Über metrische und algebraische Eigenschaften einiger beim numerischen Rechnen auftretender Räume

von

Dipl.-Math. Edgar Kaucher

aus Bauschlott

Karlsruhe 1973

Über metrische und algebraische Eigenschaften einiger beim numerischen Rechnen auftretender Räume

Zur Erlangung des akademischen Grades eines
DOKTORS DER NATURWISSENSCHAFTEN
von der Fakultät für Naturwissenschaften I der
Universität Karlsruhe

genehmigte

Dissertation

von

Dipl.-Math. Edgar Kaucher
aus Bauschlott

Tag der mündlichen Prüfung: 21. 12. 1973
Referent: . . . . Professor Dr. U. Kulisch, Karlsruhe
Korreferent: . . Priv.-Dozent Dr. J. Herzberger, Karlsruhe

Karlsruhe 1973

Druck: O. Berenz, Karlsruhe

INHALTSVERZEICHNIS

EINLEITUNG

Die Anwendung von Rechenanlagen auf mathematisch-technische Probleme wirft die Frage auf, ob die errechneten Ergebnisse in einem brauchbaren Verhältnis zum theoretischen Ergebnis stehen. Tatsächlich wird die Konvergenz von Algorithmen in Räumen untersucht, welche i.a. keine Rechenanlage verwirklichen kann. Um also z. B. die Eigenschaften von Rechenprozessen über den reellen Zahlen axiomatisch zu erfassen, ist es notwendig, mathematische Räume zu definieren, welche die Eigenschaften sinnvoller Rechner wiederspiegeln.

Einen wesentlichen Schritt in dieser Richtung tat Kulisch [Kulisch 69], [Kulisch 70], [Kulisch 72], in der abstrakten Erfassung der Ordnungsstruktur und der algebraischen Struktur. Die entscheidenden Hilfsmittel hierzu sind die Begriffe für Rundung und Raster. Die Rundung wird erklärt als eine monotone Abbildung eines vollständigen Verbandes A in einen speziellen vollständigen Teilbund B, welcher als Raster bezeichnet wird. Kulisch konnte dann ausgehend von algebraischen Strukturen in A z. B. dem Körper der reellen Zahlen entsprechende algebraische Strukturen in B herleiten, z. B. die des Ringoides, welche sich als rundungsinvariant erwiesen.

In dieser Arbeit soll dieses Konzept auf topologische Strukturen fortgesetzt werden. Die algebraischen Strukturen in den oben erwähnten Rastern zeigen jedoch keine brauchbare Verträglichkeitseigenschaft mit den Metriken der Originalräume.

Es erscheint daher notwendig, für solche Räume eine geeignete Verallgemeinerung des Begriffes Metrik vorzunehmen. Dabei werden einerseits die wesentlichen Eigenschaften wie "Dreiecksungleichung" und "Abschätzbarkeit nach oben" erhalten bleiben, andererseits Verträglichkeitseigenschaften mit den algebraischen Strukturen der oben erwähnten Raster (z. B. mit der des Rasterringoides der reellen Zahlen) entstehen.

Diese Verallgemeinerung führt zu einem System von "gestörten Abständen", das ein "prämetrisches Feld" genannt wird. Diese prämetrischen Felder erlauben schließlich eine allgemeine Vorwärts- und Rückwärtsanalyse der Rundungsfehler unter gleichzeitiger Berücksichtigung des Unterlaufes wie es in spezieller Form von Wilkinson in [Wilkinson 69] durchgeführt ist.

Im zweiten Teil dieser Arbeit wird eine Verallgemeinerung der Intervallrechnung auf sogenannte nichtreguläre Intervalle vorgenommen. Der so erweiterte Intervallraum erreicht dadurch eine höhere algebraische Abgeschlossenheit, erlaubt eine geschlossene Darstellung des modifizierten Newtonverfahrens [Alefeld 68] und die Anwendung des "verallgemeinerten Fixpunktsatzes" für mit Rundungen behaftete Iterationsverfahren im Intervallraum, so daß in speziellen Fällen eine verbesserte Einschließung des Fixpunktes möglich wird.

I. Symbole, Grundbegriffe und einige grundlegende Sätze

1. Grundbegriffe des gerundeten Rechnens

In Anlehnung an Kulisch [Kulisch 69], [Kulisch 70], [Kulisch 72] werden folgende Strukturen und Begriffe eingeführt:

Definition 1.1: Sei $(M,\leq)$ eine geordnete Menge und $T \subseteq M$ eine Teilmenge von M. Für jedes $a \in M$ werde mit $L_T(a) := \{b \mid b \in T \wedge b \leq a\}$ bzw. $U_T(a) := \{b \mid b \in T \wedge a \leq b\}$ die Menge der unteren bzw. oberen Schranken von a in T definiert.

T ist dann ein unteres bzw. oberes Raster von M wenn

(S 1) $\bigwedge_{a \in M} L_T(a) \neq \emptyset$ bzw. $\bigwedge_{a \in M} U_T(a) \neq \emptyset$

(S 2) $\bigwedge_{a \in M} \bigvee_{x \in L_T(a)} \bigwedge_{b \in L_T(a)} b \leq x$ bzw.

$\bigwedge_{a \in M} \bigvee_{x \in L_T(a)} \bigwedge_{b \in L_T(a)} x \leq b$

Die Elemente x werden auch mit $i(L_T(a))$ bzw. $o(U_T(a))$ bezeichnet.

Ist T zugleich unteres und oberes Raster von M, so wird T ein Raster genannt.

Da später oft mehrere Raster verschiedener Räume zugleich betrachtet werden, soll folgende Schreibweise vereinbart werden: Ist $(M,\leq)$ die zugrundegelegte geordnete Menge, so ist $\underline{\mathcal{R}M}$ bzw. $\overline{\mathcal{R}M}$ unteres bzw. oberes Raster von M und $\mathcal{R}M$ Raster von M .

Folgender Satz gibt ein einfaches Kriterium an, wann eine Teilmenge ein Raster ist.

Satz 1.2:

a) Eine Teilmenge $(\mathcal{R}M,\leq)$ eines vollständigen Verbandes $(M,\leq)$ ist dann und nur dann ein unteres (bzw. oberes) Raster von $(M,\leq)$, wenn gilt:

(S1')	$o(M)=o(\mathcal{R}M)$	(bzw. $i(M)=i(\mathcal{R}M)$)
(S2')	$(\mathcal{R}M,\leq)$ ist ein vollständiger sup-Teilbund von $(M,\leq)$	bzw. $(\mathcal{R}M,\leq)$ ist vollständiger inf-Teilbund von $(M,\leq)$

b) Eine Teilmenge $(\mathcal{R}M,\leq)$ eines vollständigen Verbandes $(M,\leq)$ ist dann und nur dann ein Raster von $(M,\leq)$, wenn gilt:

(S1'') $o(M)=o(\mathcal{R}M)$ $i(M)=i(\mathcal{R}M)$ und

(S2'') $(\mathcal{R}M,\leq)$ ist vollständiger Teilverband von $(M,\leq)$

Bemerkung: Tritt künftig im Zusammenhang mit einer Menge M das Symbol $\mathcal{R}M$ auf, so wird stillschweigend vorausgesetzt, daß

a) $(M,\leq)$ ein geordneter vollständiger Verband

b) $\mathcal{R}M$ Raster von M ist entsprechend den Aussagen von Satz 1.2

<u>Definition 1.3:</u> Eine <u>Rundung</u> ist eine isotone Abbildung einer geordneten Menge $(M, \leq)$ auf ein unteres bzw. oberes Raster $\mathcal{R}M \subseteq M$, welche die Eigenschaft (R1) erfüllt:

$\square : M \longrightarrow M$ heißt

(R 1) monoton, wenn

$$\bigwedge_{a \in M} \bigwedge_{b \in M} (a \leq b \implies \square a \leq \square b)$$

(R 2) optimal, wenn

$$\bigwedge_{a \in \mathcal{R}M} \square a = a$$

(R 3) gerichtet, wenn

$$\bigwedge_{a \in M} \square a \leq a \quad \text{bzw.} \quad \bigwedge_{a \in M} \square a \geq a$$

Daß zu gegebenem Raster $\mathcal{R}M$ von M stets eine monotone, optimale und gerichtete Rundung existiert, die (R 1), (R 2) und (R 3) erfüllt, besagt folgender Satz:

<u>Satz 1.4:</u> Eine Abbildung $\Delta : M \longrightarrow \mathcal{R}M$ bzw. $\nabla : M \longrightarrow \mathcal{R}M$,

die die Eigenschaft

(R) $\bigwedge_{a \in M} \Delta a := i\left(L_{\mathcal{R}M}(a)\right)$ bzw. $\bigwedge_{a \in M} \nabla a := o\left(U_{\mathcal{R}M}(a)\right)$

erfüllt, ist eine monotone, optimale nach oben bzw. nach unten gerichtete Rundung. Sie ist eindeutig bestimmt.

Kulisch betrachtet in [Kulisch 72] bezüglich solcher Rundungen algebraische Strukturen, etwa Gruppoide und Ringoide. Diese Strukturen erweisen sich als rundungsinvariant.

In dieser Arbeit werden darüber hinaus auch Abbildungen betrachtet, welche durch arithmetische Operationen in einem Ringoid definiert sind. In vielen Fällen genügt zunächst folgende Beziehung zwischen Originalfunktion über M und Rasterfunktion über $\mathcal{R}M$.

<u>Definition 1.5:</u>

Ist $(M,\leq)$ Vollverband, Ω eine Menge von Abbildungen $\omega : M \to M$, $(\mathcal{R}M,\leq)$ Raster von M und

$\tilde{\Omega} := \{\omega \in \Omega \mid \omega|_{\mathcal{R}M} : \mathcal{R}M \to \mathcal{R}M\} \neq \emptyset$ eine Teilmenge aus Ω, so

heißt der Operator $\alpha : \Omega \to \tilde{\Omega}$ eine <u>Rasterrestringierung</u> einer Abbildung $\omega \in \Omega$ bezüglich $\mathcal{R}M$, das heißt, es gilt

$$\bigwedge_{\omega \in \Omega} \; \bigvee_{\alpha(\omega) \in \tilde{\Omega}} \; \bigwedge_{x \in \mathcal{R}M} \quad \alpha\omega x := (\alpha(\omega))(x) \in \mathcal{R}M$$

und $\alpha(\omega)$ ist eindeutig bestimmt.

α heißt

(A 1) isoton: $\displaystyle\bigwedge_{\omega \in \Omega} \; \bigwedge_{x,y \in \mathcal{R}M} \quad \omega x \leq \omega y \implies \alpha\omega x \leq \alpha\omega y$

(A 2) optimal: $\displaystyle\bigwedge_{x \in \mathcal{R}M} \quad \omega x \in \mathcal{R}M \implies \alpha\omega x = \omega x$

(A 3) gerichtet: $\displaystyle\bigwedge_{x \in \mathcal{R}M} \quad \omega x \leq \alpha\omega x$

(A 4)[1] ideal, wenn eine Rundung $\square: M \to \mathcal{R}M$ existiert, so daß

$$\bigwedge_{\omega \in \Omega} \bigwedge_{x \in \mathcal{R}M} \alpha\omega x = \square(\omega x) \qquad \text{gilt.}$$

Folgende weitere Begriffe werden aus [Kulisch 72] übernommen.

Definition 1.6:

a) Eine nichtleere Menge R, in der zwei 2-stellige innere Verknüpfungen, die wir als "Addition" (+) und "Multiplikation" (·) bezeichnen, erklärt sind, heißt ein "Ringoid", i. Z. $(R,+,\cdot)$, wenn die Axiome (D1) bis (D6) gelten.

(D1) $\bigwedge_{a,b \in R} a + b = b + a$.

(D2) $\bigvee_{o \in R} \bigwedge_{a \in R} a + o = a$.

(D3) $\bigvee_{e \in R \setminus \{o\}} \bigwedge_{a \in R} a \cdot e = e \cdot a = a$.

(D4) $\bigwedge_{a \in R} a \cdot o = o \cdot a = o$.

---

1) (A4) gilt z. B. bei arithmetischen Funktionen i. a. nur, wenn sie höchstens eine Verknüpfung enthalten. Vergl. Satz 1.8.

(D5) Es existiert ein Element $x \in R\setminus\{e\}$, so daß gilt:

(a) $x \cdot x = e$

(b) $$\bigwedge_{a,b \in R} x\cdot(a\cdot b) = (x\cdot a)\cdot b = a\cdot(x\cdot b)\ .$$

(c) $$\bigwedge_{a,b \in R} x\cdot(a+b) = x\cdot a + x\cdot b\ .$$

(D6) Es gibt genau ein Element $x \in R \setminus \{e\}$, für welches (D5) gilt.

Eine nichtleere Menge R, in der eine Addition und eine Multiplikation erklärt sind, welche die Axiome (D1) bis (D5) erfüllen, heißt ein "Semiringoid". Wir bezeichnen ein Semiringoid ebenfalls durch das Symbol $(R,+,\cdot)$.
Ein Ringoid (bzw. ein Semiringoid) heißt kommutativ, wenn gilt

$$\bigwedge_{a,b \in R} a\cdot b = b\cdot a$$

Ein Ringoid (bzw. ein Semiringoid) heißt "Divisionsringoid" (bzw. "Semidivisionsringoid"), i.Z. $(R,N,+,\cdot,/)$, wenn mit $o \in N \subseteq R$ in R eine weitere 2-stellige Verknüpfung, die wir als "Division" (/) bezeichnen, erklärt ist, für welche die Axiome (D7), (D8) und (D9) gelten.

(D7) $\bigwedge_{a \in R} a/e = a$.

(D8) $\bigwedge_{a \in R \setminus N} o/a = o$.

(D9) Es existiert ein Element $x \in R \setminus \{e\}$, welches außer dem Axiom (D5) auch noch die Eigenschaft

$$\bigwedge_{a \in R} \quad \bigwedge_{b \in R \setminus N} x \cdot (a/b) = (x \cdot a)/b = a/(x \cdot b)$$

erfüllt.

b) Es sei $\{R,+,\cdot\}$ ein Ringoid (bzw. ein Semiringoid) und $(R,\leq)$ eine geordnete Menge. $(R,+,\cdot,\leq)$ heißt "geordnetes Ringoid" (bzw. "geordnetes Semiringoid"), wenn zwischen der algebraischen Struktur und der Ordnungsstruktur in R die folgenden Verträglichkeitsbedingungen erfüllt sind

(OD1) $\bigwedge_{a,b,c \in R} (a \leq b \Longrightarrow a+c \leq b+c)$.

(OD2) $\bigwedge_{a,b,c \in R} (o \leq a \leq b \wedge c \geq o \Longrightarrow a \cdot c \leq b \cdot c \wedge c \cdot a \leq c \cdot b)$.

Es existiert ein Element $x \in R \setminus \{e\}$, welches außer (D5) (a),(b),(c) auch die Eigenschaft (OD3) besitzt.

(OD3) $\bigwedge_{a,b \in R} (a < b \Longrightarrow x \cdot b < x \cdot a)$.

Ein Element a eines geordneten Ringoides (bzw. geordneten Semiringoides) heißt "negativ", wenn $a < o$ ist und "positiv", wenn $a > o$ ist.

Ist $(R,+,\cdot)$ ein Ringoid (bzw. Semiringoid) und $(R,\leq)$ ein vollständiger Verband, dann heißt $(R,+,\cdot,\leq)$ ein "vollständiges, geordnetes Ringoid" (bzw. "vollständiges, geordnetes Semiringoid"); ist $(R,\leq)$ ein vollständiger, linear geordneter Verband, so heißt $(R,+,\cdot,\leq)$ ein "vollständiges, linear geordnetes Ringoid" (bzw. "vollständiges, linear geordnetes Semiringoid").

Ein Divisionsringoid (bzw. Semidivisionsringoid) $(R,N+,\cdot,/)$ heißt ein "geordnetes Divisionsringoid" (bzw. "geordnetes Semidivisionsringoid") $(R,N,+,\cdot,/,\leq)$, wenn $(R,+,\cdot,\leq)$ ein geordnetes Ringoid (bzw. geordnetes Semiringoid) ist und zusätzlich gilt:

(OD4a) $$\bigwedge_{a,b\in R}\ \bigwedge_{c\in R\setminus N} \quad o\leq a\leq b \wedge c>o \implies o\leq a/c\leq b/c$$

(OD4b) $$\bigwedge_{a,b\in R\setminus N}\ \bigwedge_{c\in R} \quad (o\leq a\leq b \wedge c\geq o \implies c/a\geq c/b\geq o)$$

Ein geordnetes Divisionsringoid $(R,N,+,\cdot,/,\leq)$ heißt ein "linear geordnetes Divisionsringoid", wenn $(R,\leq)$ eine linear geordnete Menge ist und zudem (OD5) erfüllt ist:

(OD5) $$\bigwedge_{a\in R\setminus N} \quad (a > o \implies a/a = e) \quad \text{mit} \quad N = \{o\} .$$

Ist $(R,N,+,\cdot,/,\leq)$ ein (linear) geordnetes Divisionsringoid und $(R,\leq)$ ein vollständiger Verband, dann heißt $(R,N,+,\cdot,/,\leq)$ ein "vollständiges (linear) geordnetes Divisionsringoid".

c) Es sei $(R,+,\cdot)$ ein Ringoid und $(R,\leq)$ eine geordnete Menge. $(R,+,\cdot,\leq)$ heißt ein "durchweg isoton geordnetes Ringoid", wenn gilt:

$$\text{(OD6)} \quad \bigwedge_{a,b,c,d\in R} (a\leq b \wedge c\leq d \Longrightarrow a\circ c\leq b\circ d) \quad \text{für alle } \circ\in\{+,\cdot\}$$

Ein Divisionsringoid $(R,N,+,\cdot,/,\leq)$ heißt ein "durchweg isoton geordnetes Divisionsringoid", wenn es ein durchweg isoton geordnetes Ringoid ist mit der zusätzlichen Eigenschaft:

$$\bigwedge_{a,b\in R} \ \bigwedge_{c,d\in R\setminus N} (a \leq b \wedge c \leq d \Longrightarrow a/c \leq b/d) .$$

<u>Bemerkung:</u>

Es genügt zum Nachweis der Eigenschaft (OD6), daß

$$\text{(OD6')} \quad \bigwedge_{a,b,c\in R} a \leq b \Longrightarrow a \circ c \leq b \circ c \quad \text{für beliebige}$$

Verknüpfung $\circ \in \{+,\cdot,/\}$ gilt.

Dies folgt aus: $a \leq b \wedge c \leq d \Longrightarrow \big(a \circ c \leq b \circ c \ \wedge$

$b \circ c \leq b \circ d \big) \Longrightarrow$

$a \circ c \leq b \circ c \leq b \circ d$

Definition 1.7:

Es sei $(R,+,\cdot)$ ein Ringoid, $(R,\leq)$ ein vollständiger Verband und $\mathcal{R}R$ ein symmetrisches (unteres, oberes) Raster von R. Eine Rundung $\square : R \to \mathcal{R}R$ heißt symmetrisch, wenn gilt:

$$(R4) \qquad \bigwedge_{a \in R} \square(-a) = -\square a$$

Satz 1.8: Es sei $(R,+,\cdot,\leq)$ ein vollständiges, linear geordnetes Ringoid (bzw. $(R,\{0\},+,\cdot,/,\leq)$ ein vollständiges, linear geordnetes Divisionsringoid) mit den ausgezeichneten Elementen $\{-e,0,e\}$ und $(\mathcal{R}R,\leq)$ ein symmetrisches Raster von $(R,\leq)$ mit der Eigenschaft $0,e \in \mathcal{R}R$ sowie $\square : R \longrightarrow \mathcal{R}R$ eine optimale, symmetrische Rundung von R in $\mathcal{R}R$. In $\mathcal{R}R$ seien zwei (bzw. drei) zweistellige innere Verknüpfungen $\boxed{\circ}$, $\circ \in \{+,\cdot\}$ (bzw. $\circ \in \{+,\cdot,/\}$), erklärt durch die Vorschrift

$$(V) \qquad \bigwedge_{a,b \in \mathcal{R}R} a \,\boxed{\circ}\, b := \square(a \circ b)$$

(bzw. zusätzlich:

$$\bigwedge_{a \in \mathcal{R}R} \; \bigwedge_{b \in \mathcal{R}R \setminus \{0\}} a \,\boxed{/}\, b := \square(a/b) \;)$$

Dann ist $(\mathcal{R}R, \boxed{+}, \boxed{\cdot}, \leq)$ ein vollständiges, linear geordnetes, optimales Rasterringoid (bzw. $(\ R, \{0\}, \boxed{+}, \boxed{\cdot}, \boxed{/}, \leq)$ ein vollständiges, linear geordnetes, optimales Rasterdivisionsringoid) mit den ausgezeichneten Elementen $\{-e,o,e\}$ und es gilt

$$\bigwedge_{a,b \in \mathcal{R}R} a \boxminus b = \square(a - b) .$$

Ist ferner $\{R,+,\cdot\}$ kommutativ, so ist auch $\{\mathcal{R}R, \boxplus, \boxdot\}$ kommutativ.

<u>Bemerkung:</u> Satz 1.8 besagt unter anderem, daß bei Abbildung eines linear geordneten Ringoides (bzw. eines linear geordneten Divisionsringoides) in ein symmetrisches Raster mittels einer optimalen, symmetrischen Rundung wieder ein linear geordnetes Ringoid (bzw. ein linear geordnetes Divisionsringoid) entsteht. Wir bringen diese Tatsache einfach dadurch zum Ausdruck, daß wir sagen, ein linear geordnetes Ringoid (bzw. ein linear geordnetes Divisionsringoid) ist eine rundungsinvariante Struktur bezüglich einer optimalen, symmetrischen Rundung.

<u>Satz 1.9:</u>

Ist $(R,N,+,\cdot,/,\leq)$ ein vollständiges, durchweg isoton geordnetes Divisionsringoid mit $(-e,0,e)$ und $(\mathcal{R}R,\leq)$ ein (symmetrisches) Raster von $(R,\leq)$ und $-e,0,e \in \mathcal{R}R$ und $\square : R \longrightarrow \mathcal{R}R$ eine optimale Rundung von R nach $\mathcal{R}R$. Sind ferner in $\mathcal{R}R$ die inneren Verknüpfungen $\boxplus, \boxdot, \boxslash$ erklärt vermöge $a \;\boxed{\circ}\; b := \square(a \circ b)$ für $\circ \in \{+,\cdot,/\}$, dann ist $(\mathcal{R}R,\tilde{N}, \boxplus, \boxdot, \boxslash, \leq)$ mit $\tilde{N} := N \cap \mathcal{R}R$ ein vollständiges, durchweg isoton geordnetes, optimales Rasterdivisionsringoid.

<u>Bew.:</u> Der Beweis zum Nachweis der Eigenschaften (D1),...,(D9) verläuft genauso wie für Satz 1.8. ($\mathcal{R}R$, $\boxplus$, $\boxdot$, $\boxslash$) ist also mindestens ein Rastersemidivisionsringoid.

Die Optimalität des Rastersemidivisionsringoid ergibt sich dadurch, daß die entsprechenden Eigenschaften (RG1) (RG2) und (RG3) sich aus (R1) (R2) und (R3) der optimalen zugrundegelegten Rundung $\square$ ergeben.

(OD6) ergibt sich wie folgt: Für beliebiges $o \in \{+, \cdot, /\}$ und $a,b,c \in \mathcal{R}R$ und mit $a \leq b$ ist stets

$a \boxed{o}\, c = \square(a \circ c) \leq \square(b \circ c) = b \boxed{o}\, c$, also gilt (OD6') und damit (OD6).

<u>Bemerkung:</u> Ein durchweg isoton geordnetes Divisionsringoid ist eine rundungsinvariante Struktur bezüglich einer optimalen Rundung.

<u>Definition 1.10:</u> Eine geordnete Menge $(M,\leq)$ heißt ein Verband, wenn gilt:

$$\bigwedge_{a,b \in M} \bigvee_{c \in M} \quad c := a \sqcap b := \inf\{a,b\} \quad \text{und}$$

$$\bigwedge_{a,b \in M} \bigvee_{c \in M} \quad c := a \sqcup b := \sup\{a,b\}.$$

Eine geordnete Menge $(M,\leq)$ heißt "bedingt vollständig", wenn jede nichtleere, beschränkte Teilmenge $T \subseteq M$ eine untere und eine obere Grenze besitzt.

Eine geordnete Menge $(M,\leq)$ heißt "vollständig" (oder ein Vollverband), wenn gilt

(O6) Jede nichtleere Teilmenge $T \subseteq M$ besitzt eine untere und eine obere Grenze.

