $Z_6:=Z_1\odot Z_5$ und $Z_7:=Z_1\odot Z_2+Z_1\odot Z_3$ erhalten wir nach Satz 5

$$(z_6-z_7)\,\frac{c_{1,5}\,c_{1,2}\,c_{1,3}}{z_1r_1(|z_1|+r_1)}=\frac{z_3}{|z_3|}\,p-\frac{z_2}{|z_2|}\,q,$$

$$(r_7-r_6)\,\frac{c_{1,5}\,c_{1,2}\,c_{1,3}}{|z_1|r_1(|z_1|+r_1)}=p-q+(|z_2|+|z_3|-|z_2+z_3|)\,\frac{|z_2+z_3|}{|z_1|}\,c_{1,2}\,c_{1,3}$$

für

$$p:=|z_3|\,c_{1,2}\,(|z_3|(r_2+r_3)-|z_2+z_3|\,r_3)$$
$$q:=|z_2|\,c_{1,3}\,(|z_2+z_3|\,r_2-|z_2|(r_2+r_3))$$

und

$$p-q=|z_1|(|z_3|r_2-|z_2|r_3)^2$$
$$+(|z_2|+|z_3|-|z_2+z_3|)(|z_2|r_2 c_{1,3}+|z_3|r_3 c_{1,2}).$$

Da die Addition kommutativ ist, sei o. E. $|z_2|r_3 \le |z_3|r_2$. Für den Beweis der Ungleichung $|z_6-z_7| \le r_7-r_6$ ergeben sich dann noch zwei Fälle:

Fall a) $|z_2+z_3|r_2 \le |z_2|(r_2+r_3)$ und $|z_2+z_3|r_3 \le |z_3|(r_2+r_3)$;
Fall b) $|z_2+z_3|r_2 > |z_2|(r_2+r_3)$ und $|z_2+z_3|r_3 < |z_3|(r_2+r_3)$.
Im Fall a) ist $p \ge 0$ und $q \le 0$.

Wir erhalten folgende Abschätzung:

$$|z_6-z_7| \frac{c_{1,5}\, c_{1,2}\, c_{1,3}}{|z_1|\, r_1\, (|z_1|+r_1)} \le |p|+|q| = p-q$$
$$\le (r_7-r_6) \frac{c_{1,5}\, c_{1,2}\, c_{1,3}}{|z_1|\, r_1\, (|z_1|+r_1)}$$

Im Fall b) erhalten wir aus der Ungleichungskette

$$\frac{|z_2|}{r_2} < \frac{|z_2+z_3|}{r_2+r_3} \le \frac{|z_2|+|z_3|}{r_2+r_3} < \frac{|z_3|}{r_3}$$

die Voraussetzungen für (12) und (13), nämlich $p > q > 0$. Da

$$\frac{pq}{p-q} \le \frac{|z_2|\,|z_3|\,|z_2+z_3|}{|z_1|(|z_2|+|z_3|)}\, c_{1,2}\, c_{1,3},$$

ergibt sich unsere Behauptung aus der Abschätzung

$$|z_6-z_7| \frac{c_{1,5}\, c_{1,2}\, c_{1,3}}{|z_1|\, r_1\, (|z_1|+r_1)} = (p^2+q^2-2\,pq\cos\varphi)^{\frac{1}{2}}$$
$$\le p-q+\frac{pq}{p-q}(1-\cos\varphi)$$
$$\le p-q+\frac{|z_2|\,|z_3|\,|z_2+z_3|}{|z_1|(|z_2|+|z_3|)}\, c_{1,2}\, c_{1,3}(1-\cos\varphi)$$
$$\le p-q+\frac{|z_2+z_3|}{|z_1|}\, c_{1,2}\, c_{1,3}(|z_2|+|z_3|-|z_2+z_3|)$$
$$=(r_7-r_6)\frac{c_{1,5}\, c_{1,2}\, c_{1,3}}{|z_1|\, r_1\, (|z_1|+r_1)}.$$