In einem Verband definieren die Symbole $\sqcap$, $\sqcup$ zweistellige innere Verknüpfungen. Hierfür gelten die Regeln:

$$\left.\begin{array}{ll} \text{(L11)} & \bigwedge\limits_{a,b \in M} a \sqcap b = b \sqcap a \\[2em] \text{(L12)} & \bigwedge\limits_{a,b \in M} a \sqcup b = b \sqcup c \end{array}\right\} \text{kommutativ}$$

$$\left.\begin{array}{ll} \text{(L21)} & \bigwedge\limits_{a,b,c \in M} a \sqcap (b \sqcap c) = (a \sqcap b) \sqcap c \\[2em] \text{(L22)} & \bigwedge\limits_{a,b,c \in M} a \sqcup (b \sqcup c) = (a \sqcup b) \sqcup c \end{array}\right\} \text{assoziativ}$$

$$\text{(L31)} \quad \bigwedge_{a,b \in M} a \sqcap (a \sqcup b) = a$$

$$\text{(L32)} \quad \bigwedge_{a,b \in M} a \sqcup (a \sqcap b) = a$$

(L31), (L32): Absorptionsgesetz

In linear geordneten Verbänden gilt ein Distributivgesetz:

$$\text{(L4)} \quad \bigwedge_{a,b,c \in M} (a \sqcup b) \sqcap c = (a \sqcap c) \sqcup (b \sqcap c)$$

Zwischen der Ordnungsrelation und den Verknüpfungen eines Verbandes gelten die folgenden Beziehungen:

$$\text{(a)} \quad \bigwedge_{a,b \in M} (a \leq b \Longleftrightarrow a \sqcap b = a \Longleftrightarrow a \sqcup b = b)$$

$$\text{(b)} \quad \bigwedge_{a,b \in M} (a \leq b \Longrightarrow (\bigwedge_{c \in M} a \sqcap c \leq b \sqcap c \wedge a \sqcup c \leq b \sqcup c)).$$

<u>Satz 1.11:</u>

In einem linear geordneten Divisionsringoid $(R,+,-,*,/,\sqcap,\sqcup,\leq)$

gelten folgende Distributivgesetze:

(a) $\bigwedge\limits_{a,b,c \in R} \quad a+(b \sqcup c) = (a+b) \sqcup (a+c)$

(b) $\bigwedge\limits_{\substack{a,b,c \in R \\ a \geq o}} \quad a * (b \sqcup c) = (a * b) \sqcup (a * c)$

(c) $\bigwedge\limits_{a,b \in R_-} \quad x * (a \sqcup b) = (x * a) \sqcap (x * b)$

(d) $\bigwedge\limits_{a,b \in R \setminus N} \quad \begin{array}{l} a * b > o \Longrightarrow e/(a \sqcup b) = e/a \sqcap e/b \, , \\ a * b < o \Longrightarrow e/(a \sqcup b) = e/a \sqcup e/b \end{array}$

Sämtliche Aussagen gelten natürlich auch für die duale Aussage durch vertauschen von $\sqcap$ mit $\sqcup$.

(e) $\bigwedge\limits_{a,b,c,d \in R} \quad (a+b) \sqcup (c+d) \leq (a \sqcup c) + (b \sqcup d)$

Bew.:

Für (a) und (b) folgt die Behauptung aus (OD1), (OD2) a).

(c) Sei o.B.d.A. $a \sqcup b = a$, dann ist nach Lemma 1.12 $a \geq b$ und nach (OD3) $x * a \leq x * b$, also ist $x * a \sqcap x * b = x * a = x * (a \sqcup b)$

(d) Sei o.B.d.A. wieder $a \sqcup b = a$, also $a \geq b$ und sei

α) $b > o$, so folgt aus (OD2) b):

$e/a < e/b \implies e/a \sqcap e/b = e/a = e/(a \sqcup b)$

β) $a > o > b$, so folgt

$e/a > e/b \implies e/a \sqcup e/b = e/a = e/(a \sqcup b)$

In entsprechender Weise verifizieren sich die anderen Fälle.

(e) Aus $a \leq a \sqcup c \quad \wedge \quad b \leq b \sqcup d$ folgt $a+b \leq (a \sqcup c) + (b \sqcup d)$ und aus $c \leq a \sqcup c \quad \wedge \quad d \leq b \sqcup d$ folgt $c+d \leq (a \sqcup c) + (b \sqcup d)$ womit die Behauptung folgt.

<u>Satz 1.12:</u>

In einem durchweg isoton geordneten Divisionsringoid $(R, N, +, -, *, /, \sqcap, \sqcup, \leq)$ wobei $(R, \leq)$ linear geordnet ist, gilt für jede Verknüpfung $\circ \in \{+, -, *, /\}$

$$\bigwedge_{a,b,c \in R} \quad \begin{aligned} a \circ (b \sqcup c) &= (a \circ b) \sqcup (a \circ c) \\ a \circ (b \sqcap c) &= (a \circ b) \sqcap (a \circ c) \end{aligned}$$

bei Division
$c \in R \setminus N$

<u>Bew.:</u>

Sei o.B.d.A. $b \sqcup c = b$, so ist $b \geq c$ und damit nach (OD6):

$a \circ b \geq a \circ c \implies (a \circ b) \sqcup (a \circ c) = a \circ b = a \circ (b \sqcup c)$.

Definition 1.13:

Eine Menge M heißt ein *metrischer Raum* (M,ρ), wenn ein Funktional $\rho : M \times M \longrightarrow \mathbb{R}^{+}$ eine *Metrik* ist, d. h. ρ genügt folgenden Eigenschaften:

(M1) $$\bigwedge_{x,y \in M} \rho(x,y) = 0 \Longleftrightarrow x = y$$

(M2) $$\bigwedge_{x,y \in M} \rho(x,y) = \rho(y,x)$$

(M3) $$\bigwedge_{x,y,z \in M} \rho(x,z) \leq \rho(x,y) + \rho(y,z)$$

Ist $(M,\leq)$ geordnet, so heißt ρ verträglich mit $\leq$, wenn

(M4) $$\bigwedge_{x,y,z \in M} x \leq y \leq z \Longrightarrow \rho(x,y) \leq \rho(x,z)$$

2. Grundbegriffe der Intervallrechnung

In diesem Abschnitt sollen nur soviele Eigenschaften der Intervallrechnung aufgeführt werden, wie zum Verständnis des Abschnittes III notwendig ist. Genaueres findet man in [Kulisch 69] [Kulisch 72], [Mayer 68], [Moore 69] und [Herzberger 69].

Definition 2.1:

Ist $\mathbb{R}$ die Menge der reellen Zahlen, so ist $I\mathbb{R}$ die Menge der abgeschlossenen Intervalle $[a,b]$ mit den Intervallgrenzen $a,b \in \mathbb{R}$ und $a \leq b$. Insbesondere ist $\mathbb{R} \subseteq I\mathbb{R}$ mit der Vereinbarung, daß $[a,a]=a \in \mathbb{R}$.

Definition 2.2:

Zwei Intervalle $A=[a_1,a_2]$ und $B=[b_1,b_2]$ heißen gleich, in Zeichen $A=B$, genau dann, wenn $a_1=b_1$ und $a_2=b_2$ ist. Diese Relation in $I\mathbb{R}$ ist die Fortsetzung der Äquivalenzrelation = aus $\mathbb{R}$.

Definition 2.3:

$I\mathbb{R}$ ist Teilmenge der Potenzmenge $P\mathbb{R}$ und trägt damit dieselbe Halbordnung wie $P\mathbb{R}$ bezüglich der Inklusion $\subseteq$.

Es gilt:

$$\bigwedge_{\substack{[a,b] \in I\mathbb{R} \\ [c,d] \in I\mathbb{R}}} [a,b] \subseteq [c,d] \iff c \leq a \quad \wedge \quad b \leq d$$

Definition 2.4:

$$h(M) := \bigcap_{M \subseteq A \in \mathbb{IR}} A$$ ist die Hülle einer beliebigen Menge $M \in P\mathbb{R}$.

Sie ist das kleinste Intervall, das M gerade noch umschließt, genaueres siehe [Herzberger 69].

Definition 2.5:

Die algebraischen Verknüpfungen $\{+,-,*,/\}$ der reellen Zahlen $\mathbb{R}$ werden auf ganz $\mathbb{IR}$ fortgesetzt gemäß:

$$\bigwedge_{A,B \in \mathbb{IR}} \quad \bigwedge_{\circ \in \{+,-,*,/\}} \quad A \circ B := h(\{a \circ b \mid a \in A \wedge b \in B\})$$

wobei im Falle der Division natürlich $0 \notin B$ sein muß.

Die Hüllenbildung erübrigt sich, da $+,-,*$ und $/$ stetige Operationen (mit Ausnahme der Division für Intervalle $A \ni 0$) und Intervalle einfach zusammenhängende Mengen sind.

Diese Verknüpfungen führen also nicht aus $\mathbb{IR}$ hinaus, jedoch übertragen sich nicht alle Körpereigenschaften der reellen Zahlen $\mathbb{R}$ auf den Intervallraum $\mathbb{IR}$. Eine Zusammenstellung der verbleibenden algebraischen Struktur und weitere wichtige Eigenschaften findet man in [Kulisch 72] und hier in Abschnitt III.

Eine wichtige Folgerung aus Definition 2.5 ist

Satz 2.6:

Die oben angegebenen Intervallverknüpfungen lassen sich darstellen als Verknüpfungen der Intervallgrenzen wie folgt:

(1) $[a,b]+[c,d]:=[a+c,b+d]$

(2) $-[a,b]:=[-b,-a]$ und $A-B:=A+(-B)$

(3) $[a,b]*[c,d]:=[\inf(P),\sup(P)]$ mit $P:=\{ac,ad,bc,bd\}$

(4) $1/[a,b]:=[1/b,1/a]$ und $A/B:=A*(1/B)$ sofern $o \notin B$

Zum Beweis, siehe [Kulisch 72].

Definition 2.7:

Folgende spezielle Funktionen $I\mathbb{R} \longrightarrow \mathbb{R}$ werden später benutzt:

a) $\lambda A = \lambda(A) := \lambda([a,b]) := a$ linke Grenze von A

$\rho A = \rho(A) := \rho([a,b]) := b$ rechte Grenze von A

b) $d(A) = d([a,b]) := b-a$ Durchmesser des Intervalles A

$m(A) = m([a,b]) := (a+b)/2$ Mittelpunkt des Intervalles A

3. Abschätzungssätze für Fixpunkte bei Verwendung rasterrestringierter Abbildungen

Gegeben sei ein vollständiger metrischer Raum (M,ρ) und linear geordneter Vollverband $(M,\leq)$ und ein Raster $(\mathcal{R}M,\leq)$ von M. Ferner sei $f:M \to M$ eine bezüglich ρ kontrahierende Abbildung, d.h. f genüge in M der Lipschitzbedingung:

(1) $$\bigwedge_{x,y \in M} \rho(f(x),f(y)) \leq L\rho(x,y) \quad \wedge \quad L < 1 \ .$$

Das Iterationsverfahren

(2) $x_{i+1}:=f(x_i)$ für $i \geq o$ und beliebiges $x_o \in M$

konvergiere gegen den eindeutig bestimmten Fixpunkt $\hat{x}$ der Gleichung $x=f(x)$.

Betrachtet man nun ein zu (2) gehöriges (abgeändertes) Iterationsverfahren

(3) $y_{i+1}:=\alpha f(y_i)$ für $i \geq o$ und $y_o \in \mathcal{R}M \setminus \{-p,p\}$ [1)]

wobei $\alpha f : \mathcal{R}M \to \mathcal{R}M$ eine Rasterfunktion bezüglich $\mathcal{R}M$ ist (Def. 1.5), so ist i.a. αf nicht einmal stetig in M , so daß die Forderung einer Lipschitzbedingung der Gestalt (1) nicht möglich ist.

---

1) Betreff $\{-p,p\}$ siehe Abschnitt 4. Eine Aussage, in der $\mathcal{R}M \setminus \{-p,p\}$ auftritt, besagt u.a., daß der Überlauf ausgeschlossen sein soll!

Die nach (3) gebildete Iterationsfolge $\{y_i\}_{i \in \mathbb{N}_o}$ braucht daher nicht zu konvergieren und kann i.a. mehrere Häufungspunkte besitzen.

Für viele Probleme, u.a. bei solchen der Intervallrechnung, ist es von Wichtigkeit, zu wissen, ob bei gerichteten rasterrestringierten Abbildungen αf in entsprechender Richtung ein Häufungspunkt von (3) oberhalb bzw. unterhalb von $\hat{x}$ von (2) liegt, so daß mit diesen eine obere bzw. untere Abschätzung für $\hat{x}$ gegeben ist.

<u>Satz 3.1:</u>

<u>Voraussetzungen:</u>

1. Sei $(M,\rho,\leq)$ ein metrisch vollständiger linear geordneter Vollverband und Metrik ρ genüge der Eigenschaft (M4).

2. Sei $f:M \to M$ eine der Lipschitzbedingung (1) genügende Abbildung und sei für beliebiges $x_o \in M$ $\hat{x} \in M$ Fixpunkt von $x = f(x)$ und Grenzwert der Iterationsfolge $\{x_{i+1} = f(x_i)\}_{i \in \mathbb{N}_o}$.

3. Es sei $\mathcal{R}M$ endliches diskretes Raster von M und αf eine Rasterfunktion von f auf $\mathcal{R}M$ mit $\alpha f:\mathcal{R}M \to \mathcal{R}M$ und die Restriktion α nach oben gerichtet, d.h. es gelte

$$\text{(A3)} \quad \bigwedge_{x \in \mathcal{R}M} f(x) \leq \alpha f(x) \quad .$$

4. Die abgeänderte Iterationsfolge $\{y_{i+1} = \alpha f(y_i)\}_{i \in \mathbb{N}_0}$ sei für beliebiges $y_0 \in \mathcal{R} M \setminus \{-p,p\}$ so beschaffen, daß sie nach endlich vielen Schritten $N \leq \text{card}(\mathcal{R} M) - 2$ zyklisch in einer Zyklusmenge $Z = Z(y_0) \subseteq \mathcal{R} M \setminus \{-p,p\}$ endet. Für die Zyklusmenge gilt:

$$\bigwedge_{z \in Z(y_0)} \left\{ \bigvee_{k \in \mathbb{N}} z = (\alpha f)^{(k)}(y_0) \wedge \bigwedge_{p \in \mathbb{N}} (\alpha f)^{(p)}(z) \in Z \right\}, \text{ wobei}$$

$(\alpha f)^{(1)} := \alpha f, \quad (\alpha f)^{(i+1)}(x) := (\alpha f)((\alpha f)^{(i)}(x))$ ist.

<u>Behauptung:</u> $\hat{x}$ und $y^* := \bigsqcup_{z \in Z} z$ stehen in entsprechender Ordnungsbeziehung:

$\hat{x} \leq y^*$, d. h. es existiert ein oberer Häufungspunkt $y^* \in Z \subseteq \mathcal{R} M$, der $\hat{x}$ nach oben abschätzt.

<u>Bew.:</u>

Da $(M, \leq)$ linear geordnet ist, kann nur gelten:

$\hat{x} \leq y^*$ oder aber $y^* < \hat{x}$.

$y^* < \hat{x}$ führt zu folgendem Widerspruch:

Sei $y^* < \hat{x}$, so gilt wegen den Voraussetzungen 1.,2.,3. und 4.:

$f(y^*) \leq \alpha f(y^*) \leq y^* < \hat{x}$ daraus folgt nach (M4):

$\rho(y^*, \hat{x}) \leq \rho(f(y^*), \hat{x}) =$

$$= \rho(f(y^*), f(\hat{x})) \leq L\rho(y^*, \hat{x}) \quad \text{mit} \quad L < 1,$$

also müßte $\rho(y^*, \hat{x}) < \rho(y^*, \hat{x})$ gelten, was für keinen Abstand $\neq 0$ zutreffen kann. Somit ist stets $\hat{x} \leq y^*$ wie behauptet.

Leider ist Satz 3.1 in dieser Form nicht auf Iterationsverfahren z. B. in Intervallräumen $(\mathbb{IR}, \subseteq)$ anwendbar, da $(\mathbb{IR}, \subseteq)$ nicht linear geordnet ist. Für den Fall jedoch, daß die zugrundeliegende Abbildung f ordnungsisoton ist, läßt sich ein zu Satz 3.1 äquivalenter Satz formulieren.

Zur Vereinfachung der Schreibweise $\mathcal{R}M \setminus \{-p,p\}$ gelte künftig $\mathcal{R}M' := \mathcal{R}M \setminus \{-p,p\}$.

<u>Satz 3.2:</u>

<u>Voraussetzungen:</u>

1. Sei $(M,\leq)$ ein Vollverband.

2. Sei $f:M \to M$ eine Abbildung,

   (2.1) die ordnungsisoton ist, d. h.

$$\bigwedge_{x,y \in M} x \leq y \implies f(x) \leq f(y) \text{ , und}$$

   (2.2) für die die Iterationsfolge $\{x_{i+1} := f(x_i)\}_{i \in N_0}$ für beliebiges $x_0 \in M$ gegen den eindeutigen Fixpunkt $\hat{x} \in M$ der Gleichung $x = f(x)$ konvergiere,

3. und 4. seien dieselben Voraussetzungen wie in Satz 3.1.

5. Sei $Z(y_o)$ die Zyklusmenge, in der die Folge $\{y_{i+1} = \alpha f(y_i)\}_{i \geq o}$ ausgehend von $y_o \in \mathcal{R}_M$, endet.

Behauptung:

a) Ist $y_o \in \mathcal{R}M$ beliebig, so gilt in jedem Falle

$$\hat{x} \leq \inf Z(y_o) = y_* \in \mathcal{R}M \quad .$$

b) Ist zusätzlich αf isoton, d.h. aus $x \leq y$ folgt $\alpha f(x) \leq \alpha f(y)$ und berechnet man ausgehend von y_o' und y_o'' mit $y_o' \leq y_o''$ die Zyklusmengen $Z(y_o')$ und $Z(y_o'')$ und sind die Kardinalzahlen von $Z(y_o')$ und $Z(y_o'')$ teilerfremd, so gilt:

$$\hat{x} \leq \inf Z(y_o') \leq \sup Z(y_o') \leq \inf Z(y_o'')$$

Bew.:

a) Aus Voraussetzung (2.1) folgt unmittelbar die Beziehung:

$$x_1 = f(y_o) \leq \alpha f(y_o) = y_1$$

und aus $x_i \leq y_i$ für $i \leq k$ folgt

$$x_{k+1} = f(x_k) \leq f(y_k) \leq \alpha f(y_k) = y_{k+1} \; .$$

Sei für $k \geq N$ $y_k \in Z(y_o)$, so gilt

$$\bigwedge_{z \in Z(y_o)} \bigvee_{k \geq N} x_k \leq z \quad .$$

Es existiert also zu beliebigem $z \in Z(y_o)$ eine Partialfolge $\{x_{k_j}^{(z)}\}_{j \geq o, k_j \geq N}$ [1] mit demselben Grenzwert $\hat{x}$, für die

stets $x_{k_j}^{(z)} \leq z$ für $j \geq o$ gilt.

Somit ist für beliebiges $z \in Z(y_o)$ $\hat{x} \leq z$ und daher

$$\hat{x} \leq \inf Z(y_o) \quad .$$

b) Aus $y_o' \leq y_o''$ folgt nach Voraussetzung

$y_1' = \alpha f(y_o') \leq \alpha f(y_o'') = y_1''$ und allgemein

folgt aus $y_i' \leq y_i''$ für $i \leq k$ stets:

---

1) Ist $\operatorname{card}(Z(y_o)) = p$, so ist $x_{k_j}^{(z)} := x_{k(z) + p \cdot j}$.

$y'_{k+1} = \alpha f(y'_k) \leq \alpha f(y''_k) = y''_{k+1}$, also nach vollständiger Induktion:

$$\bigwedge_{k \geq o} y'_k \leq y''_k \quad .$$

Sei für $k \geq N$ sowohl $y'_k \in Z(y'_o)$ als auch $y''_k \in Z(y''_o)$ und sei $Z(y'_o) = \{\eta'_1, \eta'_2, \ldots, \eta'_m\}$

und $Z(y''_o) = \{\eta''_1, \eta''_2, \ldots, \eta''_n\}$, so sind nach Voraussetzung n und m teilerfremd. Es gilt somit (dies erweist sich nach spätestens $N + n \cdot m$ Iterationsschritten):

$$\bigwedge_{o \leq i \leq m} \quad \bigwedge_{o \leq j \leq n} \eta'_i \leq \eta''_j, \quad \text{d. h.}$$

$Z(y'_o)$ sind untere Schranken für $Z(y''_o)$ und $Z(y''_o)$ sind obere Schranken für $Z(y'_o)$. Es gilt daher

$$\sup Z(y'_o) \leq \inf Z(y''_o)$$

und nach a) ist $\hat{x} \leq \inf Z(y'_o)$, so daß schließlich die Behauptung folgt.

In der Praxis ist es i.a. nicht berechenbar, wann ein Zyklus auftritt bzw. wie umfangreich die Zyklusmenge wird. Deshalb ist die Anwendung der obigen Sätze i.a. nur theoretischer Art.

Der folgende Satz gibt Verfahren an, die ohne Kenntnis einer Zyklusmenge eine obere Abschätzung des exakten Fixpunktes $\hat{x}$ berechnen.

Satz 3.3:

Voraussetzungen:

1. und 2. seien dieselben Voraussetzungen wie in Satz 3.2 .

3. αf erfülle über dem endlichen Raster $\mathcal{R}M$ die Bedingung 3. aus Satz 3.2 und es sei zusätzlich αf ordnungsisoton, d.h. es gilt:

$$\bigwedge_{x,y \in \mathcal{R}M} x \leq y \Longrightarrow \alpha f(x) \leq \alpha f(y) \quad .$$

4. Die Folge $\{y_{i+1} = \alpha f(y_i)\}_{i \geq o}$, die den folgenden Aussagen zugrunde liegt, erfülle die Bedingung 4. aus Satz 3.1. $Z(\xi) \subseteq \mathcal{R}M'$ bezeichne die aus dem Startelement $\xi \in \mathcal{R}M'$ entstehende Zyklusmenge dieser Folge.

Behauptung:

a) Das Iterationsverfahren

$$v_{i+1} = \alpha f(v_i) \sqcap v_i, \quad i \geq o \quad \text{und} \quad \hat{x} \leq v_o \in \mathcal{R}M'$$

konvergiert gegen einen von v_o abhängigen Fixpunkt $\hat{v}$ und es gilt: $\hat{x} \leq \hat{v} \leq \inf Z(v_o)$.

b) Das Iterationsverfahren

$$u_{i+1} = \alpha f(u_i) \sqcup u_i, \quad i \geq o \quad \text{und beliebigem} \quad u_o \in \mathcal{R}M'$$

konvergiert, falls die Folge $\{u_i\}_{i \geq o}$ nach oben beschränkt ist, gegen einen von u_o abhängigen Fixpunkt $\hat{u}$ und es gilt

$$x \leq \inf Z(u_o) \leq \sup Z(u_o) \leq \hat{u}$$

c) Sei $\{v_i\}_{i \geq o}$ die Folge des in a) beschriebenen Iterationsverfahrens mit Fixpunkt $\hat{v}$, so konvergiert die Nachiteration

$w_{i+1} = (\alpha f(w_i) \sqcup w_i) \sqcap \hat{v}$, $i \geq o$ und $w_o \leq \hat{v} \in \mathcal{R}M$

gegen den von w_o abhängigen Fixpunkt $\hat{w}$ und es gilt:

$$\hat{x} \leq \hat{w} \leq \hat{v} \quad .$$

<u>Bew.:</u>

a) Aus $\hat{x} \leq v_o$ folgt nach Voraussetzung 2. und 3.

$\hat{x} = f(\hat{x}) \leq f(v_o) \leq \alpha f(v_o) \wedge \hat{x} \leq v_o \quad \Longrightarrow$

$\hat{x} \leq \alpha f(v_o) \sqcap v_o = v_1 \leq v_o$.

Allgemein folgt aus $\hat{x} \leq v_i \leq v_o$ für $i \leq k$:

$\hat{x} = f(\hat{x}) \leq f(v_k) \leq \alpha f(v_k) \wedge \hat{x} \leq v_i \quad \Longrightarrow$

$\hat{x} \leq \alpha f(v_k) \sqcap v_k = v_{k+1} \leq v_k \leq v_o$

durch vollständige Induktion

$\hat{x} \leq v_{k+1} \leq v_k \leq v_o$ für $k \geq o$.

Die Folge $\{v_k\}_{k \geq o}$ ist monoton fallend und beschränkt und konvergiert daher gegen den Grenzwert v^* und es gilt

$\hat{x} \leq v^* = \lim\limits_{k\to\infty} v_k$.

Da $\mathcal{R}M$ endlich ist, gilt sogar nach endlich vielen Schritten N, daß für $i \geq N$ $v_{i+1} = \alpha f(v_i) \sqcap v_i = v^*$, d.h. es ist $v^* = \alpha f(v^*) \sqcap v^* = \hat{v}$ zugleich Fixpunkt.

Ferner folgt aus

$$v_1 = \alpha f(v_o) \sqcap v_o \leq \alpha f(v_o) = y_1 \quad \text{und}$$

aus

$$v_i \leq y_i \qquad \text{für} \quad i \leq k$$

$$v_{k+1} \leq \alpha f(v_k) \sqcap v_k \leq \alpha f(v_k) = y_{k+1}, \qquad \text{d. h.}$$

$$\bigwedge_{k \geq o} v_k \leq y_k \, .$$

Sei für $k \geq N$

$v_k = \hat{v}$ und $y_k \in Z(v_o)$, so ist $\bigwedge_{z \in Z(v_o)} \hat{v} \leq z$,

d. h. $\hat{v} \leq \inf Z(v_o)$.

b) Aus

$$y_1 = \alpha f(u_o) \leq \alpha f(u_o) \sqcup u_o = u_1$$

und allgemein aus

$$y_i \leq u_i \quad \text{für} \quad i \leq k \quad \text{folgt}$$

$$y_{k+1} = \alpha f(y_k) \leq \alpha f(u_k) \leq \alpha f(u_k) \sqcup u_k = u_{k+1}$$

und durch vollständige Induktion

$$y_k \leq u_k \quad \text{für} \quad k \geq o \, .$$

Sei für $k \geq N$ $y_k \in Z(u_o)$ und $u_k = \hat{u}$ (da u_k monoton wachsend und beschränkt, also konvergent), so gilt:

$$\bigwedge_{z \in Z(u_o)} \quad z \leq \hat{u} \implies \sup Z(u_o) \leq \hat{u} \; .$$

Aus Satz 3.2 a) folgt daher $\hat{x} \leq \inf Z(u_o) \leq \sup Z(u_o) \leq \hat{u}$.

c) Mit $w_o \leq \hat{v}$ ist die Folge $\{w_i\}_{i \geq o}$ monoton wachsend und nach oben beschränkt, denn aus

$$w_i \leq \hat{v} \quad \text{folgt}$$

$\alpha f(w_i) \sqcup w_i \geq w_i$ d. h.

$w_i = w_i \sqcap \hat{v} \leq (\alpha f(w_i) \sqcup w_i) \sqcap \hat{v} = w_{i+1} \leq \hat{v}$.

$\{w_i\}_{i \geq o}$ ist daher konvergent gegen ein $w^* \leq \hat{v}$.

Da $\mathfrak{R}M'$ endlich folgt wieder $w^* = \hat{w} \leq \hat{v}$.

Nach b) aber ist $\hat{x} \leq \hat{w}$, womit die Behauptung $\hat{x} \leq \hat{w} \leq \hat{v}$ bewiesen ist.

Bemerkung 3.4:

1. Satz 3.3 ist unmittelbar auf die Intervallrechnung anwendbar.

   Die Aussage unter Punkt a) ist nichts Neues, denn dies ist die übliche Einschließungsmethode der Intervallrechnung.

   Neu in diesem Zusammenhang dürften Punkt b) und c) sein.

   Die Aussage von b) könnte einen Algorithmus liefern zur Berechnung einer Grobeinschließung, die für a) notwendige

Voraussetzung ist. Problem in b) ist, die Beschränktheit der Folge $\{u_i\}_{i \geq o}$ zu sichern. Die Folge $\{u_i\}_{i \geq o}$ ist für genügend feine Raster $\mathcal{R}M$ sicher beschränkt, wenn die Lipchitzkonstante von f kleiner 1 ist. Dies folgt aus der Tatsache, daß in einem Vollverband und Banachraum $(M,*)$ mit Metrik ρ und Eigenschaft

$$(MV) \qquad \bigwedge_{a,b,c \in M} \rho(a \sqcup b,c) \leq \rho(a,c) \sqcup \rho(b,c)$$

für eine Iterationsfunktion f mit $L < 1$ und

$$\bigwedge_{x,y \in M} \rho(f(x),f(y)) \leq L\rho(x,y) \quad \text{gilt:}$$

Die Folge

$\xi_{i+1} = f(\xi_i) \sqcup \xi_i$ mit $\xi_o \in M$ ist beschränkt im Kreis mit Radius $\rho(\hat{x},\xi_o)$ und Mittelpunkt $\hat{x}$, wenn $\hat{x}$ den Fixpunkt der Gleichung $x = f(x)$ darstellt.

Dies folgt aus:

$$\rho(\hat{x},\xi_{i+1}) = \rho(f(\hat{x}), f(\xi_i) \sqcup \xi_i) \overset{(MV)}{\leq}$$

$$\leq \rho(f(\hat{x}), f(\xi_i)) \sqcup \rho(\hat{x},\xi_i) \leq$$

$$\leq L\rho(\hat{x},\xi_i) \sqcup \rho(\hat{x},\xi_i) = (L \sqcup 1)\rho(\hat{x},\xi_i) =$$

$$= \rho(\hat{x},\xi_i)$$

d. h. für $i \geq o$ ist $\rho(\hat{x},\xi_i) \leq \rho(\hat{x},\xi_o)$.

Die Aussage unter c) liefert offenbar ein Verfahren, durch Nachiterationen eine gegebene Einschließung $\hat{v}$ eines Fixpunktes $\hat{x}$ zu verbessern. Die Variationsmöglichkeiten werden in dieser Arbeit nicht voll ausdiskutiert. Abschnitt III,6. beschränkt sich in der Anwendung auf den Fall $w_o \leq \hat{x} \leq \hat{v}$.

2. Sämtliche Überlegungen gelten aus Symmetriegründen natürlich auch für entsprechend nach unten gerichtete Ordnungsbeziehungen.

3. Die Metrik $q(A,B) = \|A - \overline{B}\|$ aus Abschnitt III,4. genügt der oben geforderten Eigenschaft (MV), denn es ist

$$q(A \sqcup B,C) = \|A \sqcup B - \overline{C}\| = \text{nach III,2.44,(1):}$$

$$= \|A - \overline{C} \sqcup B - \overline{C}\| \leq \text{nach III,4.4(e)}$$

$$\leq \|A - \overline{C}\| \sqcup \|B - \overline{C}\| = q(A,C) \sqcup \rho(B,C).$$

Die Operation $\sqcup$ ist in III,2.6 definiert. Sie ist die Supremumsoperation bezüglich $\subseteq$.

4. Das Problem des endlichen Rasters (auf Rechenanlagen)

Da die derzeit üblichen Rechenmaschinen stets mit endlichen Zahlenbereichen rechnen, wäre es unrealistisch, das Problem des endlichen Rasters zu ignorieren.

Der Usus der meisten Rechenanlagen, in Situationen den Rechenprozeß einfach abzubrechen, wenn z. B. $a \boxed{/} o$ auftritt oder $a \boxed{+} b$ den Maschinenzahlenbereich überschreitet, ist (vergleiche Abschnitt III) unter Umständen sogar störend und führt zu umständlichen Programmen.

Sicher gerechtfertigt ist, bei Bereichsüberschreitung eine Warnung auszugeben, da ab diesem Zeitpunkt bei ''falschen Programmen'' mit unendlichen Rundungsfehlern gerechnet wird. Der Rechenprozeß muß deshalb so fortgesetzt werden, daß die beim gerundeten Rechnen übliche Struktur erhalten bleibt, denn dann kann in ''richtigen Programmen'' sinnvoll weitergerechnet werden. Die Lösung dieses Problems kann in den folgenden Abschnitten nur zum Teil gegeben werden.

4.1 Vervollständigung der reellen Zahlen $\mathbb{R}$ und eine Klasse von endlichen Rastern in $\mathbb{R}$

Nach Satz 1.2 bzw. nach [Kulisch 72] muß notwendig $(\mathbb{R},\leq)$ zu einem vollständigen Verband gemacht werden, damit ein Teilverband $T \subseteq \mathbb{R}$ zu einem Raster wird. Das läßt sich leicht realisieren durch Einführung der uneigentlichen Zahlen

$-\infty$ und $+\infty$. Dann ist $\tilde{\mathbb{R}} := \mathbb{R} \cup \{-\infty,+\infty\}$ ein vollständiger Verband. Diese beiden Symbole sind bereits mit bestimmten Bedeutungen und Vorstellungen verknüpft, so daß es auch im Hinblick auf die Realisierung auf Rechenanlagen besser ist, andere Bezeichnungen zu wählen.

Da diese beiden Randzahlen auf Rechenanlagen wie eine Art Puffer bei Bereichsüberschreitung wirken, sollen sie mit $-p$ und $+p$ gekennzeichnet werden.

Definition und Satz 4.1:

Die Menge $\mathbb{R}^* := \mathbb{R} \cup \{-p,p\}$ ist mit der Festsetzung:

$$\bigwedge_{a \in \mathbb{R}^*} -p \le a \le +p \qquad \text{ein vollständiger Verband } (\mathbb{R}^*,\le).$$

Nun ist aber $(\mathbb{R}^*,\{o\},+,\cdot,/,\le)$ kein Körper mehr, denn es sind $p-p$ noch $o\cdot p$ usw. nicht ohne Widerspruch definierbar. Das bringt aber für unsere Absichten keinen Schaden, da zur Aufstellung eines Rasterringoides nach Satz 1.8 $(\mathbb{R}^*,\{o\},+,\cdot,/,\le)$ nur ein vollständiges, linear geordnetes Divisions-Ringoid zu sein braucht.

Satz 4.2:

$(\mathbb{R}^*,\{o\},+,\cdot,/,\le)$ ist gemäß folgenden Verknüpfungstabellen Tabelle 4.1, Tabelle 4.2 und Tabelle 4.3 ein vollständiges, linear geordnetes (Divisions-)Ringoid mit den ausgezeichneten Elementen $\{-p,-1,o,1,p\}$.

+	$b \in \mathbb{R}$	+p	-p
$a \in \mathbb{R}$	a+b	+p	-p
+p	+p	+p	o
-p	-p	o	-p

Tabelle 4.1

$*$	$b \in \mathbb{R}$	-p	-1	o	1	p
$a \in \mathbb{R}$	ab	-psign(a)	-a	o	a	psign(a)
-p	-psign(b)	p	p	o	-p	-p
-1	-b	p	1	o	-1	-p
o	o	o	o	o	o	o
1	b	-p	-1	o	1	p
p	psign(b)	-p	-p	o	p	p

Tabelle 4.2

/	$b \in \mathbb{R} \wedge$ $b \neq o$	-p	o	p
$a \in \mathbb{R} \wedge$ $a \neq o$	a/b	o		o
-p	-psign(b)	e		-e
o	o	o		o
+p	-p·sign(b)	-e		e

Tabelle 4.3

<u>Bew.:</u> Es werden die Eigenschaften der Definition 1.6 nachgeprüft:

(D1) Tabelle 4.1 ist symmetrisch bezüglich der Hauptdiagonalen, also kommutativ

(D2) gilt in $\mathbb{R}$ und darüberhinaus folgt für a=o:

o+p=p o+(-p)=-p

(D3) ist nach Tabelle 4.2 außer in $\mathbb{R}$ wie die 1-Zeile und 1-Spalte zeigen auch für p und -p erfüllt.

(D4) Wie die o-Zeile und o-Spalte zeigt gilt auch diese Eigenschaft.

(D5) x=-1 erfüllt a) xx=1

b) ist in $\mathbb{R}$ erfüllt und ferner ist:

-p=(-1)(ap)=(-a)p=-p=p(-a)=-p

p=(-1)(a(-p))=(-a)(-p)=p=(-p)(-a)=p

-p=(-1)(pp)=(-p)p=p(-p)=-p

p=(-1)((-p)p)=pp=(-p)(-p)=p usw.

c) In $\mathbb{R}$ gilt diese Vorzeichenregel und laut Tabelle 4.1 und 4.2 ist:

-p=(-1)(a+p)=-a-p=-p

p=(-1)(a-p)=-a+p=p

-p=(-1)(p+p)=-p-p=-p

p=(-1)(-p-p)=p+p=p

o=(-1)(p-p)=-p+p=o

(D6) x ist eindeutig bestimmt, das gilt ja in $\mathbb{R}$ und x=-p verletzt a).

Mit $N:=\{o\}$ folgt für die Divisionsringoideigenschaften:

(D7) Mit $b=1$ zeigt die erste Spalte der Tabelle 4.3, daß

$$a/1=1 \quad \text{und} \quad -p/1=-p \quad \text{und} \quad p/1=p$$

(D8) Wie die o-Zeile zeigt, ist stets $o/c=o$ für alle $c \in \mathbb{R}^*$

(D9) läßt sich analog wie in (D5) nachweisen.

Zum Nachweis, daß $\mathbb{R}^*$ ein geordnetes Ringoid ist, werden die Eigenschaften aus Definition 1.9 nachgeprüft:

(OD1) ist in $\mathbb{R}$ schon erfüllt und für die restlichen Elemente folgt im Vergleich:

$$c \neq p \quad \wedge \quad c \neq -p$$
$$-p \leq a \leq p \implies -p=-p+c \leq a+c \leq p+c=p$$

$$c=p$$
$$-p \leq a \leq p \implies o=-p+p \leq a+p=p \leq p+p=p$$

$$c=-p$$
$$-p \leq a \leq p \implies -p=-p-p \leq a-p=-p \leq p-p=o$$

(OD2) ist ebenfalls in $\mathbb{R}$ erfüllt und darüber hinaus ist:

$$c>o \quad \wedge \quad c \neq p$$
$$o \leq a \leq p \implies o=oc \leq ac \leq pc=p$$

$$c=p \qquad \implies o=op \leq ap=p \leq pp=p$$

(OD3) gilt in $\mathbb{R}$ und mit

$$-p < a < p \implies (-1)(-p)=p \implies (-1)a=-a \implies (-1)p=-p$$

Die Axiome (OD4a) und (OD4b) erfüllen sich wie folgt:

(OD4a) $o \leq a \leq b \quad \wedge \quad p > o \Longrightarrow o \leq a/p = o = b/p = o$

$o \leq a \leq p \quad \wedge \quad c > o \Longrightarrow o \leq a/c \leq p/c = p$

(OD4b) $o < a \leq b \quad \wedge \quad p > o \Longrightarrow p/a = p > p/b = p > o$

$o < a \leq p \quad \wedge \quad c > o \Longrightarrow c/a \geq c/p = o \geq o$

(OD5) gilt für alle Elemente $a \in \mathbb{R}^*$ mit Ausnahme von $a = o$.

Satz 4.3:

Jede Teilmenge $\mathcal{R}\mathbb{R}^* := \{\pm a_i \mid 1<i<N \wedge a_i \in \mathbb{R}^+\} \cup \{-p,-1,o,1,p\}$ von $\mathbb{R}^*$ ist ein symmetrisches Raster.

Bew.: Die Symmetrie liegt per Definition vor. $\mathbb{R}^*$ ist vollständiger Verband und es ist

(S1") $o(\mathcal{R}\mathbb{R}^*) = -p = o(\mathbb{R}^*) \quad \wedge \quad i(\mathcal{R}\mathbb{R}^*) = p = i(\mathbb{R}^*)$

(S2") $(\mathcal{R}\mathbb{R}^*, \leq)$ ist vollständiger Teilverband von $(\mathbb{R}^*, \leq)$

uns somit folgt die Behauptung nach Satz 1.2.

Satz 4.4:

Ist eine optimale, symmetrische Rundung $\square : \mathbb{R}^* \longrightarrow \mathcal{R}\mathbb{R}^*$ gegeben, so gilt nach Satz 1.8, daß mit

$$\text{(V)} \quad \bigwedge_{\circ \in \{+,-,*,/\}} \; \bigwedge_{a,b \in \mathcal{R}\mathbb{R}^*} \quad a \boxed{\circ} b := \square(a \circ b)$$

$(\mathcal{R}\mathbb{R}^*, \{0\}, \boxed{+}, \boxed{-}, \boxed{*}, \boxed{/}, \leq)$ ein vollständiges, linear geordnetes Raster-Divisionsringoid darstellt.

Die einzigen kritischen Situationen die das Weiterrechnen nach Berücksichtigung u.U. "sinnlos" machen, sind $0 * (\pm p) = 0$ und $p - p = 0$. In allen anderen Fällen bleiben die Ergebnisse "sinnvoll", d. h. das Ergebnis ist $+p$ oder $-p$.

KAPITEL II: PRÄMETRISCHE UND METRISCHE FELDER

1. Metrik, Norm und Prämetrik

1.1 Versagen der klassischen Metrik

Ist $\mathbb{R}$ die Menge der reellen Zahlen und $\mathcal{R}\mathbb{R} \dot{\subseteq} \mathbb{R}$ ein Raster der reellen Zahlen, so ist natürlich mit einer Metrik ρ über $\mathbb{R}$ sowohl $(\mathbb{R},\rho)$ als auch jede Teilmenge $(\mathcal{R}\mathbb{R},\rho)$ ein metrischer Raum. In der Praxis jedoch treten bei numerischen Problemen stets algebraische Operationen auf, etwa $+$ und $\cdot$ in $\mathbb{R}$, so daß bei Fehlerabschätzungen stets Abschätzungen von $\rho(a+x,a+y)$ oder $\rho(a\cdot x,a\cdot y)$ benötigt werden. Dazu werden meist gewisse Verträglichkeitseigenschaften der Metrik ρ mit diesen algebraischen Verknüpfungen (hier $+$ und $\cdot$) gefordert.

Ist $(M,+,\cdot)$ eine Menge mit zwei inneren Verknüpfungen und ist (M,ρ) linearer metrischer Raum über dem Skalarkörper $\mathbb{R}$, so heißt die Metrik $\rho: M \times M \to \mathbb{R}^+$ (in Fortsetzung von Definition 1.13).

(M5) (translationsinvariant)

$$\bigwedge_{a,x,y\in M} \rho(a+x,a+y) = \rho(x,y)$$

(M6) (supermetrisch)

$$\bigwedge_{u,v,x,y\in M} \rho(u+x,v+y) \leq \rho(u,v) + \rho(x,y)$$

(M7) (homogen bzw. subhomogen)

$$\bigwedge_{\substack{x,y\in M\\ \lambda\in\mathbb{R}}} \rho(\lambda x,\lambda y) \underset{\text{bzw.}\le}{=} |\lambda|\rho(x,y)$$

<u>Definition 1.1:</u>

Ist $(R,+,\cdot)$ Ringoid mit den ausgezeichneten Elementen $\{-e,o,e\}$ und ρ eine Metrik, also (R,ρ) ein metrischer Raum, so heißt $(R,\rho,+,\cdot)$ ein <u>metrisches Ringoid</u>, wenn gilt:

(MR1) $$\bigwedge_{a,x,y\in R} \rho(a+x,a+y) \le \rho(x,y)$$

(MR2) $$\bigwedge_{a,x,y\in R} \rho(a\cdot x,a\cdot y) \le \rho(a,o)\rho(x,y)$$

(MR3) $\rho(e,o) = \rho(-e,o) = 1$

<u>Beispiel:</u> Das Ringoid $(\mathbb{R},+,\cdot)$ der reellen Zahlen ist mit der üblichen homogenen, translationsinvarianten Metrik $\rho(a,b):=|a-b|$ ein metrisches Ringoid.

In metrischen Ringoiden gelten auf Grund der Eigenschaften (MR1),(MR2),(MR3),(D2),(D3),(D4) und (M3) folgende Beziehungen:

Satz 1.2:

(a) $\rho(x+y,o) \leq \rho(x,o) + \rho(y,o)$

(b) $\rho(a\cdot x,o) \leq \rho(a,o) \cdot \rho(x,o)$

(c) $\rho(-a,o) = \rho(a,o)$

Bew.: (a) $\rho(x+y,o) \overset{(M3)(D2)}{\leq} \rho(x+y,x+o) + \rho(x+o,o+o) \leq$

$\overset{(MR1)}{\leq} \rho(y,o) + \rho(x,o)$

(b) $\rho(a\cdot x,o) \overset{(D4)}{=} \rho(a\cdot x,a\cdot o) \overset{(MR2)}{\leq} \rho(a,o)\rho(x,o)$

(c) $\rho(-a,o) \overset{(D4)(D3)}{=} \rho((-e)\cdot a,(-e)\cdot o) \leq$

$\leq \rho(-e,o)\rho(a,o) \overset{(MR3)}{=} \rho(a,o) =$

$= \rho((-e)\cdot(-a),(-e)\cdot o) \leq \rho(-e,o)\rho(-a,o) \overset{(MR3)}{=} \rho(-a,o)$.

In der Praxis sind + und · nicht exakt, sondern, etwa auf Rechenanlagen approximiert durch abgeänderte Verknüpfungen $\boxed{+}$ und $\boxed{\cdot}$.

Das Problem, eine geeignete Metrik oder Pseudometrik zu finden, die z. B. das Rasterringoid $(\mathcal{R}R, \boxed{+}, \boxed{\cdot})$ der Gleitkommazahlen zu einem metrischen Ringoid nach Definition 1.1 macht, scheint bisher in keiner befriedigenden Weise gelöst zu sein.

Folgende Beispiele mögen veranschaulichen, woran die klassische Metrik bzw. Pseudometrik im einzelnen scheitert:

Gegeben sei das Ringoid $(\mathbb{R}^*, +, \cdot, \leq)$ der reellen Zahlen und ein Rasterringoid $(\mathcal{R}\mathbb{R}^*, \boxplus, \boxdot, \leq)$ der reellen Zahlen. Ferner sei die übliche translationsinvariante und homogene Metrik $\rho(x,y) = |x-y|$ der reellen Zahlen gegeben, für die (M5) und (M7) gilt.

Diese Eigenschaften gelten jedoch nicht mehr im Rasterringoid $(\mathcal{R}^2_{10}\,\mathbb{R}^*, \boxplus, \boxdot, \rho, \leq)$ der Gleitkommazahlen mit 2 Mantissenstellen zur Basis 10.

Es sei $\Box : R \to \mathcal{R}^2_{10}\,\mathbb{R}^*$ eine optimale Rundung und es gelte für $o \in \{+, \cdot\}$:

$$\text{(V)} \qquad \bigwedge_{x,y \in \mathcal{R}^2_{10}\mathbb{R}^*} x \boxed{o}\, y = \Box(x o y)$$

Für beliebiges $\tau \in \mathbb{N}$ gilt dann:

$$\rho(8.4 \cdot 10^\tau \boxplus 5.2 \cdot 10^\tau, 8.4 \cdot 10^\tau \boxplus 4.10^\tau) =$$

$$= |(8.4 \cdot 10^\tau \boxplus 5.2 \cdot 10^\tau) - (8.4 \cdot 10^\tau \boxplus 4.10^\tau)| = 2 \cdot 10^\tau >$$

$>\rho(5.2 \cdot 10^\tau, 4.10^\tau) = 1.2 \cdot 10^\tau$, also in Widerspruch zur gewünschten Eigenschaft (M5).

Wie bei Gleitpunktrechnung zu erwarten, hängt die Abweichung sogar vom Exponenten ab, so daß (M5) auch nicht durch einen einfachen Faktor zu retten wäre.

Eine ähnliche Abweichung gilt auch für die Homogenität von ρ.

Der Versuch, eine Pseudometrik p nach [Collatz 68] einzuführen, etwa indem man als "Meßraum" statt $\mathbb{R}^+$ den positiven Teil des entsprechenden Rasterringoides $\mathcal{R}\mathbb{R}^+$ wählt, scheitert sofort an der Ungültigkeit der Dreiecksungleichung (M3):

Es ist, im selben Raster wie oben, etwa mit

$p(a,b) := |a - b|$:

$p(1.4\cdot 10^\tau, 16\cdot 10^{\tau+1}) = |1.4\cdot 10^\tau - 16\cdot 10^{\tau+1}| =$

$= 16\cdot 10^{\tau+1} > p(1.4\cdot 10^\tau, 5.7\cdot 10^\tau) + p(5.7\cdot 10^\tau, 16\cdot 10^{\tau+1})$

$= 4.3\cdot 10^\tau + 15\cdot 10^{\tau+1} = 15\cdot 10^{\tau+1}$.

Ein anderer Ansatz, wie er in [Schäfke 70] vorliegt, ist insofern noch unbefriedigend, als er die Frage offen läßt, wie der "Meßraum" K dieser pseudometrischen Limesstruktur geeignet zu konstruieren ist, so daß zumindest die klassischen Fehlerabschätzungen berechnet werden können.

Allgemein scheint der Weg über die klassische Metrik nicht in der allgemeinen Weise zum Ziel zu führen.

In den folgenden Abschnitten wird daher ein anderer Weg, im wesentlichen über den Begriff der Prämetrik, beschritten werden, der schließlich in einfacher Weise sowohl die in [Wilkinson 69] durchgeführte Vorwärtsanalyse als auch die Rückwärtsanalyse für Rundungsfehler ermöglicht.

1.2 Norm und Prämetrik

Der klassische Zusammenhang zwischen Norm und Metrik in linearen Räumen wird hier unwesentlich verallgemeinert durch die Einführung des Begriffes Prämetrik.

Definition 1.3:

Ist L ein linearer Raum (gegebenenfalls genüge wie z. B. bei [Mayer 68] auch ein quasilinearer Raum Q) über dem Skalarkörper $\mathbb{R}$, so heißt ein Funktional $n : L \to \mathbb{R}^+$ eine <u>Norm</u>, wenn gilt:

(N1) $\bigwedge_{x \in L} \quad n(x) = o \iff x = o$

(N2) $\bigwedge_{\substack{\alpha \in \mathbb{R} \\ x \in L}} \quad n(\alpha x) = |\alpha| n(x)$

(gegebenenfalls auch mit $\alpha \in L : n(\alpha x) = n(\alpha)n(x)$)[1].

(N3) $\bigwedge_{x,y \in L} \quad n(x+y) \leq n(x) + n(y)$

Ist zusätzlich $(L,\leq)$ geordnet, so heißt n ordnungsisoton, wenn gilt:

(N4) $\bigwedge_{x,y \in L} \quad o \leq x \leq y \implies n(x) \leq n(y)$

---

[1] Diese geringfügige Verallgemeinerung gilt speziell in $(\mathbb{R},\cdot)$ und in entsprechenden Ringoiden.