Dabei ist φ der Winkel zwischen z_2 und z_3 mit $0 \le \varphi \le \pi$.

zu 3: Wir beweisen die Inklusionsmonotonie für die Multiplikation, falls $z_k \ne 0 < r_k$ für alle k. Da die Multiplikation kommutativ ist, brauchen wir nur zu zeigen: $Z_1 \odot Z_3 \subseteq Z_2 \odot Z_3$ für $Z_1 \subseteq Z_2$. Mit $Z_4 := Z_1 \odot Z_3$ und $Z_5 := Z_2 \odot Z_3$ erhalten wir nach Satz 5

$$(z_4-z_5)\frac{c_{1,3}\, c_{2,3}}{z_3(|z_3|+r_3)} = (z_1-z_2)(|z_1|+r_1)\, c_{2,3} + z_2\, r_3 (|z_2|\, r_1 - |z_1|\, r_2);$$

$$(r_5-r_4)\frac{c_{1,3}\,c_{2,3}}{|z_3|(|z_3|+r_3)}=(r_2-r_1)(|z_1|+r_1)\,c_{2,3}+|z_2|\,r_3(|z_2|\,r_1-|z_1|\,r_2)$$
$$+\big((r_2-r_1)-(|z_1|-|z_2|)\big)\frac{|z_1|\,r_3}{|z_3|}\,c_{2,3}.$$

Beim Beweis der Ungleichung $|z_4-z_5|\le r_5-r_4$ unterscheiden wir drei Fälle:

Fall a) $z_1=z_2$,
Fall b) $r_2-r_1=|z_1-z_2|>0$ und $|z_1|\,r_2\le|z_2|\,r_1$,
Fall *c*) $r_2-r_1=|z_1-z_2|>0$ und $|z_1|\,r_2>|z_2|\,r_1$.

Im Fall a) erhalten wir folgende Abschätzung:

$$|z_4-z_5|\frac{c_{1,3}\,c_{2,3}}{|z_3|(|z_3|+r_3)}=|z_2|^2\,r_3(r_2-r_1)\le(r_5-r_4)\frac{c_{1,3}\,c_{2,3}}{|z_3|(|z_3|+r_3)}.$$

Ist $z_1\neq z_2$, so führen wir den Beweis in zwei Schritten. Wir setzen

$$Z_1^*:=[z_1;\,r_2-|z_1-z_2|].$$

Da wegen $Z_1\subseteq Z_2$ auch $Z_1\subseteq Z_1^*\subseteq Z_2$ gilt, zeigen wir

$$Z_1\odot Z_3\subseteq Z_1^*\odot Z_3\subseteq Z_2\odot Z_3.$$

Sei also $r_2-r_1=|z_1-z_2|>0$. Dann erhalten wir

$$(z_4-z_5)\frac{c_{1,3}\,c_{2,3}}{z_3(|z_3|+r_3)}=\frac{z_1-z_2}{|z_1-z_2|}\,p-\frac{z_2}{|z_2|}\,q,$$

$$(r_5-r_4)\frac{c_{1,3}\,c_{2,3}}{|z_3|(|z_3|+r_3)}=p-q+\big((r_2-r_1)-(|z_1|-|z_2|)\big)\frac{|z_1|\,r_3}{|z_3|}\,c_{2,3}$$

mit

$$p:=|z_1-z_2|(|z_1|+r_1)\,c_{2,3}=(r_2-r_1)(|z_1|+r_1)\,c_{2,3},$$
$$q:=|z_2|\,r_3(|z_1|\,r_2-|z_2|\,r_1)$$

und

$$p-q=(r_2-r_1)(|z_1|+r_1)(|z_2|+r_2)|z_3|+\big((r_2-r_1)-(|z_1|-|z_2|)\big)|z_2|\,r_1\,r_3.$$