Definition 1.4:

Ist M eine beliebige Menge und $(L,\leq)$ ein geordneter linearer (quasilinearer) Raum, so heißt eine zweistellige Abbildung $p : M \times M \to L$ eine *Prämetrik*, bzw. (M,p) ist *prämetrischer Raum*, wenn gilt:

(P1) $$\bigwedge_{x,y \in M} \quad p(x,y) = o \iff x = y$$

(P2) $$\bigwedge_{x,y \in M} \quad p(x,y) + p(y,x) = o \qquad \text{(antisymmetrisch)}$$

(P3) $$\bigwedge_{x,y,z \in M} \quad p(x,z) = p(x,y) + p(y,z)\,.$$

Ist zusätzlich $(M,\leq)$ geordnet, so heißt p verträglich mit $\leq$, wenn gilt

(P4) $$\bigwedge_{x,y,z \in M} \quad x \leq y \leq z \implies o \leq p(y,x) \leq p(z,x) \quad .$$

Ist ferner $(M,+,\cdot)$ ein Ringoid mit den ausgezeichneten Elementen $\{-e,o,e\}$, so ist (M,p) ein *prämetrisches Ringoid*, wenn mindestens (P5)(P7) und (P8) der folgenden Eigenschaften gelten

(P5) $$\bigwedge_{a,x,y \in M} \quad p(a+x,a+y) = p(x,y) \qquad \text{(translationsinvariant)}$$

(P6) $$\bigwedge_{a,b,x,y\in M} p(a+x,b+y) = p(a,b) + p(x,y)$$ (supermetrisch)

(P7) $$\bigwedge_{a,x,y\in M} p(a\cdot x,a\cdot y) = p(a,o) * p(x,y)$$ (homogen)[1]

(P8) $p(e,o) = p(o,-e) = 1$ (normiert)

Unter Umständen sind natürlich in (P5),(P6) und (P7) auch Abschätzungen nach oben möglich, d.h. statt $=$ darf $\leq$ stehen.

Zwischen einer Norm n und einer Prämetrik p besteht folgender (bekannte) klassische Zusammenhang:

Ist

$$M \times M \xrightarrow{p} L \xrightarrow{n} \mathbb{R}^+$$ und

$$M \times M \xrightarrow{\quad q \quad} \mathbb{R}^+,$$ so stellt offenbar

$q := n \circ p$, d. h. $$\bigwedge_{x,y\in M} q(x,y) := n(p(x,y))$$

eine Metrik dar.

---

1) $* : L \times L \to L$ ist ein geeignetes Produkt. Für unsere Problemstellung ist $M \equiv L$ und somit meist $\cdot$ identisch mit $*$.

Die Eigenschaften (M1),(M2),(M3),(M4) leiten sich aus den entsprechenden Eigenschaften (P1),(P2),(P3),(N1),(N2) und (N3) her.

Genauer gilt:

Satz 1.5:

Sind M,L und $\mathbb{R}^+$ die Räume aus Definition 1.3 und 1.4 und ist

$p: M \times M \to L$ eine Prämetrik und

$n: L \to \mathbb{R}^+$ eine isotone Norm, so ist

$q: M \times M \to \mathbb{R}^+$ mit $\bigwedge\limits_{x,y \in M} q(x,y) := n(p(x,y))$

eine Metrik.

Bew.: (M1) folgt aus (P1) und (N1)

(M2) Aus $p(x,y) + p(y,x) = o \Longrightarrow$ da L linear,

$p(x,y) = (-1) \cdot p(y,x) \Longrightarrow$

$q(x,y) = n(p(x,y)) = n((-1) \cdot p(y,x)) \overset{(N2)}{=} n(p(y,x)) = q(y,x)$

(M3) folgt aus der Dreiecksrelation (N3) und (P3)

(M4) Ist $x \leq y \leq z$ und gilt (P4), so ist

$o \leq p(y,x) \leq p(z,x) \Longrightarrow$ nach (N4)

$q(x,y) = n(p(y,x)) \leq n(p(z,x)) = q(z,x)$

(M5) folgt aus (P5)

(M6) folgt aus (P6) und (N3)

(M7) Ist die Norm n verträglich mit $|\cdot|$ in $\mathbb{R}$ und mit p, d.h. ist $n(x) = |p(x,o)|$, so ist

$$q(a\cdot x, a\cdot y) = n(p(a\cdot x, a\cdot y)) =$$

$$\overset{(P7)}{=} n(p(a,o)) \; p(x,y)) =$$

$$\overset{(N2)}{=} |p(a,o)|\, n(p(x,y)) =$$

$$= n(a)\, q(x,y)$$

Auf Grund des in Satz 1.5 dargestellten Zusammenhanges stellt offenbar die Prämetrik eine Vorstufe zu einer Metrik dar. Daher wurde auch der Name so gewählt. Eine Prämetrik unterscheidet sich in $\mathbb{R}$ von einer Metrik z. B. nur im Vorzeichen. Es ist dort z. B. $p(x,y) = x - y$ eine Prämetrik, $|x| = n(x)$ eine Norm und $n(p(x,y)) = |x-y|$ ist bekanntlich eine Metrik. Man könnte daher in diesem Zusammenhang die Prämetrik auch als signierten Abstand bezeichnen.

Die Prämetrik fand bisher in der Literatur nur selten Anwendung und dann meist nicht in dieser expliziten Form. Eine Anzahl von Anwendungsmöglichkeiten bringt z. B. Walter in [Walter 64] und zählt dort auch weitere Literatur auf. Walter verwendet die Prämetrik nicht unter diesem Begriff in expliziter Form, benutzt sie aber dazu, die einseitigen Lipchitzabschätzungen aufzustellen um damit verbesserte Fehlerabschätzungen bei A.W.P. zu erlangen.

Dort ist $p(x,y) = x - y$ eine Prämetrik und dann ist

$$p(f(t,z),f(t,\bar{z})) \leq L\, p(z,\bar{z}) \quad \text{mit} \quad L \in \mathbb{R}$$

eine sogenannte einseitige Lipchitzabschätzung. Bei dieser Art der Abschätzung ist natürlich darauf zu achten, daß p nicht symmetrisch und vorzeichenabhängig ist.

Auch im Zusammenhang mit Rundungsfehlern wird sich zeigen, daß mit Hilfe einer verallgemeinerten Prämetrik, dem prämetrischen Feld, bessere und einfachere Fehlerabschätzungen möglich sind, als mit Hilfe einer entsprechenden verallgemeinerten Metrik, dem metrischen Feld. Das prämetrische Feld ermöglicht sowohl die <u>Vorwärtsanalyse</u> als auch die <u>Rückwärtsanalyse</u> [Wilkinson 69] von Rundungsfehlern, das metrische Feld jedoch nur noch die Vorwärtsanalyse. Dies ist der wesentliche Grund, warum das prämetrische Feld im folgenden einen breiten Raum einnimmt.

2. Das prämetrische Feld

In diesem und in den folgenden Abschnitten wird bereits wesentlicher Gebrauch von den in Kapitel III abgeleiteten Eigenschaften des erweiterten Intervallraumes $\mathbb{H}^*$ bzw. Intervallvektorraumes $V_n\mathbb{H}$ gemacht.

2.1 Definition des prämetrischen Feldes

Zur Beschreibung des prämetrischen Feldes wird der Raum $\mathcal{L}_n$ der quasilinearen Abbildungen über $V_n\mathbb{IR}$ benötigt.

Definition 2.1:

Die Menge

$$\mathcal{L}_n := \{\phi \mid \bigwedge_{X \in V_n\mathbb{IR}} \phi(X) := \mathfrak{A} * X + B \text{ mit } \mathfrak{A} \in M_n\mathbb{IR},\ \mathfrak{A} \text{ Diagonalmatrix},$$

$$\mathfrak{E} \subseteq \mathfrak{A},\quad 0 < \lambda\mathfrak{A},\ B \in V_n\mathbb{IR},\quad 0 \subseteq B\}$$ [1)]

heißt die Menge der quasilinearen Abbildungen von $V_n\mathbb{IR}$ nach $V_n\mathbb{IR}$.

Bemerkung 2.2

1. Der Begriff "quasilineare Abbildung" ist eine Anlehnung an den in [Mayer 68] eingeführten Begriff des quasilinearen Raumes $V_n\mathbb{IR}$. Die Bedeutung der einzelnen Bezeichnungen sind in Abschnitt III,3. nachzusehen.

---

1) Es wird identifiziert: $0 = [0,0]$. Vergl. Abschnitt I,2..
$\mathfrak{E}$ ist die Einheitsmatrix (δ_{ij}) identifiziert mit $([\delta_{ij},\delta_{ij}]) \in M_n\mathbb{IR}$.

2. Da $0 < \lambda\mathcal{A}$ und $\mathcal{A}$ Diagonalmatrix, so besitzt $\mathcal{A}$ eine Inverse. Es ist mit

$$\mathcal{A} = (A_{ii})_{i=1,\dots,n} \qquad \mathcal{A}^{-1} := (1/\overline{A_{ii}})_{i=1,\dots,n}$$

und mit $0 < \lambda A_{ii}$ ist nach Abschnitt III.3 und Satz III,2.26
$\mathcal{A} * \mathcal{A}^{-1} = \mathcal{A}^{-1} * \mathcal{A} = \mathcal{E}$.

Praktisch und beweistechnisch unterscheiden sich die Fälle $\mathcal{A} \in M_n\mathbb{I}\mathbb{R}$ und $\mathcal{A} = A \in \mathbb{I}\mathbb{R}$ formal nicht, wenn man vereinbart, daß $1/\bar{\mathcal{A}} := \mathcal{A}^{-1}$ ist.

-

<u>Definition 2.3</u> (des prämetrischen Feldes)

Sei M eine beliebige Menge und $(V_n\mathbb{I}\mathbb{R}^*, \subseteq)$ der n-dimensionale Intervallvektorraum über dem Ringoid $(\mathbb{R}^*, +, \cdot)$ der reellen Zahlen. Eine Menge ϑ von Abbildungen $D : M \times M \to V_n\mathbb{I}\mathbb{R}^*$ heißt ein <u>prämetrisches Feld</u> über M bzw. (M, ϑ, p) ist ein <u>Raum mit prämetrischem Feld</u>, wenn gilt:

(PF1) $p \in \vartheta$ und (M,p) ist prämetrischer Raum.

(PF2) $\displaystyle\bigwedge_{D \in \vartheta} \; \bigwedge_{x,y,z \in M} D(x,z) = p(x,y) + D(y,z)$

Ist $(M,+,\cdot)$ Ringoid, $(\mathfrak{R}M, \boxplus, \boxdot)$ Rasterringoid von M mit den ausgezeichneten Elementen $\{-e, o, e\}$, (M,p) prämetrisches Ringoid, so heißt ϑ verträglich mit dem Ringoid-Rasterringoid $M \times \mathfrak{R}M$, wenn gilt:

(PF3) $$\bigwedge_{D\in\vartheta}\ \bigvee_{D'\in\vartheta}\ \bigwedge_{\substack{x\in M\\ a,y\in\mathfrak{R}M}} D(a+x,a\boxplus y)\subseteq D'(a,a)+D'(x,y)-\overline{D'(o,o)}$$

(PF4) $$\bigwedge_{D\in\vartheta}\ \bigvee_{D'\in\vartheta}\ \bigwedge_{\substack{x\in M\\ a,y\in\mathfrak{R}M}} D(a\cdot x,a\boxdot y)\subseteq p(a,o)*(D'(x,y)-\overline{D'(o,o)})+D'(o,o)$$

Je nach Problemstellung kann folgende zusätzliche Forderung gestellt werden:

(PF5) ϑ heißt quasilinear, wenn gilt

$$\bigwedge_{D\in\vartheta}\ \bigvee_{\phi\in\mathcal{L}_n}\ \bigwedge_{x,y\in M} D(x,y) = p(x,o) + \phi(p(o,y))$$

<u>Bemerkungen und Folgerungen 2.4:</u>

(1) Vergleicht man die Eigenschaften eines prämetrischen Feldes ϑ mit denen der zugrundegelegten Prämetrik $p\in\vartheta$, so unterscheiden sich beide prinzipiell darin, daß eine Prämetrik ein einelementiges prämetrisches Feld $\vartheta' = \{p\}$ ist, ϑ i.a. dagegen beliebig viele Elemente enthalten kann. ϑ ist also eine Verallgemeinerung einer Prämetrik, denn es ist nach (PF1) $p\in\vartheta$, (PF2) eine entsprechende Verallgemeinerung von (P3), (PF3) von (P5) und (PF4) von (P7).

(2) Die Eigenschaft (PF5) ist nicht unbedingt erforderlich, wird aber in der Praxis stets erfüllt sein, vor allem

in den Fällen, in denen eine Rückwärtsanalyse von Rundungsfehlern erforderlich ist. Quasilineare prämetrische Felder sind additiv in die beiden Argumente zerlegbar, so daß eine rechnerische Auflösung nach einem Argument stets möglich ist. Vergleiche hierzu Abschnitt 2.2.

(3) Aus (P5) (PF2) und (PF3) folgt unmittelbar die Eigenschaft

$$(\mathrm{PF3}^*) \quad \bigwedge_{D\in\vartheta} \; \bigvee_{D'\in\vartheta} \; \bigwedge_{\substack{a,x\in M\\ b,y\in \Re M}} D(a+x,b \boxplus y) \subseteq D'(a,b) + D'(x,y) - \overline{D'(o,o)}$$

Es ist $D(a+x,b \boxplus y) \overset{(PF2)}{=} p(a+x,b+x) + D(b+x,b \boxplus y)$

$$\overset{(P5)}{=} p(a,b) + D(b+x,b \boxplus y)$$

$$\overset{(PF3)}{\subseteq} p(a,b) + D'(b,b) + D'(x,y) - \overline{D'(o,o)}$$

$$\overset{(PF2)}{=} D'(a,b) + D'(x,y) - \overline{D'(o,o)}$$

(4) Entsprechend folgt aus (P7) (PF2) und (PF4) die Eigenschaft:

$(\mathrm{PF4}^*)$

$$\text{a)} \quad \bigwedge_{D\in\vartheta} \; \bigvee_{D'\in\vartheta} \; \bigwedge_{\substack{a,x\in M\\ b,y\in \Re M}} D(a\cdot x,b \boxdot y) \subseteq p(a,o) * p(x,y) + p(y,o) * (D'(a,b)-\overline{D'(o,o)})+D'(o,o)$$

b) $$\bigwedge_{D\in\vartheta}\ \bigvee_{D'\in\vartheta}\ \bigwedge_{\substack{a,x\in M\\ b,y\in \mathcal{R}M}} D(a\cdot x,b\boxdot y) \subseteq p(x,o)*p(a,b) + p(b,o)*(D'(x,y)-\overline{D'(o,o)}) + D'(o,o)$$

a) ergibt sich wie folgt:

$$D(a\cdot x,b\boxdot y) \overset{(PF2)}{=} p(a\cdot x,a\cdot y) + D(a\cdot y,b\boxdot y)$$

$$\overset{(P7)(PF4)}{\subseteq} p(a,o)p(x,y) + p(y,o)*(D'(a,b)-\overline{D'(o,o)}) + D'(o,o)$$

b) ergibt sich analog aus Symmetriegründen.

(5) a) Vermöge der Beziehung $-a:=(-e)\cdot a$ gilt in Ringoiden bei Gültigkeit von (PF4) die Abschätzung:

$$\bigwedge_{D\in\vartheta}\ \bigvee_{D'\in\vartheta}\ \bigwedge_{x,y\in M} D(-x,-y) \subseteq p(-e,o)*(D'(x,y)-\overline{D'(o,o)}) + D'(o,o)$$

und bei Gültigkeit von (PF5) und (P7) speziell:

$$\bigwedge_{D\in\vartheta}\ \bigwedge_{x,y\in M} D(-x,-y) \subseteq p(-e,o)*(D(x,y)-\overline{D(o,o)}) + D(o,o)$$

Es ist $D(-x,-y) = p(-x,o) + \phi p(o,-y)$

$$\overset{(P7)}{=} p(-e,o)p(x,o) + \phi(p(-e,o)p(o,y))$$

nach Satz 2.6(4) $\subseteq p(-e,o)p(x,o)+p(-e,o) * (\phi p(o,y)-\overline{\phi,o}))+\phi(o)$

$$= p(-e,o) * (p(x,o)+\phi p(o,y)-\overline{\phi(o)}+\phi(o)$$

$$= p(-e,o) * (D(x,y))-\overline{D(o,o)})+D(o,o).$$

b) Ist $(R,N,+,\cdot,/)$ ein Divisionsringoid, so werden i.a. folgende Eigenschaften gegeben sein:

$$\bigwedge_{D\in\mathcal{V}} \ \bigwedge_{x,y\in M\setminus N} D(1/x,1/y)=D(y,x)/p(x\cdot y,o) \quad \text{und}$$

(D9) $$\bigwedge_{\substack{a,b\in M\\ b\notin N}} a/b = a \cdot 1/b \quad \text{und}$$

(V) $$\bigwedge_{\substack{a,b\,\in\mathcal{R}M\\ b\notin N}} a \boxed{/} b = \square(a/b)$$

2.2 Die Vorwärts- und Rückwärtsanalyse von Rundungsfehlern mit Hilfe prämetrischer Felder

In [Wilkinson 69] wird das Prinzip der Vorwärts- und Rückwärtsanalyse von Rundungsfehlern zusammenfassend beschrieben und in vielen praktischen Problemen angewendet. Hier soll nun gezeigt werden, wie mit dem Kalkül des prämetrischen Feldes ein Mechanismus entwickelt ist, der diese beiden Analysenverfahren ermöglicht.

Das Prinzip kann hier nicht in voller Allgemeinheit vorgeführt werden, da hierzu eine schärfere Formulierung des Begriffes "Algorithmus zur Berechnung einer Größe" notwendig wäre, was in diesem Rahmen zu weit führen würde und im wesentlichen keinen Einfluß auf den Kalkül der prämetrischen Felder hat.

Der Einfachheit halber möge daher angenommen werden, daß ausgehend von einem Datensatz $a_1,\dots,a_n$ über dem Rasterringoid $(\mathcal{R}R, \boxplus, \boxdot)$ eines Ringoides $(R,+,\cdot)$ eine Zahl $x \in R$ mit Hilfe der arithmetischen Grundoperationen $+$ und $\cdot$ des Ringoides R berechnet werden soll. Es sei also eine arithmetische Funktion $A : \mathcal{R}R^n \to R$ gegeben, zusammengesetzt mit den Verknüpfungen $+$ und $\cdot$ des zugrundegelegten Ringoides. Es ist dann

$$x := A(a_1,\dots,a_n)$$

zu berechnen.

In der Praxis findet die Berechnung jedoch mit abgeänderten Verknüpfungen $\boxplus$ und $\boxdot$ des Rasterringoides (etwa einer Rechenanlage [Kulisch 72]) statt.

An Stelle von A liegt nun die abgeänderte arithmetische Funktion $\alpha A : \mathcal{R}R^n \to \mathcal{R}R$, die sogenannte Rasterrestriktion von A der Berechnung zu Grunde. (Vergl. Def. I,1.5) αA entsteht in diesem Falle aus A indem $+$ und $\cdot$ durch die entsprechenden Rasterverknüpfungen $\boxplus$ und $\boxdot$ ersetzt werden. Der mit αA berechnete Wert

$$\overline{x} := \alpha A(a_1,\dots,a_n)$$

weicht i.a. vom "exakten" Wert x ab. Diese Abweichung, der sogenannte Rundungsfehler ist dann von der Größe $\rho(x,\overline{x})$, wenn ρ eine geeignete Metrik im Ringoid R ist.

Zur Berechnung der Größe $\rho(x,\overline{x})$ gibt es nach Wilkinson zwei Verfahren, die sogenannten Vorwärts- und Rückwärtsanalysen. Sie lassen sich nun in Anlehnung an [Wilkinson 69] in einem geeigneten prämetrischen Feld ϑ, das mit dem Ringoid-Rasterringoid $R \times {}^{\mathcal{R}} R$ verträglich ist, wie folgt beschreiben.

Die <u>Vorwärtsanalyse</u>

berechnet durch sukzessive Anwendung der Eigenschaften (PF1) bis (PF4) des zugrundegelegten prämetrischen Feldes eine arithmetische Funktion $B : V_n \mathbb{I}\mathbb{R} \to V_n \mathbb{I}\mathbb{R}$ derart, daß die Abschätzung

$$\begin{aligned} p(x,\overline{x}) = p(A(a_1,\ldots,a_n),\ \alpha A(a_1,\ldots,a_n)) &\subseteq \\ \subseteq B(D_1(a_1,a_1), D_2(a_2,a_2),\ \ldots\ ,\ D_n(a_n,a_n),& \\ p(a_1,o),\ p(a_2,o),\ \ldots\ ,\ p(a_n,o))& \end{aligned}$$

mit $D_i \in \vartheta$ für $i = 1, \ldots, n$ gilt.

Die so berechnete Abschätzung für $p(x,\overline{x})$ liefert dann nach Satz 1.5 und nach Lemma III,4.4(f) mit Hilfe der Intervallnorm $\| \cdot \|$ die gesuchte Abschätzung

$$\rho(x,\overline{x}) = \|p(x,\overline{x})\| \leq \|B(\ldots)\| .$$

Beispiel:

Sei $x = A(a,b,c) = a \cdot b + c$ in $\mathbb{R}^*$ zu berechnen mit $a,b,c \in \mathcal{R}\mathbb{R}^*$. Wird das Ergebnis

$$\overline{x} = \alpha A(a,b,c) = a \boxdot b \boxplus c \text{ in } \mathcal{R}\mathbb{R}^* \text{ berechnet, so ist}$$

$$\begin{aligned}
p(x,\overline{x}) &= p(a \cdot b + c, a \boxdot b \boxplus c) \subseteq \\
&\overset{(PF3)}{\subseteq} D'(a \cdot b, a \boxdot b) + D'(c,c) - \overline{D'(o,o)} \subseteq \\
&\overset{(PF4)}{\subseteq} p(a,o) * (D''(b,b) - \overline{D''(o,o)}) + D''(o,o) + D'(c,c) - \overline{D'(o,o)} = \\
&=: B(D'(c,c), D''(b,b), p(a,o)).
\end{aligned}$$

Die Rückwärtsanalyse

läßt sich wie folgt beschreiben. Statt den tatsächlichen Rundungsfehler $\rho(x,\overline{x})$ zu berechnen, werden (meist in einem geeigneten Zwischenschritt) abgeänderte Ausgangsdaten $\tilde{a}_1,\dots,\tilde{a}_n$ gesucht, so daß gilt:

(I) $A(\tilde{a}_1,\dots,\tilde{a}_n) = \alpha A(a_1,\dots,a_n)$

Die neuen Daten $\tilde{a}_1,\dots,\tilde{a}_n$ sind i.a. nicht eindeutig bestimmt. Es genügt, Schranken für die Abweichung $p(\tilde{a}_i,a_i)$ für $i = 1,\dots,n$ anzugeben. Durch zusätzliche Überlegungen wird dann auf den Rundungsfehler des Endergebnisses geschlossen. (x und $\overline{x}$ sind in diesem Zusammenhang nur Zwischenergebnisse!) Das Vorgehen im zugrundegelegten prämetrischen Feld ϑ ist das folgende: Man gewinnt durch sukzessive Anwendung der Eigenschaften (PF1) bis (PF4) eine arithmetische Funktion $C : V_n\mathbb{IR} \to V_n\mathbb{IR}$ so, daß gilt:

(II) $p(A(\tilde{a}_1,\dots,\tilde{a}_n),\ \alpha A(a_1,\dots,a_n)) \subseteq$

$$\subseteq C(D_1(\tilde{a}_1,a_1),\dots,D_m(\tilde{a}_n,a_n))$$

mit $D_i \in \mathcal{V}$, $i = 1,\dots,m$ und $o \subseteq C(o,\dots,o)$.

Nach Eigenschaft (P1) der dem prämetrischen Feld $\mathcal{V}$ zugrundegelegten Prämetrik p ist $x = y \iff p(x,y) = o$, also gilt (I) genau dann, wenn die linke Seite von (II) Null ist, d.h. es entsteht die (notwendige) Forderung:

(III) $o \subseteq C(D_1(\tilde{a}_1,a_1),\ \dots,\ D_m(\tilde{a}_n,a_n))$

Da, wie erwähnt, i.a. keine eindeutige Bestimmung der $\tilde{a}_i$ gefordert ist, genügen die hinreichenden Bedingungen

(IV) $$\bigwedge_{j=1,\dots,m} o \subseteq D_j(\tilde{a}_i,a_i) ,$$

zur Erfüllung von (III)[1]. Dies folgt aus der Gültigkeit des Einschließungssatzes (Satz III,1.2) in $V_n\mathbb{IR}$ für arithmetische Funktionen und der in (II) geforderten Eigenschaft $o \subseteq C(o,o,\dots,o)$ von C .

---

1) An dieser Stelle muß gesichert sein, daß wegen (PF3) und (PF4) $\bigwedge_{D\in\mathcal{V}} o \subseteq D(o,o)$ gilt, was im Falle der Gültigkeit von (PF5) mit $\phi(o) = B \geq o$ in $\mathcal{L}_n$ stets gewährleistet ist.