Im Fall b) ist $p\ge 0$ und $q\le 0$. Wir erhalten folgende Abschätzung:

$$|z_4-z_5|\frac{c_{1,3}\,c_{2,3}}{|z_3|(|z_3|+r_3)}\le|p|+|q|=p-q\le(r_5-r_4)\frac{c_{1,3}\,c_{2,3}}{|z_3|(|z_3|+r_3)}.$$

Im Fall c) sind die Voraussetzungen von (12) und (13) erfüllt, nämlich $p>q>0$. Unsere Behauptung erhalten wir aus der folgenden Abschätzung, da

$$\frac{pq}{p-q}\le\frac{(r_2-r_1)|z_2|\,|z_1|\,r_3}{(|z_2|+r_2-r_1)|z_3|}\,c_{2,3}.$$

$$|z_4-z_5|\frac{c_{1,3}\,c_{2,3}}{|z_3|(|z_3|+r_3)}=(p^2+q^2-2\,pq\cos\varphi)^{\frac{1}{2}}$$

$$\le p-q+\frac{pq}{p-q}(1-\cos\varphi)$$

$$\leq p-q+\frac{(r_2-r_1)\,|z_2|\,|z_1|\,r_3}{(|z_2|+r_2-r_1)\,|z_3|}\,c_{2,3}\,(1-\cos\varphi)$$

$$\leq p-q+\big((r_2-r_1)-(|z_1|-|z_2|)\big)\,\frac{|z_1|\,r_3}{|z_3|}\,c_{2,3}$$

$$=(r_5-r_4)\,\frac{c_{1,3}\,c_{2,3}}{|z_3|\,(|z_3|+r_3)}.$$

Dabei ist φ der Winkel zwischen z_1-z_2 und z_2 mit $0\leq\varphi\leq\pi$. Aus der Inklusionsmonotonie für Inversion und Multiplikation erhalten wir diese Eigenschaft auch für die Division.

Bei der Kreisarithmetik nach Definition 3.2 berechnen wir die kleinsten strukturerhaltenden Einschließungskreise mit geringerem Rechenaufwand bei der Multiplikation. Deshalb haben wir sie optimale Kreisarithmetik genannt.

5. Zentrierte Kreisarithmetik

Die in Satz 2 angegebenen Formeln können wir mit Ausnahme der Inversion auch für die zentrierten Einschließungskreise nach Definition 3.3 übernehmen. Bei dem in Satz 3 behandelten Spezialfall der Multiplikation erhalten wir den zentrierten Einschließungskreis für $x_m=0$, also $[1;a]\,(\cdot)\,[1;b]=[1;a+b+ab]$. Das ergibt im allgemeinen Fall auch für $z_k\neq 0<r_k$ die in Satz 2 angegebene Formel.

Satz 7:

Bei der Verknüpfung zweier Kreise nach Definition 3.3 (*zentrierte Kreisarithmetik*) *erhalten wir*

1. *bei Addition und Subtraktion die Formel aus Satz* 2,

2. $1\,(:)\,Z=\left[\frac{1}{z};\frac{r}{|z|\,(|z|-r)}\right]$, *falls* $|z|>r$,

3. $Z_1\,(\cdot)\,Z_2=[z_1 z_2;\,|z_1|\,r_2+|z_2|\,r_1+r_1 r_2]$ *und*

4. $Z_1\,(:)\,Z_2=Z_1\,(\cdot)\,(1\,(:)\,Z_2)=\left[\frac{z_1}{z_2};\frac{|z_1|\,r_2+|z_2|\,r_1}{|z_2|\,(|z_2|-r_2)}\right]$, *falls* $|z_2|>r_2$.

Alle in Satz 1 aufgezeigten Strukturmerkmale bleiben auch bei der zentrierten Kreisarithmetik erhalten. Dasselbe gilt für den Ansatz von Gargantini und Henrici [3], einer Modifikation der minimalen Kreisarithmetik mit der Multiplikation nach Definition 3.3, die wir durch ein G unter den Operationszeichen kennzeichnen wollen. Für die arithmetischen Grundoperationen und damit für beliebige rationale Ausdrücke gilt die Inklusionsrelation:

$$Z_1 * Z_2 \subseteq Z_1 \circledast Z_2 \subseteq Z_1 \underset{G}{*} Z_2 \subseteq Z_1\,(*)\,Z_2.$$

6. Radiusvergleich der Einschließungskreise

Bei den drei Ansätzen nach Definition 3 erhalten wir unterschiedliche Kreise bei Inversion, Multiplikation und Division. Bei der Inversion wird nur der Radius des zentrierten Einschließungskreises vergrößert. Wir erhalten

$$\frac{r(1\,(:)\,Z)}{r(1:Z)}=1+\frac{r}{|z|}<2.$$

Der Radius bei der zentrierten Inversion kann also bis zu 100% größer werden als bei der minimalen und optimalen Inversion.

Auch bei der Multiplikation ist das Verhältnis der Radien nur von den Quotienten $a_k=\frac{r_k}{|z_k|}$ abhängig. Extremwerte erhalten wir für $a_1=a_2=1$.

$$\frac{r(Z_1 \odot Z_2)}{r(Z_1 \mathbin{\triangle} Z_2)}\le\left(\frac{256}{243}\right)^{\frac{1}{2}}\approx 1.02640$$

$$\frac{r(Z_1\,(\cdot)\,Z_2)}{r(Z_1 \mathbin{\triangle} Z_2)}\le\left(\frac{4}{3}\right)^{\frac{1}{2}}\approx 1.15470$$

$$\frac{r(Z_1\,(\cdot)\,Z_2)}{r(Z_1 \odot Z_2)}\le\frac{9}{8}=1.125$$

Gegenüber der minimalen Multiplikation ist der Radius bei der optimalen Multiplikation maximal 2.64% größer, bei der zentrierten bis zu 15.47% größer.

Ausgehend von dem zentrierten Einschließungskreis beträgt die relative Radiusverbesserung bei der optimalen Multiplikation

$$\frac{r(Z_1\,(\cdot)\,Z_2)-r(Z_1 \odot Z_2)}{r(Z_1\,(\cdot)\,Z_2)}=\frac{1}{(1+a_1+a_2)(1+\frac{1}{a_1}+\frac{1}{a_2})}.$$

Diese ist bei der minimalen Multiplikation höchstens 20.6% größer. Mit der optimalen Multiplikation erreichen wir also über 82.9% der möglichen Radiusverbesserung.

Bei der Division gilt für die Verhältnisse der Radien

$$\frac{r(Z_1 \odot Z_2)}{r(Z_1 \mathbin{\triangle} Z_2)}\le\left(\frac{256}{243}\right)^{\frac{1}{2}}\approx 1.02640,$$

$$\frac{r(Z_1\,(:)\,Z_2)}{r(Z_1 \odot Z_2)}\le\frac{r(Z_1\,(:)\,Z_2)}{r(Z_1 \mathbin{\triangle} Z_2)}<2.$$

Der Radius des Einschließungskreises bei der Division wird entscheidend bestimmt durch den Nullabstand des Divisors

$$N(Z_2):=\inf\{|c| \mid c\in Z_2\}.$$

Vorteile der minimalen und optimalen Multiplikation zeigen sich auch in der Vergrößerung dieses Nullabstandes gegenüber der zentrierten Kreisarithmetik. Bei der Multiplikation erhalten wir für die verschiedenen Einschließungskreise die folgenden Abschätzungen, falls die rechte Seite positiv ist und $|z_k|>r_k$ für $k=1,2$.