I.a. hängt es nun von der Struktur des prämetrischen Feldes ab, wie die Auflösung in (IV) nach $p(\tilde{a}_i, a_i)$ möglich ist. Ist jedoch das zugrundegelegte prämetrische Feld ν quasilinear, d.h. es gilt (PF5):

$$D_i(\tilde{a}_i, a_i) = p(\tilde{a}_i, o) + \phi_i(p(o, a_i)), \text{ so}$$

folgt aus (IV) unmittelbar:

$$o \subseteq p(\tilde{a}_i, o) + \phi_i(p(o, a_i)) \quad \text{für } i = 1, \ldots, n$$

und mit $p(\tilde{a}_i, o) \in V_n\mathbb{R}$ das Ergebnis

$$\text{(V)} \quad p(\tilde{a}_i, o) \subseteq -\phi_i(p(o, a_i)) \quad \text{für } i = 1, \ldots, n$$

bzw. mit Hilfe von (P3)

$$p(\tilde{a}_i, a_i) \subseteq p(o, a_i) - \phi_i(p(o, a_i))$$

Vergleicht man das Ergebnis mit dem in [Wilkinson 69], so erhält man im Falle $p(x,y) := x-y$ mit $p(\tilde{a}_i, o) = \tilde{a}_i = a_i(1+\varepsilon_i)$ und $\varepsilon_{i1} \leq \varepsilon_i \leq \varepsilon_{i2}$ bzw. $\varepsilon_i \subseteq [\varepsilon_{i1}, \varepsilon_{i2}] =: E_i$ die Abschätzung

$p(\tilde{a}_i, o) = \tilde{a}_i \subseteq a_i(1+E_i)$.

Aus (V) folgt in diesem Falle

$$\tilde{a}_i \subseteq \phi_i(a_i) = \mathfrak{A}_i * a_i + B_i$$

Wie sich später zeigen wird, ist $B_i \neq o$ nur bei Berücksichtigung des Unterlaufes, der bei Wilkinson jedoch nicht berücksichtigt wird. Damit erhält man mit $B_i = o$ und $\mathfrak{A}_i = (1+E_i)$

eine äquivalente Abschätzung.

Beispiel:

Sei wieder $A(a,b,c) = a \cdot b + c$ und

$\alpha A(a,b,c) = a \boxed{\cdot} b \boxed{+} c$ vorgelegt, so folgt für (II) die Beziehung

$$p(\tilde{a}\cdot\tilde{b}+\tilde{c}, a \boxed{\cdot} b \boxed{+} c) \overset{(PF3^*)}{\subseteq} D'(\tilde{a}\cdot\tilde{b}, a \boxed{\cdot} b) + D'(\tilde{c},c) - \overline{D'(o,o)}.$$

Nach (PF4*) kann der 1. Summand in folgende zwei Arten abgeschätzt werden:

a) $D'(\tilde{a}\cdot\tilde{b}, a \boxed{\cdot} b) \overset{(PF4^*)a)}{\subseteq} p(\tilde{a},o)p(\tilde{b},b) +$

$$p(b,o) * (D''(\tilde{a},a) - \overline{D''(o,o)}) + D''(o,o)$$

b) $D'(\tilde{a}\cdot\tilde{b}, a \boxed{\cdot} b) \overset{(PF4^*)b)}{\subseteq} p(\tilde{b},o)p(\tilde{a},a) +$

$$p(a,o) * (D''(\tilde{b},b) - \overline{D''(o,o)}) + D''(o,o)$$

zusammenfassend also ergibt sich im Falle a) die Funktion

$C_a(p(\tilde{b},b), D''(\tilde{a},a), D'(\tilde{c},c)) =$

$$= p(\tilde{a},o)p(\tilde{b},b) + p(b,o) * (D''(\tilde{a},a) - \overline{D''(o,o)} + D''(o,o) + D'(\tilde{c},c) - \overline{D'(o,o)}$$

mit $C_a(o,o,o) = o$. Damit ist (II) erfüllt und nach (V) ergeben sich z. B. mit $p(x,y) := x - y$ die Abschätzungen:

$$\tilde{b} = b$$

$$\tilde{a} \subseteq \phi''(a) \quad \text{wenn} \quad D''(\tilde{a},a) = \tilde{a} - \phi''(a) \quad \text{und}$$

$$\tilde{c} \subseteq \phi'(c) \qquad D'(\tilde{c},c) = \tilde{c} - \phi'(c) \quad \text{ist.}$$

Im Falle b) ergibt sich auf entsprechende Weise

$$\tilde{a} = a$$

$$\tilde{b} \subseteq \phi''(b)$$

$$\tilde{c} \subseteq \phi'(c) \; .$$

Hier zeigt sich auf natürliche Weise, daß es gleichgültig ist, ob man den Fehler des Produktes $a \boxdot b = \tilde{a} \cdot \tilde{b}$ $\tilde{a}$ oder $\tilde{b}$ auflastet. Unter geeigneten zusätzlichen Voraussetzungen an ϑ kann diese Unsymmetrie vermieden werden.

Wie die bisherige Untersuchung zeigt, ist die Fehleranalyse eines arithmetischen Problems nur noch ein einfacher Abschätzalgorithmus, wenn in dem betreffenden Raum ein geeignetes prämetrisches Feld ϑ mit den Eigenschaften (PF1) bis (PF4) gegeben ist.

Die Eigenschaft (PF5) ist nur ein Zusatz für den Fall der Rückwärtsanalyse, damit eine einfache Auflösung in den Argumenten gewährleistet ist.

Der folgende Abschnitt widmet sich dem Problem, wie in Ringoiden und zugehörigen Rasterringoiden und gegebener Rundung prämetrische Felder zu konstruieren sind.

2.3 Konstruktion prämetrischer Felder

Zur Darstellung und Beschreibung der im Folgenden auftretenden speziellen prämetrischen Felder werden folgende Begriffe und Eigenschaften benötigt.

Definition 2.5:

(1) Sei $\mathcal{L}_n$ der in Definition 2.1 eingeführte Raum der quasilinearen Abbildungen $\phi : V_n\mathbb{IR} \to V_n\mathbb{IR}$. $\mathcal{L}_n^*$ bezeichne dann die Menge der durch die Potenzbildung

$$\phi^{(k)}(X) := \phi(\phi^{(k-1)}(X)) \quad \text{für} \quad k \geq 1, \quad \phi \in \mathcal{L}_n \quad \text{und}$$
$$\phi^{(o)}(X) = X$$

über $\mathcal{L}_n$ erzeugten Funktionen:

$\mathcal{L}_n^* := \{\psi = \phi^{(k)} \mid k \in \mathbb{N} \wedge \phi \in \mathcal{L}_n\}$.

$(\mathcal{L}_n^*, \subseteq)$ ist geordnet mit $\subseteq$ gemäß: Mit $\psi_1, \psi_2 \in \mathcal{L}_n^*$, $\psi_1 = \phi_1^{(p)}$ und $\psi_2 = \phi_2^{(q)}$ ist $\psi_1 \subseteq \psi_2 :\Longleftrightarrow \phi_1 \subseteq \phi_2 \wedge p \leq q$

(2) Es sei die Funktionenabbildung $\delta : \mathcal{L}_n^* \to \mathcal{L}_n$ definiert als die Anwendung des Subdistributivgesetzes in $\mathbb{IR}$ nach Bemerkung III,2.24 auf alle Klammerungen, die durch die Potenzbildung $\phi^{(k)} \in \mathcal{L}_n^*$ der arithmetischen Funktion $\phi \in \mathcal{L}_n$ entstanden sind, so daß alle Klammern aufgelöst sind und eine Funktion $\delta\phi^{(k)} = \tilde{\phi} \in \mathcal{L}_n$ entsteht.

Formal wirkt also δ wie folgt:

Sei $\phi(X) = \mathfrak{A} * X + B$, so ist

$$\phi^{(k)}(X) = \underbrace{\mathfrak{A} * (\mathfrak{A} * \ldots\ldots * (\mathfrak{A}}_{k\text{-mal}} * X + \underbrace{B) + \ldots + B) + B}_{k\text{-mal}}$$

und es ist dann

$$\delta\phi^{(k)}(X) := \mathfrak{A}^{(k)} * X + \sum_{\nu=0}^{k-1} (\mathfrak{A}^{(\nu)} * B)$$ [1)]

mit $\mathfrak{A}^{(k)} := \mathfrak{A}^{(k-1)} * \mathfrak{A}$ für $k \geq 1 \wedge \mathfrak{A}^{(0)} = 1$.

<u>Beispiel:</u>

Sei $\phi(X) = A * X + B \in \mathcal{L}_1$, so ist

$$\phi^{(3)}(X) = A * (A * (A * X + B) + B) + B \in \mathcal{L}_1^* .$$

Es ist dann

$$\delta\phi^{(3)}(X) = \delta\{A * (A * (A * X + B) + B) + B\} =$$

$$= A^{(3)} * X + A^{(2)} * B + A * B + B =$$

$$= A' * X + B' \in \mathcal{L}_1 \quad \text{mit}$$

$A' := A^{(3)} = A * A * A$ und $B' := A * A * B + A * B + B$.

---

1) δ ist nicht trivial, denn in $V_n\mathbb{IR}$ gilt kein Distributivgesetz.

Die Abbildung δ, und die Räume $\mathcal{L}_n$ und $\mathcal{L}_n^*$ stehen in folgender Beziehung und haben folgende Eigenschaften:

<u>Satz 2.6:</u>

(1) Es ist $\mathcal{L}_n \subseteq \mathcal{L}_n^*$ und die Ordnungsrelation $\subseteq_{\mathcal{L}_n}$ von $(\mathcal{L}_n, \subseteq_{\mathcal{L}_n})$ ist dieselbe wie die von $(\mathcal{L}_n^*, \subseteq_{\mathcal{L}_n^*})$ im Teilraum $\mathcal{L}_n$ induzierte Ordnungsrelation.

(2) $(\mathcal{L}_n^*, \subseteq)$ und $(\mathcal{L}_n, \subseteq)$[1)] sind Vollverbände und $\mathcal{L}_n = \mathcal{R}\mathcal{L}_n^*$ ist ein Raster von $\mathcal{L}_n^*$.

(3) $\delta : \mathcal{L}_n^* \to \mathcal{L}_n$ ist eine isotone, optimale und (nach oben) gerichtete Rundung bezüglich $\subseteq$, d.h. es gilt:

(R1) $\psi_1 \subseteq \psi_2 \Longrightarrow \delta\psi_1 \subseteq \delta\psi_2$

(R2) $\psi \in \mathcal{L}_n \Longrightarrow \delta\psi = \psi$

(R3) $\psi \subseteq \delta_{\psi}$

(4) $$\bigwedge_{\phi \in \mathcal{L}_n} \phi(X+Y) \subseteq \phi(X) + \phi(Y) - \overline{\phi(0)}$$

$$\bigwedge_{\phi \in \mathcal{L}_n} \phi(X * Y) \subseteq X * (\phi(Y) - \overline{\phi(0)}) + \phi(0) \quad \text{b.z.w.}$$
$$\subseteq Y * (\phi(X) - \overline{\phi(0)}) + \phi(0)$$

$$\bigwedge_{\phi \in \mathcal{L}_n} \bigwedge_{X \in V_n\mathbb{IR}} X \subseteq \phi(X)$$

(5) $$\bigwedge_{\phi \in \mathcal{L}_n} \bigwedge_{X \in V_n\mathbb{IR}} \delta\phi^{(i)}(\phi(X)) \subseteq \delta\phi^{(i+1)}(X)$$

---

1) Der Index bei $\subseteq$ wird wegen der Aussage (1) künftig weggelassen.

Bew.:

(1) Es ist per Definition $\phi^{(1)} \in \mathcal{L}^*_n \wedge \phi^{(1)} = \phi \in \mathcal{L}_n$.

Ferner ist per Definition mit $\phi_1, \phi_2 \in \mathcal{L}_n$ $\phi_1 \subseteq_{\mathcal{L}^*_n} \phi_2 \Longleftrightarrow$

$\left(\phi_1{}^{(1)} \subseteq_{\mathcal{L}^*_n} \phi_2{}^{(1)}\right) \Longleftrightarrow \left(\phi_1 \subseteq_{\mathcal{L}_n} \phi_2 \wedge 1 \leq 1\right)$, d. h.

$\phi_1 \subseteq_{\mathcal{L}^*_n} \phi_2 \Longleftrightarrow \phi_1 \subseteq_{\mathcal{L}_n} \phi_2$.

(2) Der Raum $(\mathcal{L}_n, \subseteq_{\mathcal{L}})$ ist ordnungsisomorph zu dem Produktraum $(\underbrace{\mathcal{L}_1 \times \mathcal{L}_1 \ \ldots \ \times \mathcal{L}_1}_{n\text{-mal}}, \subseteq_{\mathcal{L}_n})$, da $\subseteq_{\mathcal{L}_n}$ komponentenweise definiert ist.

$(\mathcal{L}_1, \subseteq_{\mathcal{L}_1})$ ist aber ordnungsisomorph zum $(V_4\mathbb{R}^{*+}, \leq)$, wobei $\leq$ komponentenweise über $(\mathbb{R}^{*+}, \leq)$ zu verstehen ist.

Diese Behauptung ergibt sich wie folgt:

$\phi \in \mathcal{L}_1$ hat die Gestalt $\phi(X) = A * X + B$ mit $1 \subseteq A \wedge 0 \subseteq B$.

Nun ist $\phi_1 \subseteq_{\mathcal{L}_1} \phi_2 \Longleftrightarrow \bigwedge_{X \in \mathbb{IR}} \phi_1(X) \subseteq_{\mathbb{IR}} \phi_2(X)$,

und speziell folgt für $i = 1,2$ $\phi_i(X) = A_i * X + B_i$ mit $X = 0$.

$\phi_1 \subseteq \phi_2 \longrightarrow B_1 \subseteq B_2$.

Ferner folgt für $\xi \in \mathbb{R}^{*+}$

$$\frac{\phi_1(\xi)}{\xi} \subseteq \frac{\phi_2(\xi)}{\xi} \Longleftrightarrow A_1 + \frac{B_1}{\xi} \subseteq A_2 + \frac{B_2}{\xi}$$

und für $\xi \to +p(= +\infty)$ $A_1 \subseteq A_2$.

Nun ist mit $A_i = [1-a_i^1,\ 1+a_i^2]$ und

$B_i = [-b_i^1,\ +b_i^2]$ für $i = 1,2$ und $a_i, b_i \in \mathbb{R}^{*+}$

$$A_1 \subseteq A_2 \ \wedge\ B_1 \subseteq B_2 \Longleftrightarrow$$

$$a_1^1 \leq a_2^1 \ \wedge\ a_1^2 \leq a_2^2 \ \wedge\ b_1^1 \leq b_2^1 \ \wedge\ b_1^2 \leq b_2^2$$

$$\Longleftrightarrow$$

$$(a_1^1,\ a_1^2,\ b_1^1,\ b_1^2) \ \leq\ (a_2^1,\ a_2^2,\ b_2^1,\ b_2^2) \text{ in } (V_4\mathbb{R}^{*+}, \leq) \ .$$

$(V_4\mathbb{R}^{*+}, \leq)$ ist ein Vollverband, also ist $(\mathcal{L}_1, \subseteq_{\mathcal{L}_1})$ ein Vollverband und schließlich wie behauptet $(\mathcal{L}_n, \subseteq_{\mathcal{L}_n})$ ein Vollverband.

$(\mathcal{L}_n^*, \subseteq_{\mathcal{L}_n^*})$ ist per Definition von $\subseteq_{\mathcal{L}_n^*}$ ordnungsisomorph zu $(\mathcal{L}_n \times \mathbb{N}^*, \subseteq_o)$ bezüglich des Isomorphismus:

$$\mathcal{L}_n^* \ni \phi^k \to (\phi, k) \in \mathcal{L}_n \times \mathbb{N}^*$$

mit $\mathbb{N}^* := \mathbb{N} \cup \{\infty\}$ und

$$(\phi_1, p) \subseteq_o (\phi_2, q) \Longleftrightarrow \phi_1 \subseteq_{\mathcal{L}_n} \phi_2 \ \wedge\ p \leq q$$

$$\Longleftrightarrow \phi_1^{(p)} \subseteq_{\mathcal{L}_n^*} \phi_2^{(q)} \ .$$

$(\mathcal{L}_n, \subseteq)$ ist nach obigem Beweis ein Vollverband ebenso wie $(\mathbb{N}^*, \leq)$. Somit ist $(\mathcal{L}_n \times \mathbb{N}^*, \subseteq_o)$ ein Vollverband und damit auch $(\mathcal{L}_n^*, \subseteq_{\mathcal{L}_n^*})$.

Nun bleibt noch zu zeigen, daß $\mathcal{L}_n$ Teilverband von $\mathcal{L}_n^*$ ist.

Für das Supremum gilt nach obiger Überlegung:

$$\sup_{\mathcal{L}_n^*}(\phi_1^{(p)}, \phi_2^{(q)}) = \sup_{\mathcal{L}_n^*}((\phi_1, p), (\phi_2 q))$$

(Def)

$$= (\sup_{\mathcal{L}_n}(\phi_1, \phi_2), \sup_{\mathbb{N}^*}(p, q))$$

$$= (\sup_{\mathcal{L}_n}(\phi_1, \phi_2))^{(\sup_{\mathbb{N}^*}(p,q))}$$

Speziell für $p = q = 1$ aber ist

$\sup_{\mathcal{L}_n^*}(\phi_1, \phi_2) = \sup_{\mathcal{L}_n}(\phi_1, \phi_2)$. Dasselbe folgt für das Infimum.

Somit ist $\mathcal{L}_n$ Teilverband von $\mathcal{L}_n^*$ und es gilt (S_2'').

Mit $i(\mathcal{L}_n^*) = i(\mathcal{L}_n) = [-p, p]$ und

$o(\mathcal{L}_n^*) = o(\mathcal{L}_n) = \emptyset$ gilt auch (S_1'')

und damit ist $\mathcal{L}_n = \mathcal{R}\mathcal{L}_n^*$ Raster von $\mathcal{L}_n^*$.

(3) (R1) Sei $\psi_1 = \phi_1^{(p)}$ und $\psi_2 = \phi_2^{(q)}$, so gilt:

$$\psi_1 \subseteq \psi_2 \iff \phi_1 \subseteq \phi_2 \quad \wedge \quad p \leq q \ .$$

Sei $\phi_i(X) = \mathfrak{A}_i * X + B_i$ für $i = 1,2$, so gilt:

$$\psi_1 \subseteq \psi_2 \iff \mathfrak{A}_1 \subseteq \mathfrak{A}_2 \quad \wedge \quad B_1 \subseteq B_2 \quad \wedge \quad p \leq q \ .$$

Nun ist für beliebiges $X \in V_n I\!IR$

$$\delta\psi_1(X) = \mathfrak{A}_1^{(p)} * X + \sum_{\nu=1}^{p-1} B_1 * \mathfrak{A}_1^{(\nu)}$$

$$\subseteq \mathfrak{A}_2^{(p)} * X + \sum_{\nu=1}^{p-1} B_2 * \mathfrak{A}_2^{(\nu)} = \delta\phi_2^{(p)}(X) \ .$$

Wegen $X \subseteq \phi_2^{(q-p)}(X)$ folgt schließlich

$$\delta\psi_1(X) \subseteq \delta\phi_2^{(p)}(\phi_2^{(q-p)}(X)) \subseteq$$

$$\overset{(5)}{\subseteq} \delta\phi_2^{q}(X) = \delta\psi_2(X) \ , \quad \text{also ist}$$

$$\delta\psi_1 \subseteq \delta\psi_2 \quad .$$

(R2) Nach Definition ist $\delta(\mathfrak{A} * X + B) = \mathfrak{A} * X + B$, d. h. δ ist auf $\mathcal{L}_n$ wirkungslos.

(R3) ist eine unmittelbare Folge des Subdistributivgesetzes in $V_n\mathbb{IR}$ nach Bemerkung III,2.24 und des Einschließungssatzes Satz III,1.2.

(4) Sei $\phi \in \mathcal{L}_n$, so ist

$$\begin{aligned}\phi(X + Y) &= A * (X + Y) + B \subseteq \\ &\subseteq A * X + A * Y + B = \\ &= A * X + A * Y + B = \\ &= (A * X + B) + (A * Y + B) - \overline{B} = \\ &= \phi(X) + \phi(Y) - \overline{\phi(0)}\end{aligned}$$

$$\begin{aligned}\phi(X * Y) &= A * X * Y + B = \\ &= X * (A * Y) + B = \\ &= X * (\phi(Y) - \overline{\phi(0)}) + \phi(0) .\end{aligned}$$

Die Inklusionsisotonie gilt für arithmetische Funktion^en in $V_n\mathbb{IR}$. Es folgt daher

mit $\mathfrak{E} \subseteq \mathfrak{A} \Longrightarrow X \subseteq \mathfrak{A} * X$ und schließlich

mit $0 \subseteq B \Longrightarrow X \subseteq \mathfrak{A} * X + B = \phi(X)$.

(5) $\delta\phi^{(i)}(\phi(X)) = \mathfrak{A}^{(i)} * (\mathfrak{A} * X + B) + \sum_{\nu=o}^{i-1} \mathfrak{A}^{(\nu)} * B \subseteq$

$\subseteq \mathfrak{A}^{(i+1)} * X + \mathfrak{A}^{(i)} * B + \sum_{\nu=o}^{i-1} \mathfrak{A}^{(\nu)} * B =$

$= \mathfrak{A}^{(i+1)} * X + \sum_{\nu=o}^{i} \mathfrak{A}^{(\nu)} * B = \delta\phi^{(i+1)}(X)$

<u>Definition 2.7:</u>

(1) Ein prämetrisches Feld heißt <u>verträglich mit einem Ringoid</u> $(R,+,\cdot)$, wenn

(1.1) $(R,\mathcal{V},p)$ ein Raum mit prämetrischem Feld ist und

(1.2) (PF31) $\bigwedge_{D\in\mathcal{V}} \ \bigwedge_{a,x,y\in R} \ D(a+x,a+y) \subseteq D(a,a) + D(x,y) - \overline{D(o,o)}$

(PF41) $\bigwedge_{D\in\mathcal{V}} \ \bigwedge_{a,x,y\in R} \ D(a\cdot x,a\cdot y) \subseteq p(a,o) * (D(x,y)-\overline{D(o/o)}+D(o,o))$

gilt.

(2) Ein prämetrisches Feld $\mathcal{V}$ heißt <u>rechtsverträglich mit einem Rasterringoid</u> $(\mathcal{R}R, \boxplus, \boxdot)$ bezüglich eines Ringoides $(R,+,\cdot)$, wenn gilt:

(2.1) $(R,\mathcal{V},p)$ ist ein Raum mit prämetrischem Feld

(2.2) (PF32) $\bigwedge_{D\in\mathcal{V}} \ \bigvee_{D'\in\mathcal{V}} \ \bigwedge_{\substack{a,x\in R\\ x,y\in\mathcal{R}R}} \ D(a,x \boxplus y) \subseteq D'(a,x+y)$

(PF42) $\bigwedge\limits_{D\in\vartheta} \bigvee\limits_{D'\in\vartheta} \bigwedge\limits_{\substack{a\in R\\ x,y\in \mathcal{R}R}} D(a,x \boxdot y) \subseteq D'(a,x \cdot y)$

(3) Ein prämetrisches Feld ϑ heißt rechtsverträglich mit einer Rundung $\square : R \to \mathcal{R}R$, wenn

(3.1) (R,ϑ,p) ein Raum mit prämetrischem Feld ist und

(3.2) (PF6) $\bigwedge\limits_{D\in\vartheta} \bigvee\limits_{D'\in\vartheta} \bigwedge\limits_{x,y\in R} D(x,\square y) \subseteq D'(x,y)$ gilt.

Folgendes Lemma zeigt nun den Zusammenhang zwischen der Rechtsverträglichkeit von prämetrischen Feldern mit einer Rundung und der Rechtsverträglichkeit mit einem geeigneten Rasterringoid.

<u>Lemma 2.8:</u>

Ist $(R,+,\cdot)$ Ringoid, $(\mathcal{R}R, \boxplus, \boxdot)$ Rasterringoid, $\square : R \to \mathcal{R}R$ eine Rundung und gilt

(V) $\bigwedge\limits_{\circ\in\{+,\cdot\}} \bigwedge\limits_{a,b\in\mathcal{R}R} a \mathbin{\boxed{\circ}} b = \square(a \circ b)$,

so ist jedes mit der Rundung $\square$ rechtsverträgliches prämetrisches Feld stets auch rechtsverträglich mit dem Rasterringoid $(\mathcal{R}R, \boxplus, \boxdot)$, d. h. in der Nummerung von Definition 2.7:

Aus (3) und (V) folgt (2).

Bew.: Seien $a \in R$, $x,y \in \mathcal{R}R$ beliebig, so folgt:

(PF32) $D(a,x \boxplus y) \overset{(V)}{=} D(a, \square(x + y)) \subseteq$

$\overset{(PF6)}{\subseteq} D'(a,x + y)$ mit $D' \in \vartheta$

(PF42) $D(a,(x \boxdot y)) \overset{(V)}{=} D(a, \square(x \cdot y)) \subseteq$

$\overset{(PF6)}{\subseteq} D'(a,x \cdot y)$ mit $D' \in \vartheta$.

Satz 2.9:

Ist $(R,+,\cdot)$ Ringoid, $(\mathcal{R}R, \boxplus, \boxdot)$ Rasterringoid von $(R,+,\cdot)$ und gilt mit der Rundung $\square: R \to \mathcal{R}R$ die Beziehung:

(V) $$\bigwedge_{o \in \{+,\cdot\}} \bigvee_{a,b \in \mathcal{R}R} a \boxed{o} b = \square(a \circ b),$$

und ist ϑ verträglich mit dem Ringoid R und rechtsverträglich mit der Rundung $\square$, so ist ϑ ein mit dem Ringoid - Rasterringoid $R \times \mathcal{R}R$ verträgliches prämetrisches Feld, d.h. es gilt (PF3), (PF4) und insbesondere (PF31),(PF32),(PF41) und (PF42).

Bew.: Nach Voraussetzung gelten die Voraussetzungen des Lemmas 2.8, so daß ϑ rechtsverträglich mit $(\mathcal{R}R, \boxplus, \boxdot)$ bezüglich $(R,+,\cdot)$ ist. Somit gelten (PF32) und (PF42).

Für beliebige $x \in R$, $a,y \in \mathcal{R}R$ folgt daher:

(PF3) $D(a + x, a \boxplus y) \overset{(PF32)}{\subseteq} D'(a + x, a + y) \subseteq$

$\overset{(PF31)}{\subseteq} D'(a,a)+D'(x,y)-\overline{D'(o,o)}$

$$\text{(PF4)}\quad D(a \cdot x, a \boxdot y) \overset{\text{(PF42)}}{\subseteq} D'(a \cdot x, a \cdot y) \subseteq$$
$$\overset{\text{(PF41)}}{\subseteq} p(a,o) * (D'(x,y) - \overline{D'(o,o)}) + D'(o,o)$$

Nach diesen vorbereitenden Sätzen können nun einfache Kriterien aufgestellt werden, unter denen stets ein quasilineares prämetrisches Feld leicht angebbar ist, das zugleich verträglich mit dem zugrundegelegten Ringoid - Rasterringoid ist.

Definition 2.10:

Ist (M,p) prämetrischer Raum mit Prämetrik $p : M \times M \to V_n\mathbb{IR}$, $(\mathcal{R}M, \leq)$ Raster von $(M, \leq)$ und $\square : M \to \mathcal{R}M$ eine Rundung, so heißt $\square$ quasilinear isoton beschränkt bezüglich einem vorgegebenen Element $O \in M$, wenn eine Funktion $\phi \in \mathcal{L}_n$ existiert, mit

$$\text{(R6)}\quad \bigwedge_{y \in M} p(O, \square y) \subseteq \phi p(O,y)$$

Der folgende Satz gibt nun konstruktiv an, wie in Ringoiden - Rasterringoiden unter der Bedingung (V) und (R6) stets ein prämetrisches Feld zu erhalten ist:

Satz 2.11:

Ist $(R, +, \cdot, p, \leq)$ prämetrisches Ringoid, $(\mathcal{R}R, \boxplus, \boxdot, \leq)$ Rasterringoid von R mit

$$\text{(V)}\quad \bigwedge_{o \in \{+,\cdot\}} \bigwedge_{a,b \in \mathcal{R}R} a \,\boxed{o}\, b = \square(a \circ b)$$

und den ausgezeichneten Elementen $\{-e,0,e\}$ und ist die Rundung

$\square : R \to \mathcal{R}R$ quasilinear isoton beschränkt nach (R6) bezüglich $0 \in R$ und $\phi \in \mathcal{L}_n$,

so ist

$$\vartheta^* := \{D_i \mid D_i(x,y) := p(x,0) + \delta\phi^{(i)} p(0,y) \;\wedge\; i \in \mathbb{N}_0\}$$

ein quasilineares prämetrisches Feld, das mit dem Ringoid - Rasterringoid $R \times \mathcal{R}R$ verträglich ist.

Es gelten zusammenfassend folgende Eigenschaften:

(PF1) $p \in \vartheta^*$

(PF2) $D_i(x,z) = p(x,y) + D_i(y,z)$

(PF3) $D_i(a + x, a \boxplus y) \subseteq D_{i+1}(a,a) + D_{i+1}(x,y) - \overline{D_{i+1}(o,o)}$

(PF4) $D_i(a \cdot x, a \boxdot y) \subseteq p(a,o) * (D_{i+1}(x,y) - \overline{D_{i+1}(o,o)}) + D_{i+1}(o,o)$

(PF5) $D_i(x,y) = p(x,o) + \delta\phi^{(i)} p(o,y)$, $(\delta\phi^{(i)} \in \mathcal{L}_n)$

(PF6) $D_i(x, \square y) \subseteq D_{i+1}(x,y)$

<u>Bemerkung:</u>

Darüber hinaus gelten natürlich nach Bemerkung 2.4 die Eigenschaften $(PF3^*)$, $(PF4^*)$ und nach Satz 2.9 die Eigenschaften (PF31),(PF32),(PF41) und (PF42).

<u>Bew.:</u>

(PF1) $p \in \vartheta^*$ mit $i = o$ und (R,p) ist nach Voraussetzung prämetrischer Raum.

(PF2) $D_i(x,z) = p(x,o) + \delta\phi^{(i)}p(o,z) =$ nach (P3)

$$= p(x,y) + p(y,o) + \delta\phi^{(i)}p(o,z) =$$

$$= p(x,y) + D_i(y,z) \ .$$

Zum Nachweis der Verträglichkeit mit dem Ringoid - Rasterringoid genügt es nach Satz 2.9, die in Definition 2.7 eingeführten Eigenschaften (PF31), (PF41) und (PF6) nachzuweisen.

Es gilt (mit $\psi_i = \delta\phi^{(i)} \in \mathcal{L}_n$):

(PF31) $D_i(a + x,a + y) = p(a + x,o) + \psi_i p(o,a+y) =$

nach Bemerkung (2) zu Definition 1.4:

$$= p(a,o) + p(x,o) + \psi_i(p(o,a) + p(o,y))$$

nach Satz 2.6(4) $\subseteq p(a,o) + p(x,o) +$

$$\psi_i(p(o,a)) + \psi_i(p(o,y)) - \overline{\psi_i(o)} =$$

$$= D_i(a,a) + D_i(x,y) - \overline{D_i(o,o)}$$

(PF41) $D_i(a \cdot x,a \cdot y) = p(a \cdot x,o) + \psi_i p(o,a \cdot y) =$

nach Bemerkung (2) zu Definition 1.4

$$= p(a,o)p(x,o) + \psi_i(p(a,o)p(o,y)$$

nach Satz 2.6(4) $\subseteq p(a,o)p(x,o) +$

$$p(a,o) * (\psi_i(p(o,y)) - \overline{\psi_i(o)}) + \psi_i(o)$$

$$\overset{1)}{=} p(a,o)(p(x,o) + \psi_i(p(o,y)) - \overline{\psi_i(o)}) + \psi_i(o)$$

$$= p(a,o)\ (D_i(x,y) - \overline{D_i(o,o)}) + D_i(o,o)$$

---

1) Für $\alpha \in R$, $A,B \in V_n IIR$ gilt $\alpha(A+B) = \alpha * A + \alpha * B$ (Semidistributivgesetz in [Kulisch 69])

(PF6) $D_i(x, \Box y) = p(x,o) + \delta\phi^{(i)} p(o, \Box y)$

$$\overset{(R6)}{\subseteq} p(x,o) + \delta\phi^{(i)}(\phi(p(o,y)))$$

$$\overset{\text{Satz 2.6(5)}}{\subseteq} p(x,o) + \delta\phi^{(i+1)} p(o,y) =$$

$$= D_{i+1}(x,y)$$

(PF5) $\mathcal{V}^*$ ist quasilinear per Definition, denn es ist $\delta\phi^{(i)} \in \mathcal{L}_n$ und es ist

$$D_i(x,y) = p(x,y) + \delta\phi^{(i)} p(o,y) \ .$$

2.4 Einige spezielle prämetrische Felder in $V_n\mathbb{R}^* \times V_n\mathcal{R}\mathbb{R}^*$ und Anwendungsbeispiele bei der Rundungsfehleranalyse

2.4.1 Prämetrische Felder über $\mathbb{R}^* \times \mathcal{R}\mathbb{R}^*$

Die reellen Zahlen $\mathbb{R}$ bilden einen Körper bezüglich + und $\cdot$. Nach verbandstheoretischer Vervollständigung nach I.4 ist $\mathbb{R}^*$ ein Ringoid.

Ferner ist $p(x,y) = x - y$ eine Prämetrik, denn es gilt für $a,b,x,y,z \in R$:

(P1) $x - y = o \iff x = y$

(P2) $x - y + y - x = o$

(P3) $x - z = x - y + y - z$

(P4) $x \leq y \leq z \implies o \leq y - x \leq z - x$

(P5) $a + x - (a + y) = x + y$ $\qquad a,x,y \neq -p,p$

(P6) $a + x - (b + y) = a - b + x - y$ $\qquad a,b,x,y \neq -p,p$

(P7) $a \cdot x - a \cdot y = a \cdot (x - y) = (a - o)\,(x - y)$ $\quad a,x,y \neq -p,p$

(P8) $1 - o = o - (-1) = 1$

Es ist also $(\mathbb{R}^*, p)$ ein prämetrisches Ringoid mit $p(x,o) = x$ und $p(o,y) = -y$.

Zur Erfüllung der weiteren Voraussetzungen von Satz 2.11 mögen sich die folgenden Betrachtungen nur auf die am häufigsten realisierten Raster der Gleitkommazahlen $\mathcal{R}_b^\tau\mathbb{R}^*$ (Basis $b \in \mathbb{N}$ und τ Mantissenstellen) und das Raster der Festkommazahlen $F\mathcal{R}_b^\tau\mathbb{R}^*$ beschränken.

Der Einfachheit halber wird der Überlauf nicht berücksichtigt, d.h. die Forderung (R6) an die Rundung wird nur für die Punkte $|y| \leq p$ gefordert. Aus diesem Grunde können die betrachteten Raster als unbeschränkt angenommen werden.

2.4.1.1 Das Gleitkommazahlenraster $\mathcal{R}_b^\tau\mathbb{R}^*$ und zugehörige Rundungen

<u>Definition 3.1:</u> $\mathcal{R}_b^\tau\mathbb{R}^*$ ist Gleitkommazahlenraster der reellen Zahlen mit der Darstellung:

$$\bigwedge_{x \in \mathcal{R}_b^\tau\mathbb{R}^*} x = m\, b^t$$

wobei $m \in \mathbb{Z}$ und $|m| \leq b^\tau - 1$ und τ ist Stellenzahl der Mantisse m und $t \in \mathbb{Z}$ ist Exponent der Basis $b \in \mathbb{N}$.

Die oben definierten Gleitkommazahlen haben folgende Eigenschaften:

- $\mathcal{R}_b^\tau\mathbb{R}^*$ ist abzählbar
- $\mathcal{R}_b^\tau\mathbb{R}^*$ ist bezüglich der natürlichen Ordnungsrelation linear geordnet
- $\mathcal{R}_b^\tau\mathbb{R}^*$ ist ein symmetrisches Raster ([Kulisch 72]) .

Die üblichen Rundungen $\square : R \to \mathcal{R}_b^\tau\mathbb{R}^*$ lassen sich in folgender Weise beschreiben:

Seien $\nabla : \mathbb{R} \to \mathcal{R}_b^\tau\mathbb{R}^*$ bzw. $\Delta : R \to \mathcal{R}_b^\tau\mathbb{R}^*$ die nach unten bzw. nach oben gerichteten Rundungen wie sie in [Kulisch 72] definiert sind gemäß:

$$\bigwedge_{x \in \mathbb{R}} \nabla x := i(L_{\mathcal{R}\mathbb{R}}(x)) \quad \text{bzw.} \quad \bigwedge_{x \in \mathbb{R}} \Delta x := o(U_{\mathcal{R}\mathbb{R}}(x))$$

Zur Bezeichnung vergleiche Abschnitt I,1.

Damit kann folgende Zwischenstellenfunktion definiert werden:
μ sei eine für die vorgesehene Rundung fest vorgegebene ganze Zahl mit $o \leq \mu \leq b$, so ist

$$\bigwedge_{x \in \mathbb{R}} M(x) := \frac{\Delta x - \nabla x}{b}\mu + \nabla x \; .$$

Dann lassen sich die üblichen monotonen optimalen Rundungen

$\square_\mu : \mathbb{R} \to \mathcal{R}^{T}_{b}\mathbb{R}$ mit

$$\bigwedge_{x \in \mathbb{R}} \quad \square_\mu(x) := \begin{cases} \nabla x & \text{für alle } x \in [\nabla x, M(x)] \\ \Delta x & \text{für alle } x \quad [M(x), \Delta x] \end{cases}$$

darstellen. Es ist unerheblich, ob die speziellen Zahlen M(x) selbst nach unten oder nach oben gerundet werden.

Wie man leicht verifizieren kann gilt speziell:

$\square_o(x) = \Delta x$ und $\square_b(x) = \nabla x$.

Die Eigenschaften der oben definierten Rundungen sind in [Kulisch 72] nachgewiesen mit (R1) Monotonie, (R2) Optimalität und der folgende Satz 2.15 wird zeigen, daß sie auch die Eigenschaft der quasilinearen isotonen Beschränktheit (R6) besitzen.

Da künftig stets auch der Unterlauf berücksichtigt werden soll, möge folgender allgemeine Unterlauf angenommen sein:

<u>Definition 2.14:</u>

Die Rundung $\square_{\mu,\Theta}$ ist eine optimale Rundung mit Unterlauf bezüglich des Unterlaufintervalles

$\Theta := [-\theta, \theta]$, $\theta \in \mathbb{R}^+$, wenn gilt:

$$\square_{\mu,\Theta}(x) := \begin{cases} \square_\mu(x) & \text{für } x \in \mathbb{R} \setminus \Theta \\ \omega \in \Omega \subseteq [-m,m] & \text{für } x \in \Theta \text{ und} \end{cases}$$

m = kleinste positive Maschinenzahl und $\square_{\mu,\Theta}(\omega) = \omega$, $\Omega = -\Omega$.

<u>Lemma 2.15:</u>

Im symmetrischen Raster $\mathfrak{R}^{\tau}_{b}\mathbb{R}^{*}$ gilt:

(1) $$\bigwedge_{x \in \mathbb{R}} \square_{\mu}(-x) = - \square_{b-\mu}(x)$$

(2) $$\bigwedge_{x \in \mathbb{R}} \square_{\mu,\Theta}(-x) = - \square_{b-\mu,\Theta}(x)$$

<u>Bew.:</u>

Zunächst möge gezeigt werden, daß für

$M_{\mu}(x) = (\Delta x - \nabla x)\frac{\mu}{b} + \nabla x$ gilt:

$M_{\mu}(-x) = -M_{b-\mu}(x)$.

Es ist

$$M_{\mu}(-x) = (\Delta(-x) - \nabla(-x))\frac{\mu}{b} + \nabla(-x) =$$

[Kulisch 72]
$$= (-\nabla x + \Delta x)\frac{\mu}{b} - \Delta x =$$

$$= \frac{-\nabla x\mu + \Delta x\mu - \Delta x b}{b} + (\nabla x - \nabla x) =$$

$$= \frac{-\nabla x(\mu - b) + \Delta x(\mu - b)}{b} - \nabla x =$$

$$= -\left\{(\Delta x - \nabla x)\frac{(b-\mu)}{b} + \nabla x\right\} = -M_{b-\mu}(x)$$

Damit folgt für (1):

$$\square_{\mu}(-x) = \begin{cases} \nabla(-x) & \text{für} \quad -x \in [\nabla(-x), M_{\mu}(-x)] \\ \\ \Delta(-x) & \text{für} \quad -x \in [M_{\mu}(-x),\ \Delta(-x)] \end{cases}$$

$$= \begin{cases} -\Delta x & \text{für} \quad x \in [M_{b-\mu}(x),\ \Delta x] \\ \\ -\nabla x & \text{für} \quad x \in [\nabla x,\ M_{b-\mu}(x)] \end{cases}$$

$$= \quad -\square_{b-\mu}(x)\ .$$

(2) folgt auf Grund der Symmetrie von R, Θ, Ω:

$$\square_{\mu,\Theta}(-x) = \begin{cases} \square_{\mu}(-x) & \text{für} \quad x \in R \setminus \Theta \\ \\ \omega \in \Omega & \text{für} \quad x \in \Theta \end{cases}$$

$$= \begin{cases} -\square_{b-\mu}(x) & \text{für} \quad x \in R \setminus \Theta \\ \\ -\omega \in \Omega & \text{für} \quad x \in \Theta \end{cases}$$

$$= \quad -\square_{b-\mu,\Theta}(x)$$

Satz 2.16:

Die in Definition 2.14 gegebenen Rundungen sind quasilinear isoton beschränkt nach (R6) bezüglich der Prämetrik $p(x,y) := x - y$ in $\mathbb{R}^*$.

Es gilt insbesondere:

$$\bigwedge_{x \in \mathbb{R}} p(o, \square_{\mu,\Theta} x) \subseteq \mathfrak{A} * p(o,x) + B \quad \text{mit}$$

$$\mathfrak{A} := \frac{1 \pm [o, b^{1-\tau}]}{1 + K\, b^{-\tau}} \quad \text{mit} \quad K := [\mu \sqcap (b-\mu), \mu \sqcup (b-\mu)]$$

und

$$B := \overline{\mathfrak{A}} * \Theta + \Omega = \lambda(\mathfrak{A}) * \Theta + \Omega$$

Bew.:

Zum Beweis wird zunächst die Rundung $\square_\mu = \square_{\mu,o}$ (ohne Berücksichtigung des Unterlaufes) abgeschätzt durch eine Ober- und Untergerade. Siehe Abbildung 2.1.

Mit $a_1, a_2 \in \mathbb{R}$, $a_1 \leq a_2$ soll gelten:

$$\bigwedge_{x \in \mathbb{R}^+} a_1 \cdot x \leq \square_\mu(x) \leq a_2 \cdot x \;\wedge\; \bigwedge_{x \in \mathbb{R}^-} a_2 \cdot x \leq \square_\mu(x) \leq a_1 \cdot x$$

oder unter Ausnutzung der Intervallarithmetik:

$$\bigwedge_{x \in \mathbb{R}} \quad \square_\mu(x) \subseteq [a_1, a_2] * x \quad .$$

Auf Grund von Lemma 2.15 genügt es zunächst, für positive x eine Abschätzung zu finden:

$$\bigwedge_{x \in \mathbb{R}^+} \quad \square_\mu(x) \subseteq \mathfrak{A}_\mu * x \qquad \text{mit } \mathfrak{A} \in \mathbb{IR}$$

Zur Konstruktion dieser Abschätzung definiere für $y \in \mathcal{R}_b^\tau \mathbb{R}^*$ $\nu(y)$ den Nachfolger von y mit $y \leq \nu(y)$ und y ist unmittelbarer Nachbar von $\nu(y)$ bezüglich der Ordnungsrelation in $\mathbb{R}$. Es ist z. B.

$$\nu(b^t) = b^t + b^{t-\tau+1} = b^t(1+b^{1-\tau}) \quad .$$

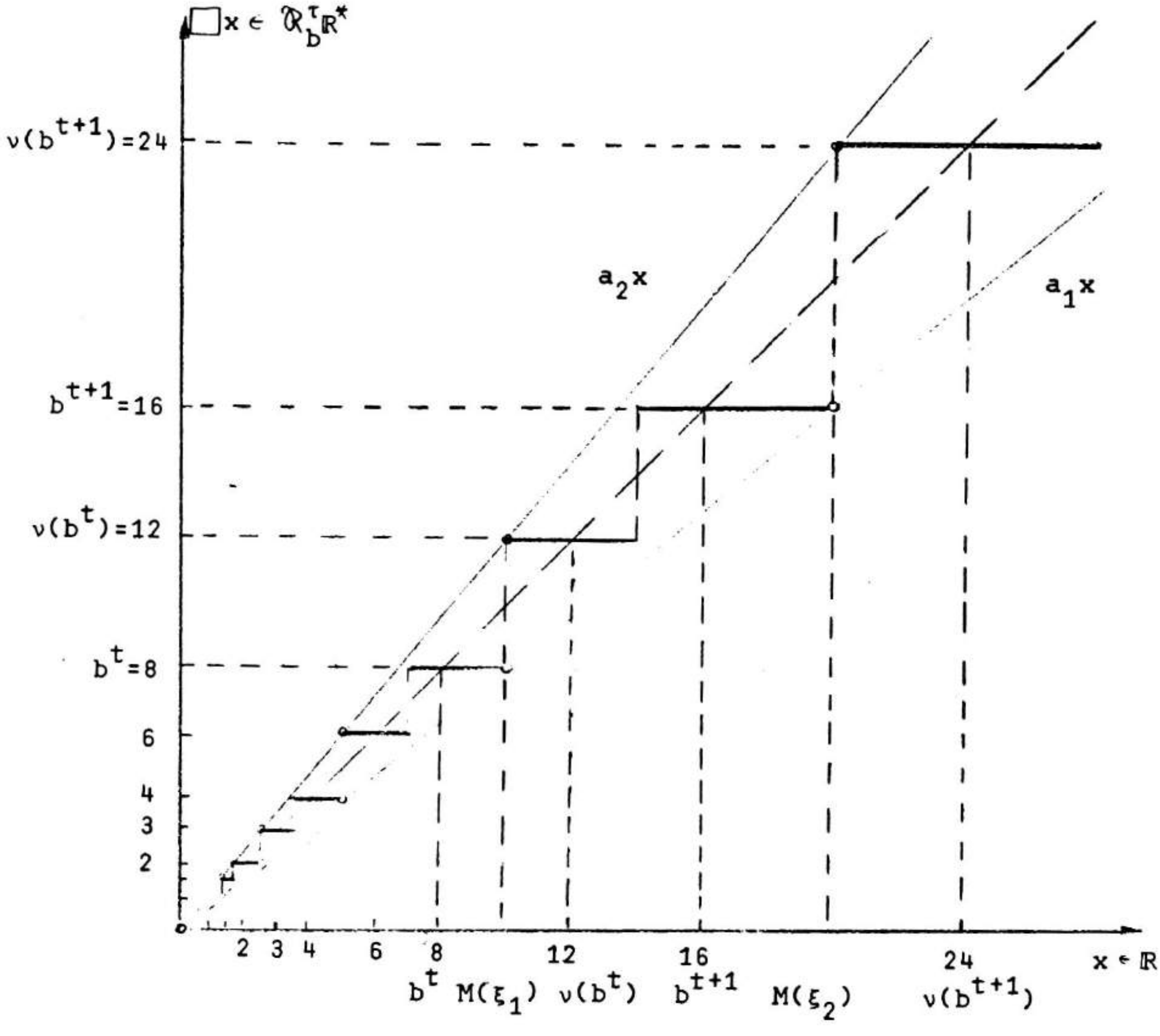


Abb. 2.1

(Für die Zeichnung gilt speziell: $b = 2$, $\tau = 2$, $\mu = b/2 = 1$,

also $\mathcal{R}_2^2\mathbb{R} = \{\ldots,10,11,100,110,1000,1100,10\,000,11\,000,\ldots\}$(dual)

$= \{\ldots,2\ ,3\ ,\ 4\ ,\ 6\ ,\ 8\ ,\ 12\ ,\ 16\ ,\ 24\ ,\ldots\}$(dezimal)

womit schließlich die Steigungen $a_1 = 4/5$ und $a_2 = 6/5$ berechnet werden.)

Die für die Abschätzung kritischen Stellen liegen, wie die Abbildung 2.1 zeigt, in den Sprungstellen $M(x)$. Eine weitere Betrachtung zeigt, daß unter den Sprungstellen diejenigen ausschlaggebend sind, die im Intervall $[b^t, \nu(b^t)]$ liegen. Berechnet man, wie in der Abbildung gezeigt, die Steigung der Obergeraden:

$$a_2 = \frac{\nu(b^t) - \nu(b^{t-1})}{M(\xi_2) - M(\xi_1)} \quad \text{mit} \quad \xi_1 \in [b^{t-1}, \nu(b^{t-1})] \quad \text{und}$$

$$\xi_2 \in [b^t \;, \nu(b^t)] \quad \text{ist}$$

$$M(\xi_2) = \frac{\nu(b^t) - b^t}{b} + b^t$$

$$= b^t + b^{t-\tau} \quad \text{und entsprechend ist}$$

$$M(\xi_1) = b^{t-1} + b^{t-\tau-1} \quad \text{womit schließlich}$$

$$a_2 = \frac{b^t(1 + b^{1-\tau}) - b^{t-1}(1 + b^{1-\tau})}{b^t(1 + \mu b^{t-\tau}) - b^{t-1}(1 + \mu b^{-\tau})}$$

$$= \frac{b^{t-1}(b-1)(1 + b^{1-\tau})}{b^{t-1}(b-1)(1 + \mu b^{-})} = \frac{1 + b^{1-\tau}}{1 + \mu b^{-\tau}} \qquad \text{folgt.}$$

$a_2(t) = a_2$ ist vom Exponenten unabhängig, d.h. die Steigung ist für alle t gleich und es ist daher für $x \in \mathbb{R}^+$

$$\sqsubset_{\mu} x \leq a_2 x = \frac{1 + b^{1-\tau}}{1 + \mu b^{-\tau}} x$$

In entsprechender Weise folgt für a_1 (siehe Abbildung 2.1):

$$a_1 = \frac{b^t - b^{t-1}}{M(\xi_2) - M(\xi_1)} = \frac{b^{t-1}(b-1)}{b^{t-1}(b-1)(1 + \mu b^{-\tau})} = \frac{1}{1 + \mu b^{-\tau}}$$

und damit für $x \in \mathbb{R}^+$

$$x \geq a_1 x = \frac{1}{1 + \mu b^{-\tau}} x \quad .$$

Damit ist für $x \in \mathbb{R}^+$ die Abschätzung

$$(*) \quad \square_\mu(x) \subseteq [a_1, a_2]_\mu * x = \mathfrak{A}_\mu * x = \frac{1 + [o, b^{1-\tau}]}{1 + \mu b^{-\tau}} * x$$

gewonnen.