$$N(Z_1 \cdot Z_2)=(|z_1|-r_1)(|z_2|-r_2)$$

$$N(Z_1 \mathbin{\triangle\!\!\!\cdot} Z_2)>(|z_1|-r_1)(|z_2|-r_2)-0.7187\, r_1 r_2$$

$$N(Z_1 \odot Z_2)>(|z_1|-r_1)(|z_2|-r_2)-1\, r_1 r_2$$

$$N(Z_1 (\cdot) Z_2)=(|z_1|-r_1)(|z_2|-r_2)-2\, r_1 r_2$$

Werden die Radien im Vergleich zu den Zentren klein, dann verlieren — wie alle Abschätzungen zeigen — die Verbesserungen der minimalen und optimalen Inversion und Multiplikation an Bedeutung.

7. Testprogramm und Ergebnisse

Das im folgenden beschriebene Programm wurde in Algol 60 geschrieben und auf dem Telefunken-Rechner TR 440 des Rechenzentrums der Ruhr-Universität Bochum gerechnet. Für die Ansätze der optimalen und zentrierten Kreisarithmetik sowie für eine Arithmetik mit achsenparallelen Rechtecken nach Alefeld [1] und Rokne und Lancaster [9] wurden Unterprogramme für die arithmetischen Grundoperationen erstellt. Da die Programme sichere Fehlerschranken angeben sollten, mußten Rundungsfehler der Gleitkommaarithmetik berücksichtigt werden.

In **KCT** ist die optimale Kreisarithmetik auf der Grundlage der Triplex-Prozeduren für reelle Intervalle programmiert. Der Fehler bei der Berechnung des Zentrums wird zur oberen Schranke des Radius addiert. Der Hauptwert wird einfachgenau mit Rundung zur nächsten Maschinenzahl berechnet.

In **KCD** ist die zentrierte Kreisarithmetik mit doppeltgenauer Berechnung des mit dem Zentrum zusammenfallenden Hauptwertes programmiert. Der Fehler bei der Hauptwertberechnung wird durch den größten möglichen Rundungsfehler abgeschätzt und zum Radius addiert.

In **ICT** werden achsenparallele Rechtecke verknüpft auf der Grundlage der Triplex-Prozeduren für reelle Intervalle.

Das Unterprogramm **GAUSS** löst lineare Gleichungssysteme und invertiert reguläre Matrizen. Da der Name des gewünschten arithmetischen Unterprogramms als Parameter übergeben wird, können wir durch mehrfachen Aufruf von GAUSS das gleiche Problem mit unterschiedlichen Arithmetiken lösen.

Wir invertieren eine 10×10 Matrix mit willkürlich gewählten Hauptwerten betragsmäßig zwischen 0.04 und 1.3. Als Ausgangsfehler wählen wir a) 10^{-5} und b) den maximalen Abstand zur nächsten Maschinenzahl einfacher Länge, etwa 10^{-11}. Der Ausgangsfehler bezeichnet die halbe Kantenlänge des Fehlerquadrats, das wir durch einen Fehlerkreis einschließen. Wir vergleichen die Rechenzeiten und bei den Ergebnissen die Fehlerradien mit dem Radius des kleinsten, das Rechteck einschließenden Kreises.

Bei KCD haben wir bei doppeltgenauer Berechnung des Hauptwertes die günstigsten Rechenzeichen und die kleinsten Rundungsfehler. Da KCT mehr arithmetische Operationen erfordert als ICT, haben wir bei KCT größere Rechenzeiten und größere Rundungsfehler.

Im Lauf a) fallen bei relativ großem Ausgangsfehler die Rundungsfehler kaum ins Gewicht. Die Fehlerradien der optimalen Kreisarithmetik sind um 5% kleiner als die der zentrierten Kreisarithmetik, die „Radien" bei der Rechteckarithmetik etwa 13-mal so groß.