Für $-x \in \mathbb{R}^-$ ist $x \in \mathbb{R}^+$ und so folgt nach Lemma 2.15 und nach ($*$) die Abschätzung:

$$\square_\mu(-x) = -\square_{b-\mu}(x) \subseteq -(\mathfrak{A}_{b-\mu} * x) = \mathfrak{A}_{b-\mu} * (-x) \quad .$$

Soll also ein Intervall $\mathfrak{A}$ gesucht werden, für das gilt:

$$\bigwedge_{x \in \mathbb{R}} \square_\mu(x) \subseteq \mathfrak{A} * x \quad ,$$ so folgt für

$$x \geq o : \quad \mathfrak{A}_\mu * x \subseteq \mathfrak{A} * x \quad \text{und für}$$

$$x \leq o : \quad \mathfrak{A}_{b-\mu} * x \subseteq \mathfrak{A} * x \quad .$$

Diese Bedingungen sind äquivalent mit

$$\mathfrak{A}_\mu \subseteq \mathfrak{A} \quad \wedge \quad \mathfrak{A}_{b-\mu} \subseteq \mathfrak{A}$$

und nach Satz III,2.42 (3) und (6) gilt dies sicher für

$\mathfrak{A} := \mathfrak{A}_\mu \mathbin{\underline{\cup}} \mathfrak{A}_{b-\mu}$.

Nach (*) ist dann

$$\mathfrak{A} = \frac{1 + [o,b^{1-\tau}]}{1 + \mu b^{-\tau}} \mathbin{\underline{\cup}} \frac{1 + [o,b^{1-\tau}]}{1 + (b-\mu)b^{-\tau}}$$

Satz III,2.44(1)

$$= (1 + [o,b^{1-\tau}]) * \{1/(1+\mu b^{-\tau}) \mathbin{\underline{\cup}} 1/(1+(b-\mu)b^{-\tau})\}$$

Satz III,2.44(3)

$$= (1+[o,b^{1-\tau}]) * \frac{1}{(1+\mu b^{-\tau}) \mathbin{\underline{\cup}} (1+(b-\mu)b^{-\tau})}$$

Satz III,2.44(1)

$$= (1 + [o,b^{1-\tau}]) * \frac{1}{1 + (\mu \mathbin{\underline{\cup}} (b-\mu))b^{-\tau}}$$

Nach Definition III,2.41 ist

$$\mu \mathbin{\underline{\cup}} (b-\mu) = [\mu \sqcap (b-\mu),\ \mu \sqcup (b-\mu)] =: K \ .$$

Damit ist für $x \in \mathbb{R}$

$$\square_\mu(x) \subseteq \mathfrak{A} * x = \frac{1 + [o,b^{1-\tau}]}{1 + K\, b^{-\tau}} * x \ .$$

Soll nun der Fall des Unterlaufes miteinbezogen werden, d. h. für alle $x \in \Theta$ ist $\square_{\mu,\Theta}(x) \subseteq \Omega$ so führt der Ansatz

$$\square_{\mu,\Theta}(x) \subseteq \mathfrak{A} * x + B \qquad \text{mit} \quad \mathfrak{A} = \frac{1 + [o, b^{1-\tau}]}{1 + Kb^{-\tau}}$$

zu der Bedingung:

$$\bigwedge_{\xi \in \Theta} \quad \square_{\mu,\Theta}(\xi) \subseteq \Omega \subseteq \mathfrak{A} * \xi + B$$

Daraus folgt nach Satz III,2.9

$$\bigwedge_{\xi \in \Theta} \quad B \supseteq \Omega - \bar{\mathfrak{A}} * \xi \quad , \text{ nach Satz III,2.42(6)}$$

$$B \supseteq \bigcup_{\xi \in \Theta} (\Omega - \bar{\mathfrak{A}} * \xi), \text{ nach Satz III,2.51(1)}$$

$$B \supseteq \Omega - \bar{\mathfrak{A}} * \Theta \qquad , \text{ mit } -\Theta = \Theta \quad \text{schließlich}$$

$$B \supseteq \Omega + \bar{\mathfrak{A}} * \Theta$$

und nach Tabelle III, 2.5 mit $o \leq a_1 = \lambda(\mathfrak{A})$

$$B \supseteq \Omega + \rho(\bar{\mathfrak{A}}) * \Theta = \Omega + \lambda(\mathfrak{A}) * \Theta$$

Es genügt daher $B := \Omega + \bar{\mathfrak{A}} * \Theta = \dfrac{\Theta}{1 + \rho K b^{-\tau}} + \Omega$ zu setzen.

Nun bedarf es noch des Nachweises, daß

$\mathfrak{A} * X + B =: \Phi(X) \in \mathcal{L}_n$, d. h., daß gilt:

$1 \subseteq \mathfrak{A}\,,\quad o < \lambda\mathfrak{A} = a_1$ und $o \subseteq B$.

Tatsächlich ist mit $o \leq \mu \leq b \quad \rho(K) = \mu \sqcup (b-\mu) \geq o$ und somit

$$o \leq \lambda\mathfrak{A} = \frac{1}{1 + \rho(K)b^{-\tau}}$$

und es ist $o \leq \mu \sqcap (b-\mu) \quad \wedge \quad \mu \sqcup (b-\mu) \leq b$

und damit nach Folgerung (6) im Anschluß von Satz III,2.6 und Definition III,2.5:

$$\overline{K} \subsetneq K \subseteq [o,b] \qquad \Longrightarrow$$

$$\overline{K}b^{-\tau} \subseteq [o,b^{1-\tau}] \qquad \Longrightarrow$$

$$1 + \overline{K}b^{-\tau} \subseteq 1 + [o,b^{1-\tau}] \qquad \Longrightarrow \qquad \text{nach Satz III,2.26(1)}$$

$$1 \subseteq \frac{1 + [o,b^{1-\tau}]}{1 + Kb^{-\tau}} = \mathfrak{A}\ .$$

Ferner ist mit $o \subseteq \Theta$ stets $o \subseteq \overline{\mathfrak{A}} * \Theta$ und mit $\Omega = -\Omega$ ist $o \subseteq \Omega$ und somit insgesamt $o \subseteq \overline{\mathfrak{A}} * \Theta + \Omega = B$.

Mit $p(o, \Box_{\mu,\Theta}x) = - \Box_{\mu,\Theta}(x) \subseteq \mathfrak{A} * (-x) + B = \mathfrak{A} * p(o,x) + B$

ist die Aussage von Satz 2.16 bewiesen.

2.4.1.2 Das Festpunktzahlenraster $F\mathcal{R}_b^\tau[-1,1]$

Definition 2.17: $F\mathcal{R}_b^\tau[-1,1]$ ist Festpunktzahlenraster mit

$$\bigwedge_{x \in F\mathcal{R}_b^\tau[-1,1]} |x| = o \cdot \xi_1 \xi_2 \xi_3 \ldots \xi_\tau \quad \text{mit} \quad \bigwedge_{1 \leq i \leq \tau} \xi_i \in N_o \,,$$

τ Mantissenstellen und Zahlbasis b .

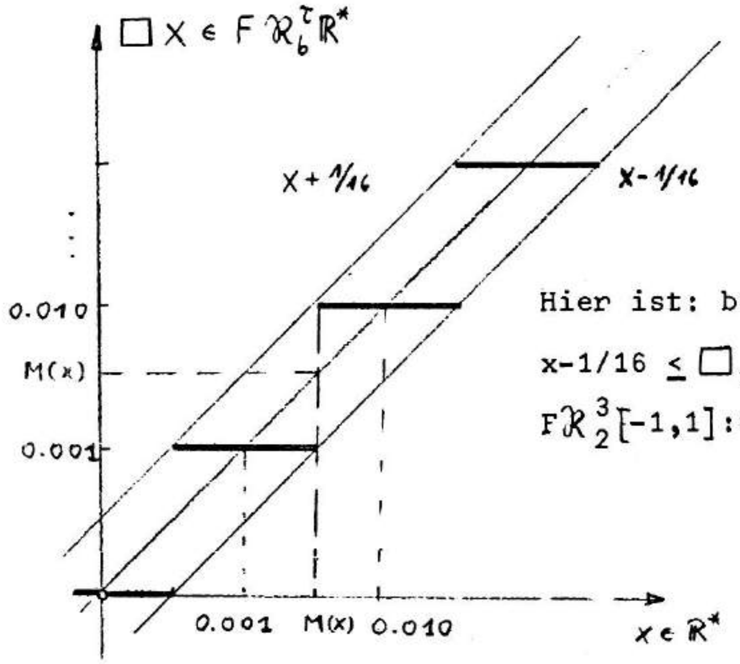
Die Eigenschaften (S1) und (S2) sind erfüllt als Teilmenge des Rasters der Gleitkommazahlen und $i(F\mathcal{R}_b^\tau[-1,1]) = p$ und $o(F\mathcal{R}_b^\tau[-1,1]) = -p$.

Das Raster ist nach Definition symmetrisch.

Die entsprechende Rundung ist mit den Bezeichnungen aus 2.4.1.1 die Restriktion der Rundung $\square_\mu : \mathbb{R} \to \mathcal{R}_b^\tau\mathbb{R}^*$ auf das Intervall $[-1,1]$:

$$\square_\mu : [-1,1] \to F\mathcal{R}_b^\tau[-1,1] \,.$$

Diese Rundung besitzt dieselben Eigenschaften (R1), (R2) und (R6). Die Abschätzung für die lineare Beschränktheit läßt sich jedoch wesentlich verschärfen. Wie eine kurze Betrachtung zeigt und wie aus Abbildung 2,2 zu ersehen ist, genügen diese Rundungen folgender Abschätzung:



Hier ist: $b=2$, $\tau=3$, $\mu=1$,

$x-1/16 \leq \square_1 x \leq x+1/16$

$F\mathcal{R}_2^3[-1,1] := \{\ldots, 0.000, 0.001, 0.010, \ldots$

$\ldots, 0.101, 0.110, 0.111\}$

Abb. 2.2

Für $x \in [-1,1]$ folgt für $\nabla x \leq x \leq M(x)$ die Abschätzung

$$x - \square_\mu(x) \leq M(x) - \square_\mu(x) = M(x) - \nabla x$$

und für $M(x) \leq x \leq \Delta x$ die Abschätzung

$$x - \square_\mu(x) \geq M(x) - \square_\mu(x) = M(x) - \Delta x \ .$$

Insgesamt also gilt die Abschätzung

$$M(x) - \Delta x \leq x - \square_\mu(x) \leq M(x) - \Delta x$$

oder mit Hilfe der Intervallrechnung

$$x - \square_\mu(x) \subseteq M(x) - [\nabla x, \Delta x]$$

und schließlich

$$-\square_{\mu}(x) \subseteq -x + M(x) - [\nabla x, \Delta x].$$

Nun ist nach Definition von $M(x)$ in Abschnitt 2.4.1.1

$$M(x) - \Delta x = (\Delta x - \nabla x)\frac{\mu}{b} + (\nabla x - \Delta x) = (\Delta x - \nabla x)(\frac{\mu}{b} - 1)$$

und

$$M(x) - \nabla x = (\Delta x - \nabla x)\frac{\mu}{b} \quad .$$

Damit erhält man die Abschätzung

$$-\square_{\mu}(x) \subseteq -x + (\Delta x - \nabla x) * [\frac{\mu}{b} - 1, \frac{\mu}{b}]$$

und mit $\Delta x - \nabla x = b^{-\tau}$ schließlich

$$p(o,\square_{\mu}(x) = -\square_{\mu}(x) \subseteq -x + [\frac{\mu}{b} - 1, \frac{\mu}{b}]b^{-\tau} = p(o,x) + B$$

mit $B := [\frac{\mu}{b} - 1, \frac{\mu}{b}]b^{-\tau}$.

Die so definierte Funktion $\phi(X) = \mathfrak{A} * X + B$ mit $\mathfrak{A} = 1$ erfüllt alle Eigenschaften aus Definition 2.1, denn es ist mit $\mu \leq b$ $\frac{\mu}{b} \leq 1$, also $\frac{\mu}{b} - 1 \leq o \leq \frac{\mu}{b}$, $o \subseteq B$.

Wie Abschnitt 2.4.1.1 und 2.4.1.2 zeigten, sind die Rundungen, die sowohl bei der Gleitkommarechnung als auch bei der Festkommarechnung auftreten nach (R6) quasilinear beschränkt. Im ersten Fall ist mit $\phi(X) = \mathfrak{A} * X + B$

$$\mathfrak{A} := \frac{1 + [o, b^{1-\tau}]}{1 + K\, b^{-\tau}} , \qquad K := [\mu \sqcap (b-\mu),\ \mu \sqcup (b-\mu)]$$

$$B := \overline{\mathfrak{A}} * \Theta + \Omega$$

und im zweiten Fall speziell

$\mathfrak{A} := 1$ und $B := [\mu - b, \mu] b^{-\tau-1}$.

Ferner ist $(\mathbb{R}^*, +, \cdot, p, \leq)$ ein prämetrisches Ringoid mit den ausgezeichneten Elementen $\{-1, o, 1\}$ und mit

$$(V) \quad \bigwedge_{o \in \{+,\cdot\}} \quad \bigwedge_{a,b \in \mathcal{R}^{\tau}_{b}\mathbb{R}^*} \quad a \boxed{o} b := \square_{\mu,\theta}(a \circ b)$$

ist nach Satz I,1.11 bei symmetrischer Rundung [1] (d.h. $\mu = \frac{b}{2}$) $(\mathcal{R}^{\tau}_{b}\mathbb{R}^*, \boxplus, \boxdot, \leq)$ ein Rasterringoid.

Damit sind alle Voraussetzungen des Satzes 2.11 erfüllt und man erhält die beiden prämetrischen Felder:

<u>Lemma 2.18:</u>

Es sind im Raster der Gleitkommazahlen $\mathcal{R}^{\tau}_{b}\mathbb{R}^*$:

$$\vartheta^* := \{D_i \mid D_i(x,y) := p(x,o) + \mathfrak{A}^{(i)} * p(o,y) + \sum_{\nu=o}^{i-1} \mathfrak{A}^{(\nu)} * B \;\wedge\; i \in \mathbb{N}_o\}$$

mit $\mathfrak{A} := \dfrac{1 + [o, b^{1-\tau}]}{1 + \frac{1}{2} b^{1-\tau}}$ und $B = \dfrac{\theta}{1 + \frac{1}{2} b^{1-\tau}} + \Omega$

---

1) Diese Forderung ist eigentlich unnötig, da die Eigenschaften (D1) bis (D4) aus Definition I,1.6 für die Problemstellung dieser Arbeit voll ausreichen. (D1) bis (D4) gelten aber auch bei unsymmetrischer Rundung, so daß prinzipiell eine Art "schwaches Rasterringoid" für diesen Rahmen ausreichend wäre.

und im Raster der Festpunktzahlen $F\mathcal{R}_b^\tau[-1,1]$:

$$\vartheta^* := \{D_i \mid D_i(x,y) := p(x,y) + iB \quad \wedge \quad i \in \mathbb{N}_0\}$$

mit $B := \frac{[-1,1]}{2} b^{-\tau}$ mit dem Ringoid-Rasterringoid $\mathbb{R}^* \times \mathcal{R}_b^\tau \mathbb{R}^*$ verträgliche quasilineare prämetrische Felder.

2.4.1.3 Beispiele von Rundungsfehlerberechnungen mit Hilfe des prämetrischen Feldes

Folgende Beispiele zeigen inwieweit sich Rundungsfehlerabschätzungen in prämetrischen Feldern mit denen der klassischen Methoden decken.

Beispiel 2.20: Die Gleitkommasumme mit und ohne Berücksichtigung eines Unterlaufes

Es sei das Raster der Dualzahlen $\mathcal{R}_2^\tau \mathbb{R}^*$ mit Basis $b = 2$ und der üblichen Rundung $\square_{\mu,\Theta}$ mit $\mu = b/2 = 1$ und einmal mit $\Omega \neq o$, $\Theta = o$ und schließlich speziell $\Theta = o$, $\Omega = o$ betrachtet.

In der Schreibweise von [Wilkinson 69] ist $g\,l\,(x + y) = x \boxplus y$ in der hier gebrauchten Darstellungsweise im Rasterringoid $(\mathcal{R}_2^\tau \mathbb{R}^*, \boxplus, \boxdot)$.

Es wird nun in der Vorwärtsanalyse der Rundungsfehler $s - \tilde{s}$ berechnet, der bei der Summe $s = (\ldots(x_1 + x_2) + \ldots) + x_n$ im Raster $\mathcal{R}_2^\tau \mathbb{R}^*$ durch Berechnung von $\tilde{s} = (\ldots(x_1 \boxplus x_2) + \ldots) \boxplus x_n$

entsteht.

In dem nach Lemma 2.18 angegebenen prämetrischen Feld mit Prämetrik $p(x,y) = x - y$ folgt für die Differenz:

$$(((\ldots(x_1+x_2)+\ldots)+x_{n-2})+x_{n-1})+x_n-((\ldots(x_1 \boxplus x_2) \boxplus \ldots) \boxplus x_{n-1}) \boxplus x_n =$$

$$p(((\ldots(x_1+x_2)+\ldots)+x_{n-1})+x_n,\ ((\ldots(x_1 \boxplus x_2) \boxplus \ \ldots) \boxplus x_{n-1}) \boxplus x_n)$$

$$= D_o\Big(\sum_{\nu=1}^{n-1} x_\nu + x_n,\quad \underset{\mu=1}{\overset{n-1}{\boxed{\Sigma}}} x_\mu \boxplus x_n\Big)$$

(PF3)

$$\subseteq D_1\Big(\sum_{\nu=1}^{n-2} x_\nu + x_{n-1},\ \underset{\mu=1}{\overset{n-2}{\boxed{\Sigma}}} x_\mu \boxplus\ x_{n-1}\Big) + \big(D_1(x_n,x_n) - \overline{D_1(o,o)}\big)$$

PF3)

$$\subseteq D_2\Big(\sum_{\nu=1}^{n-3} x_\nu + x_{n-2},\ \underset{\mu=1}{\overset{n-3}{\boxed{\Sigma}}} x_\mu \boxplus\ x_{n-2}\Big) + \big(D_2(x_{n-1},x_{n-1}) -$$

$$\overline{D_2(o,o)}\big) + \big(D_1(x_n,x_n) - \overline{D_1(o,o)}\big)$$

$$\subseteq$$

.
.
.
.

$$\subseteq \sum_{\nu=2}^{n} \big(D_{n+1-\nu}(x_\nu,x_\nu) - \overline{D_{n+1-\nu}(o,o)}\big) + D_{n-1}(x_1,x_1)$$

Nun gilt nach Lemma 2.18:

$$D_k(x_i,x_i) = (1- \mathfrak{A}^{(k)}) * x_i + \sum_{\nu=o}^{k-1} \mathfrak{A}^{(\nu)} * B \ ,$$

$$D_k(o,o) = \sum_{\nu=o}^{k+1} \mathfrak{A}^{(\nu)} * B \quad \text{und}$$

$$D_k(x_i,x_i) - \overline{D_k(o,o)} = (1- \mathfrak{A}^{(k)}) * x_i \quad .$$

Es gilt daher die Abschätzung:

$$(\mathrm{I})\quad p\Big(\sum_{\nu=1}^{n} x_\nu , \boxed{\Sigma}_{\mu=1}^{n} x_\mu\Big) \subseteq \sum_{\nu=2}^{n} (1-\mathfrak{A}^{(n+1-\nu)}) * x_\nu + (1-\mathfrak{A}^{(n-1)}) * x_1 +$$

$$+ \sum_{\mu=0}^{n-2} \mathfrak{A}^{(\mu)} * B$$

oder umgeformt

$$\sum_{\nu=1}^{n} x_\nu - \boxed{\Sigma}_{\mu=1}^{n} x_\mu = \sum_{\nu=1}^{n} x_\nu - \sum_{\mu=2}^{n} \mathfrak{A}^{(n+1-\nu)} * x_\mu - \mathfrak{A}^{(n-1)} * x_1 +$$

$$+ \sum_{\mu=0}^{n-2} \mathfrak{A}^{(\mu)} * B$$

erhält man die Abschätzung:

$$(\mathrm{II})\quad \boxed{\Sigma}_{\mu=1}^{n} x_\mu \subseteq \sum_{\mu=2}^{n} \mathfrak{A}^{(n+1-\nu)} * x_\mu + \mathfrak{A}^{(n-1)} * x_1 + \sum_{\mu=0}^{n-2} \mathfrak{A}^{(\mu)} * B$$

Mit $b = 2$ und $\mu = 1$ folgt für $\mathfrak{A}$:

$$\mathfrak{A} = \Big[\frac{1}{1+2^{-\tau}} , \frac{1+2^{1-\tau}}{1+2^{-\tau}}\Big] \quad \text{und es ist}$$

$$\frac{1}{1+2^{-\tau}} \geq 1 - 2^{-\tau} \quad \text{und}$$

$$\frac{1+2^{1-\tau}}{1+2^{-\tau}} = \frac{2(1+2^{-\tau}) - 1}{1+2^{-\tau}} = 2 - \frac{1}{1+2^{-\tau}} \leq 2 - 1 + 2^{-\tau} =$$

$$= 1 + 2^{-\tau} ,$$

so daß für $\mathfrak{A}$ die einfache Abschätzung

$$\mathfrak{A} \leq \tilde{\mathfrak{A}} := [1 - 2^{-\tau}, 1 + 2^{-\tau}] \quad \text{gilt.}$$

Für $D_k(o,o)$ folgt nach Tabelle III,2.5 mit $o \subseteq B$:

$$\sum \mathfrak{A}^{(\nu)} * B \subseteq \sum \tilde{\mathfrak{A}}^{(\nu)} * B = \rho(\sum \tilde{\mathfrak{A}}^{(\nu)}) * B = \sum \rho(\tilde{\mathfrak{A}}^{(\nu)}) * B =$$
$$= \sum \rho(\tilde{\mathfrak{A}})^{\nu} * B \quad .$$

Mit $1 + 2^{-\tau} = \rho(\tilde{\mathfrak{A}}) =: q$ und

$$\sum_{\nu=o}^{k-1} q^{\nu} = \frac{q^k - 1}{q - 1} = \frac{(1+2^{-\tau})^k - 1}{2^{-\tau}}$$

und $B = \lambda \mathfrak{A} * \Theta + \Omega = (1-2^{-\tau}) * \Theta + \Omega$

folgt:

$$D_k(o,o) \subseteq ((1 + 2^{-\tau})^k - 1)((2^{\tau} - 1) * \Theta + \Omega\, 2^{-\tau}).$$

In den meisten Fällen ist $\Omega = o$, so daß sich also die Unterlaufsgröße im wesentlichen für $\tau \approx k$ in der Größenordnung von $\| D_k(o,o) \| \sim 2^{\tau} * \Theta$ beläuft. (Sie spielt nur dann eine größere Rolle, wenn sich die Summanden in Θ in größerer Zahl häufen.)

Für $\tilde{\mathfrak{A}}^{(k)} = [(1-2^{-\tau})^k,\ (1+2^{-\tau})^k]$ folgt aus (II) daher die Abschätzung:

$$\mathop{\boxed{\Sigma}}_{\mu=1}^{n} x_\mu = \sum_{\mu=2}^{n} (1 + \eta_{n+1-\mu})x_\mu + (1 + \eta_{n-1})x_1 + \vartheta_n$$

mit $(1-2^{-\tau})^{n+1-\mu} \leq 1+\eta_{n+1-\mu} \leq (1+2^{-\tau})^{n+1-\mu}$ für $\mu=2,\ldots,n$

und $(1-2^{-\tau})^{n-1} \leq 1 + \eta_{n-1} \leq (1 + 2^{-\tau})^{n-1}$

und $\vartheta_n \in ((1+2^{-\tau})^{n-2}-1)((2^\tau-1) * \Theta + \Omega 2^{-\tau})$.

Vergleicht man dieses Ergebnis mit dem in [Wilkinson 69] S. 21, Formel (25.5), (25.8) und (25.9), so stellt man Übereinstimmung fest. Dabei ist natürlich bei Wilkinson $\Theta = \Omega = o$, d.h. $\vartheta_n = o$.

Beispiel 2.21: Abschätzung des Rundungsfehlers für das abgeänderte Iterationsverfahren $y_{i+1} := a \boxed{*} y_i \boxed{+} b$.

Bei diesem Beispiel läßt sich sehr schön zeigen, wie man mit prämetrischen Feldern die Vorwärts- und Rückwärtsanalyse durchführt.