Im Lauf b) bilden die Rundungsfehler bei KCT und ICT den Hauptteil des Ergebnisfehlers. Deshalb ist bei ICT der Ergebnisfehler nur etwa 6-mal so groß wie bei KCT und bei KCT 11-mal so groß wie bei KCD.

Die besseren Ergebnisse bei der Kreisarithmetik erklären sich dadurch, daß sie eine exakte Einschließung bei der Inversion ermöglicht und bei der Multiplikation wegen der Rotationsinvarianz der Kreise bessere Einschließungen als bei der Reckteckarithmetik.

Die Radiusverbesserung der optimalen gegenüber der zentrierten Kreisarithmetik gewinnt nur bei größeren Fehlerradien Bedeutung. Dann kann sie aber entscheidend für den sicheren Nachweis einer Lösung sein.

	Lauf a	**Lauf b**
Ausgangsdaten:		
Hauptwertbetrag	0.04 — 1.3	0.04 — 1.3
Fehler	10^{-5}	0.22 — 58.2 $\cdot 10^{-12}$
rel. Fehler	7.7 — $250 \cdot 10^{-6}$	2.57 — 58.2 $\cdot 10^{-12}$
Ergebnisdaten:		
Hauptwertbetrag	0.398 — 10.53	0.398 — 10.53
rel. Fehler KCD	0.0424 — 0.3242	0.0146 — 0.111 $\cdot 10^{-6}$
Fehlerradius KCD	0.0548 — 0.893	0.0188 — 0.3065 $\cdot 10^{-6}$
Fehlerradius KCT	0.0522 — 0.849	0.21 — 3.403 $\cdot 10^{-6}$
Fehlerradius ICT	0.628 — 14.18	1.11 — 26.64 $\cdot 10^{-6}$

Verhältnis der Rechenzeiten und Fehlerradien:

	KCD	:	KCT	:	ICT
Rechenzeit:	13.5 s	:	25.5 s	:	15.0 s
	9	:	17	:	10
Fehlerradien:					
Lauf a	1	:	0.95	:	9.3 — 22.6
	1.052	:	1	:	9.8 — 23.8
Lauf b	1	:	11.1	:	48 — 123
	0.09	:	1	:	4.32 — 11

Literatur

[1] Alefeld, G.: Über Eigenschaften und Anwendungsmöglichkeiten einer komplexen Intervallarithmetik. Z. Angew. Math. Mech. *50*, 455—465 (1970).

[2] Fischer, H.: Intervall-Arithmetiken für komplexe Zahlen. Z. Angew. Math. Mech. *53*, T190—Z191 (1973).

[3] Gargantini, I., Henrici, P.: Circular arithmetic and the determination of polynomial zeros. Numer. Math. *18*, 305—320 (1972).
[4] Hauenschild, M.: Ansätze zur komplexen Kreisarithmetik. Arbeitsbericht des Rechenzentrums Nr. 7304, Universität Bochum, 1973.
[5] Krier, N.: Komplexe Kreisarithmetik. Dissertation, Universität Karlsruhe, 1973.
[6] Kulisch, U.: Grundzüge der Intervallrechnung. In: Überblicke Mathematik 2 (Laugwitz, D., Hrsg.), S. 51—98 (Hochschultaschenbücher, Band 232/232a). Mannheim-Zürich: Bibl. Institut 1969.
[7] Moore, R. E.: Interval analysis. Englewood Cliffs: Prentice Hall Inc. 1966.
[8] Nickel, K.: Zeros of polynomials and other topics. In: Topics in interval analysis (Hansen, E., Hrsg.), S. 25—34. Oxford: Clarendon Press 1969.
[9] Rokne, J., Lancaster, P.: Complex interval arithmetic. Comm. ACM *14*, 111—112 (1971).

Dipl.-Math. M. Hauenschild
Rechenzentrum
Ruhr-Universität Bochum
Postfach 2148
D-4630 Bochum
Bundesrepublik Deutschland