Das Iterationsverfahren $x_{i+1} = ax_i+b$ für $x_o \in \mathbb{R}$ und $i \geq o$ konvergiert mit $|a| < 1$ gegen den Fixpunkt $\hat{x} = \frac{b}{1-a}$ der Gleichung $x = ax + b$. Die i-te Näherung berechnet sich zu

$x_i = a^i x_o + b\frac{a^i-1}{a-1}$ und somit ist

$$p(\hat{x},x_i) = \hat{x} - x_i = \frac{b}{1-a} - (a^i x_o + b\frac{1 - a^i}{1 - a})$$

$$= \frac{ba^i}{1-a} - a^i x_o$$

$$= a^i(\frac{b}{1-a} - x_o)$$

$$= \frac{a^i}{1-a}(b + ax_o - x_o)$$

$$= \frac{a^i}{1-a}(x_1 - x_o)$$

$$= \frac{a^i}{1-a}\, p(x_1,x_o) \ .$$

Es gilt also

$$\text{(I)} \quad p(\hat{x},x_i) = \frac{a^i}{1-a}\, p(x_1,x_o) \ .$$

Betrachtet man das im Raster $(\mathcal{R}_b^{\tau}\mathbb{R}^*, \boxplus, \boxdot)$ abgeänderte Iterationsverfahren

$$y_{i+1} := a \boxdot y_i \boxplus b \quad \text{mit} \quad y_o \in \mathcal{R}_b^{\tau}\mathbb{R}^* \quad \text{und} \quad i \geq o,$$

so lassen sich mit Hilfe des prämetrischen Feldes ϑ^* folgende Fehlerabschätzungen angeben:

<u>Vorwärtsanalyse:</u>

$$p(x_i,y_i) = p(a \cdot x_{i-1} + b,\ a \boxdot y_{i-1} \boxplus b) \leq$$

$$\overset{\text{(PF3)}}{\subseteq} D_1(a \cdot x_{i-1},\ a \cdot y_{i-1}) + (D_1(b,b) - \overline{D_1(o,o)}) \leq$$

$$\overset{\text{(PF4)}}{\subseteq} a * (D_2(x_{i-1},y_{i-1}) - \overline{D_2(o,o)}) + D_2(o,o) + (D_1(b,b) - \overline{D_1(o,o)}$$

Schätzt man nun $D_2(x_{i-1},y_{i-1})$ wieder auf dieselbe Weise ab usw., so erhält man schließlich:

$$\text{(II)}\quad p(x_i,y_i) \subseteq a^i * (D_{2i}(x_o,y_o) - \overline{D_{2i}(o,o)}) +$$

$$+ \sum_{\nu=o}^{i-1} a^\nu * (D_{2\nu+1}(b,b) - \overline{D_{2\nu+1}(o,o)})$$

$$+ \sum_{\nu=o}^{i-1} a^\nu \, (D_{2(\nu+1)}(o,o) - \overline{D_{2\nu}(o,o)})$$

<u>Lemma:</u>

Mit $D_k(o,o) := (\varepsilon^k - 1) * 2^\tau * B$ und $\varepsilon := 1 + 2^{-\tau}$
und $a \in [-1,1]$

ist $\sum_{\nu=o}^{i-1} a^\nu (D_{2(\nu+1)}(o,o) - \overline{D_{2\nu}(o,o)}) \subseteq D_{2i}(o,o)$

<u>Bew.:</u>

Es ist $D_{2(\nu+1)}(o,o) - \overline{D_{2\nu}(o,o)} = (\varepsilon^{2\nu+2} - 1)2^\tau * B - (\varepsilon^{2\nu} - 1)2^\tau * \overline{B} =$

$$= (\varepsilon^{2\nu+2} - 1 - (\varepsilon^{2\nu} - 1))2^\tau * B =$$

$$= (\varepsilon^{2\nu+2} - \varepsilon^{2\nu})2^\tau * B =$$

$$= \varepsilon^{2\nu}(\varepsilon^2 - 1)2^\tau * B$$

und $a^\nu * B \subseteq [-1,1] * B = B$, so daß schließlich gilt:

$$\sum_{\nu=o}^{i-1} a^{\nu} * B * (\varepsilon^{2\nu}(\varepsilon^2-1)2^{\tau}) \leq \sum_{\nu=o}^{i-1} \varepsilon^{2\nu}(\varepsilon^2-1)2^{\tau} * B = \frac{\varepsilon^{2i}-1}{\varepsilon^2-1}(\varepsilon^2-1)2^{\tau} * B =$$

$$= (\varepsilon^{2i} - 1)2^{\tau} * B =$$

$$= D_{2i}(o,o) \quad .$$

Nimmt man der Einfachheit halber an, daß $x_o = y_o \in \mathcal{R}\,{}_b^{\tau}\mathbb{R}^*$, so folgt mit Hilfe des Lemmas die Abschätzung

$$p(x_i,y_i) \leq b\frac{1-a^i}{1-a} - \sum_{\nu=o}^{i-1} a^{\nu} * \mathfrak{A}^{(2\nu+1)} * b + D_{2i}(o,o)$$

Für $i \to \infty$ gehe $x_i \to \hat{x}$ und $y_i \to y^*$ und ist nach m Iterationsschritten gesichert, daß kein Unterlauf mehr stattfindet, so gilt die Fehlerabschätzung:

$$p(\hat{x},y^*) \leq \hat{x} + D_{2m}(o,o) - \sum_{\nu=o}^{\infty} a^{\nu} * \mathfrak{A}^{(2\nu+1)} * b < \infty$$

Dies gilt offenbar nur, falls $\|a\,\mathfrak{A}^{(2)}\| < 1$, d. h. $|a| < 1/\|\mathfrak{A}^{(2)}\|$ gilt.

Durch Subtraktion von $\hat{x}$ auf beiden Seiten folgt

$$\text{(III)} \quad y^* \leq D_{2m}(o,o) - \sum_{\nu=o}^{\infty} a^{\nu} * \mathfrak{A}^{(2\nu+1)} * b \ .$$

Der Übergang zur Fehlerabschätzung im Festpunktraster $F\mathcal{R}_b^\tau\mathbb{R}^*$ ist in Formel (II) mit $\mathfrak{A} = 1$ nach Lemma 2.18 nicht schwierig. Es ist mit $x_o = y_o \in F\mathcal{R}_b^\tau\mathbb{R}^*$ und $D_k(x,x) = k * B$

$$p(x_i,y_i) \subseteq \sum_{\nu=o}^{i-1} a^\nu(2\nu + 2 - 2\nu) * B$$

$$= \sum_{\nu=o}^{i-1} 2a^\nu * B$$

$$= \frac{2B}{1 - a} .$$

Ein Vergleich mit [Henrici 64] S. 393 zeigt eine gute Übereinstimmung, wenn man bedenkt, daß nach (PF3) auch bei der Addition ein Rundungsfehler berücksichtigt wird, der bei Festkommarechnung aber nicht vorliegt. Dies läßt sich speziell leicht berücksichtigen durch Beachtung der Tatsache, daß in Festkommarastern gilt:

(V) $a \boxplus b = a + b$ und $a \boxdot b = \square(a \cdot b)$

Nach Lemma 2.18 gilt dann speziell:

(PF3') $D_k(a+k, a \boxplus y) = p(a+x, a \boxplus y) + k * B$

$$= p(a+x,a+y) + k * B$$

$$= D_k(a,a) + D_k(x,y) - \overline{D_k(o,o)}$$

Rückwärtsanalyse:

Zur Rückwärtsanalyse werden ein $\tilde{a}$ und $\tilde{b}$ gesucht, so daß nach Abschnitt 2.2 gilt:

$$p(\tilde{a}y + \tilde{b}, a \boxed{\cdot} y \boxed{+} b) \overset{(PF3^*)}{\leq} D_1(\tilde{a}y, a \cdot y) + (D_1(\tilde{b}, b) - D_1(o,o) \leq$$

$$\overset{(PF4)}{\leq} p(o,y) * (D_2(\tilde{a},a) - \overline{D_2(o,o)}) + D_2(o,o) +$$

$$+ D_1(\tilde{b},\tilde{b}) - \overline{D_1(o,o)}$$

Nach Formel (IV) aus Abschnitt 2.2 folgt schließlich:

$\tilde{a} \leq \mathfrak{A}^{(2)} * a + \mathfrak{A} * B + B$ und

$\tilde{b} \leq \mathfrak{A} * b + B$

Für den Fixpunkt des so abgeänderten Iterationsverfahrens

$z_{i+1} = \tilde{a} z_i + \tilde{b}$ erhält man mit

$$|\tilde{a}| \leq \| \mathfrak{A}^{(2)} * a + \mathfrak{A} * B + B \| < 1$$

d. h. für $|a| < \dfrac{1 - \| \mathfrak{A} * B + B \|}{\| \mathfrak{A}^{(2)} \|}$, die Abschätzung:

$$\hat{z} = \frac{\tilde{b}}{1 - \tilde{a}} \leq \sum_{\nu=o}^{\infty} (\mathfrak{A}^{(2)} * a + \mathfrak{A} * B + B)^{(\nu)} * (\mathfrak{A} * b + B).$$

Diese Abschätzung hat im Vergleich zur Abschätzung (III) den

Vorteil, daß für genügend kleines B mit

$$|a| < \frac{1 - \| \mathcal{A} * B + B \|}{\| \mathcal{A}^{(2)} \|}$$

beliebig oft Unterlauf stattfinden kann.

Für $B = o$ sind beide Abschätzungen gleichwertig!

2.4.2 Die prämetrischen Felder über den Rastern des $V_n\mathbb{R}$.

Es sei $(V_n\mathbb{R}, +, *, \cdot)$ der n-dimensionale euklidische Vektorraum über $(\mathbb{R}, +, \cdot)$ mit Vektoren $x = (x_i) \in V_n\mathbb{R}$ und mit den Verknüpfungen:

$$\bigwedge_{x,y \in V_n\mathbb{R}} \quad x + y := (x_i + y_i)$$

$$\bigwedge_{A \in M_n\mathbb{R}} \; \bigwedge_{x \in V_n\mathbb{R}} \quad A * x := \left(\sum_{j=1}^{n} (A_{ij} \cdot x_j) \right)$$

$$\bigwedge_{\lambda \in \mathbb{R}} \; \bigwedge_{x \in V_n\mathbb{R}} \quad \lambda \cdot x = (\lambda \cdot x_i)$$

Entsprechend sei in Anlehnung an [Ullrich 72] das Vektoid $(V_n\mathcal{R}\mathbb{R}^*, \boxplus, \boxast, \boxdot)$ über dem Rasterringoid $(\mathcal{R}\mathbb{R}^*, \boxplus, \boxdot)$ gegeben mit

$x,y \in V_n \mathcal{R}\mathbb{R}^*$ $\qquad x \boxed{+} y := (x_i \boxed{+} y_i)$

$A \in M_n \mathcal{R}\mathbb{R}^*$ $\qquad x \in V_n \mathcal{R}\mathbb{R}^*$ $\qquad A \boxed{*} x := (\boxed{\sum_{j=1}^{n}} (A_{ij} \boxed{\cdot} x_j))$

$\lambda \in \mathcal{R}\mathbb{R}^*$ $\qquad x \in V_n \mathcal{R}\mathbb{R}^*$ $\qquad \lambda \boxed{\cdot} x := (\lambda \boxed{\cdot} x_i)$.

Ferner sei die Prämetrik $p(x,y) := (p(x_i,y_i)) = x - y$ in $V_n\mathbb{R}$ und $p(A,B) = (p(A_{ij},B_{ij})) = A - B$ in $M_n\mathbb{R}$ gegeben. Es gilt dann für $A \in M_n\mathbb{R}$, $a,x,y \in V_n\mathbb{R}$ und $\lambda \in \mathbb{R}$:

$$p(A * x, A * y) = p(A,o) \ast p(x,y)$$

$$p(a + x, a + y) = p(x,y)$$

$$p(\lambda \cdot x, \lambda \cdot y) = p(\lambda,o)p(x,y)) \ .$$

Mit diesen Vereinbarungen läßt sich dann ohne Berücksichtigung eines Unterlaufes zeigen, daß gilt:

$$p(o,x \boxed{+} y) \subset \mathfrak{A} * p(o,x+y) = \phi p(o,x+y) \quad \text{mit } \phi \in \mathcal{L}_1$$

$$p(o,A \boxed{*} x) \subseteq p(A,o) * \mathcal{L} * p(o,x) = p(A,o)\Psi p(o,x) \text{ mit } \Psi \in \mathcal{L}_n$$

und

$$p(o,\lambda \boxed{\cdot} x) \subseteq \mathfrak{A} * p(o,\lambda \cdot x) = \Lambda p(o,\lambda \cdot x) \quad \text{mit } \Lambda \in \mathcal{L}_1 .$$

Damit ist

$$\vartheta^{*}_{n} := \{D_{ijk}(x,y) := p(x,o) + \mathfrak{A}^{(i)} * \mathcal{L}^{(j)} * \mathcal{L}^{(k)} * p(o,y) \quad \text{mit} \quad i,j,k \in \mathbb{N}_0\}$$

ein (dreischichtiges) prämetrisches Feld, da mit

$$\mathfrak{A}^{(i)} * \mathcal{L}^{(j)} * \mathcal{L}^{(k)} * X = \phi^{(i)}\psi^{(j)}\Lambda^{(k)}(X) = \mathfrak{z} \in \mathcal{L}_n \quad \text{ist.}$$

Es gelten in diesem quasilinearen prämetrischen Feld die speziellen Eigenschaften:

$$D_{i,j,k}(a + x, a \boxplus y) \subseteq D_{i+1,j,k}(a,a) + D_{i+1,j,k}(x,y)$$

$$D_{i,j,k}(A * x, A \boxed{*} y) \subseteq p(A,o) * D_{i,j+1,k}(x,y)$$

$$D_{i,j,k}(\lambda \cdot x, \lambda \boxdot y) \subseteq p(\lambda,o) * D_{i,j,k+1}(x,y) \ .$$

In vielen Fällen genügt es $\mathfrak{A}, \mathcal{L}, \mathcal{L}$ durch eine universelle Konstante $\tilde{\mathfrak{A}}$ abzuschätzen, dann kann das obige dreischichtige prämetrische Feld zu einem in den vorhergehenden Abschnitten behandelten Typ vereinfacht werden.(Dabei kann ohne Schwierigkeit der Unterlauf miteinbezogen werden.)

Möchte man schließlich eine Fehlerabschätzung nicht über $V_n\mathbb{IR}$ sondern über $\mathbb{IR}$ erreichen, so ist eine geeignete neue eindimensionale Prämetrik $\overline{p}$ einzuführen, z.B.:

$$\bigwedge_{x,y \in V_n \mathbb{R}} \overline{p}(x,y) := \bigcup_{i=1}^{n} p(x_i,y_i)$$

Es gilt dann für alle $i = 1, \ldots, n$ $p(x_i,y_i) \subseteq \overline{p}(x,y)$.

Diese Prämetrik $\overline{p}$ induziert ein entsprechendes eindimensinales prämetrisches Feld $\overline{\vartheta}$.

Dieser Abschnitt zeigt, daß das Konzept des prämetrischen Feldes auch auf Räume mit mehrstelligen inneren und äußeren Verknüpfungen erweitert werden kann. Liegen r verschiedene algebraische Verknüpfungen vor, so ist i.a. das prämetrische Feld r-schichtig.

3. Das metrische Feld

In Analogie zu Abschnitt 1.2 Satz 1.5, in dem gezeigt wird, daß aus einer geeigneten Prämetrik p und Norm n eine Metrik $q := n \circ p$ erzeugt werden kann, kann es auch für Fehlerabschätzungen von Interesse sein, keine Abschätzung in $V_n \mathbb{I}\mathbb{R}$ mit $\subseteq$ sondern eine Abschätzung in $\mathbb{R}^+$ bezüglich $\leq$ zu besitzen.

Auf Grund des Verlustes von Vorzeichenkonstellationen zeigt es sich, daß das metrische Feld nur für Vorwärtsanalysen geeignet ist.

Prinzipiell kann im Anscnluß an eine Abschätzung in einem prämetrischen Feld mit der in Kapitel III eingeführten Norm $\| \cdot \|$ über $V_n \mathbb{I}\mathbb{R}$ stets eine reelle Abschätzung gewonnen werden.

Ist $p(x,y) \subseteq B$, so ist nach Satz III,2.42 und Satz 1.5

$q(x,y) := \| p(x,y) \| \leq \|B\| \in \mathbb{R}^+$ und q eine Metrik.

Für den Fall jedoch, daß von vornherein in $(\mathbb{R}^+, \leq)$ abgeschätzt werden soll, wird man im wesentlichen analoge Eigenschaften wie in Definition 2,3 benötigen.

Definition 3.1:

Die Menge

$$L := \{ l \mid \bigwedge_{x \in \mathbb{R}^+} l(x) := ax+b \quad \text{mit} \quad a,b \in \mathbb{R}^+ \}$$

heißt die Menge der positiv definiten isotonen linearen Abbildungen von $\mathbb{R}^+$ nach $\mathbb{R}^+$.

Definition 3.2: (des metrischen Feldes)

Sei M eine beliebige Menge und $\mathbb{R}^*$ das Ringoid der reellen Zahlen. Eine Menge U von Funktionalen $u : M \times M \to \mathbb{R}^{*+}$ heißt ein metrisches Feld von M bzw. (M,U,ρ) ist ein Raum mit metrischem Feld, wenn gilt:

(MF1) $\rho \in U$ und (M,ρ) ist metrischer Raum

(MF2) $$\bigwedge_{u \in U} \; \bigwedge_{x,y,z \in M} u(x,z) \leq \rho(x,y) + u(y,z)$$

Ist $(M,+,\cdot)$ Ringoid, $(\mathcal{R}M, \boxplus, \boxdot)$ Rasterringoid von M mit den ausgezeichneten Elementen $\{-e,o,e\}$, (M,ρ) metrisches Ringoid nach Definition 1.1, so heißt U verträglich <u>mit dem Ringoid-Rasterringoid</u> $M \times \mathcal{R}M$, wenn gilt:

$$\text{(MF3)} \quad \bigwedge_{u \in U} \; \bigvee_{u' \in U} \; \bigwedge_{\substack{x \in M \\ a,y \in \mathcal{R}M}} u(a+x, a \boxplus y) \leq u'(a,a) + u'(x,y) - u'(o,o)$$

$$\text{(MF4)} \quad \bigwedge_{u \in U} \; \bigvee_{u' \in U} \; \bigwedge_{\substack{x \in M \\ a,y \in \mathcal{R}M}} u(a\cdot x, a \boxdot y) \leq \rho(a,o)(u'(x,y) - u'(o,o)) + u'(o,o)$$

Zur Vereinfachung kann folgende zusätzliche Forderung gestellt werden:

(MF5) U heißt linear, wenn gilt:

$$\bigwedge_{u \in U} \; \bigvee_{l \in L} \; \bigwedge_{x,y \in M} u(x,y) = \rho(x,y) + l(\rho(o,y))$$

Ist ein prämetrisches Feld ϑ gegeben, so lassen sich auf verschiedene Weise daraus metrische Felder $U(\vartheta)$ gewinnen.

Die einfachste Methode wäre folgende:

$U_1(\vartheta) := \{u \mid u = \|D\| \wedge D \in \vartheta\}$

$U_1(\vartheta)$ ist ein metrisches Feld, denn es gilt:

(MF1) Mit $\rho := \|p\|$ und $p \in \vartheta$ ist $\rho \in U$, und nach Satz

1.5 ist dann (M,ρ) ein metrischer Raum.

$$\text{(MF2)} \quad u(x,z) = \|D(x,z)\| \overset{\text{(PF2)}}{=} \|p(x,y) + D(y,z)\| \leq$$

$$\leq \|p(x,y)\| + \|D(y,z)\| =$$

$$= \rho(x,y) + u(y,z) \; .$$

Eine Verträglichkeit mit einem Ringoid-Rasterringoid $M \times \mathcal{R}\, M$ ist in dieser Allgemeinheit i.a. nicht möglich.

Lemma 3.3:

Für alle $A,B \in I\!R$ mit $-A = A$, $-B = B$ und $A \subseteq B$ gilt bezüglich der in III,4. eingeführten Norm:

$$\|B - \overline{A}\| = \|B\| - \|A\|$$

Bew.:

Es ist $\|B - \overline{A}\| = |\lambda B - \lambda A| \sqcup |\rho B - \rho A|$ und aus $A \subseteq B$ folgt mit $\lambda B \leq \lambda A \leq 0 \;\wedge\; \rho A \leq \rho B \leq 0$:

$$|\lambda B - \lambda A| \sqcup |\rho B - \rho A| = (|\lambda B| - |\lambda A|) \sqcup (|\rho B| - |\rho A|)$$

Aus $-A = A$ folgt $|\lambda A| = |\rho A| = \|A\|$ und

aus $-B = B$ $\quad |\lambda B| = |\rho B| = \|B\|$, so daß

schließlich

$$\|B - \overline{A}\| = (\|B\| - \|A\|) \sqcup (\|B\| - \|A\|) = \|B\| - \|A\|$$

die Behauptung folgt.

Satz 3.4:

Ist das prämetrische Feld ϑ verträglich mit einem Ringoid-Rasterringoid $M \times_{\mathcal{R}} M$ und für alle $D \in \vartheta$ ist $D(o,o) = -D(o,o)$[1], so ist

$$U_2(\vartheta) := \{u \mid u(x,y) := \|D(x,y) - \overline{D(o,o)}\| + \|D(o,o)\| \wedge D \in \vartheta\}$$

ein mit diesem Ringoid-Rasterringoid $M \times_{\mathcal{R}} M$ verträgliches metrisches Feld.

Bew.:

Nach Definition ist $u(o,o) = \|D(o,o)\|$.

(MF1) Es ist $\rho = \|p\|$ und (M,ρ) metrischer Raum.

(MF2) $u(x,z) = \|D(x,z) - \overline{D(o,o)}\| + \|(Do,o)\| =$

$$\overset{(PF2)}{=} \|p(x,y) + D(y,z) - \overline{D(o,o)}\| + \|D(o,o)\| \leq$$

$$\leq \|p(x,y)\| + \|D(y,z) - \overline{D(o,o)}\| + \|D(o,o)\| =$$

$$= \rho(x,y) + u(y,z)$$

(MF3) $u(a+x, a \boxplus y) = \| D(a+x, a \boxplus y) - \overline{D(o,o)}\| + \|D(o,o)\|$

Nach (PF3) aber ist

$$D(a+x, a \boxplus y) - \overline{D(o,o)} \subseteq (D'(a,a) - \overline{D'(o,o)}) + (D'(x,y) - \overline{D'(o,o)}) + $$
$$+ (D'(o,o) - \overline{D(o,o)})$$

---

1) Diese Bedingung ist, wie Abschnitt 2.4 zeigt, i.a. erfüllt.

daraus folgt:

$$\| D(a+x, a \boxed{+} y) - \overline{D(o,o)} \| \leq$$

$$\| D'(a,a) - \overline{D'(o,o)} \| + \| D'(x,y) - \overline{D'(o,o)} \|$$

$$+ \| D'(o,o) - \overline{D(o,o)} \|$$

Nach Voraussetzung ist $-D'(o,o) = D'(o,o)$ und $-D(o,o) = D(o,o)$ und nach (PF3) gilt speziell mit $a = x = y = o$ $D(o,o) \subseteq D'(o,o)$. Somit folgt nach Lemma 3.3

$$\| D'(o,o) - \overline{D(o,o)} \| = \| D'(o,o) \| - \| D(o,o) \|$$

Durch Addition von $u(o,o) = \| D(o,o) \|$ auf beiden Seiten folgt somit:

$$u(a+x, a \boxed{+} y) \leq u'(a,a) + u'(x,y) - u'(o,o)$$

(MF4) Aus (PF4) folgt

$$D(a \cdot x, a \boxed{\cdot} y) - \overline{D(o,o)} \subseteq$$

$$p(a,o) * (D'(x,y) - \overline{D'(o,o)}) + (D'(o,o) - \\ - \overline{D(o,o)})$$

und damit

$$\| D(a \cdot x, a \boxed{\cdot} y) - \overline{D(o,o)} \| \leq$$

$$\| p(a,o) \| \quad \| D'(x,y) - \overline{D'(o,o)} \| + \| D'(o,o) - \overline{D(o,o)} \|$$

und mit Hilfe von Lemma 3.3:

$$u(a \cdot x, a \boxed{\cdot} y) \leq \rho(a,o)(u'(x,y) - u'(o,o)) + u'(o,o) .$$

(MF5) ist i.a. nicht nachweisbar.

Eine direkte Konstruktionsmethode bietet sich in Analogie zu Satz 2.11 an:

<u>Satz 3.5:</u>

Ist $(R, +, \cdot, \rho, \leq)$ metrisches Ringoid,

$(\mathcal{R}R, \boxplus, \boxdot, \leq)$ Rasterringoid von R mit

$$\text{(V)} \quad \bigwedge_{\circ \in \{+,\cdot\}} \bigwedge_{a,b \in \mathcal{R}R} a \boxed{\circ} b = \Box(a \circ b)$$

und den ausgezeichneten Elementen $\{-e, o, e\}$ und ist die Rundung

$\Box : R \to R$ linear isoton beschränkt mit

$$\text{(R6')} \quad \bigwedge_{a \in R} \rho(y, \Box y) \leq \alpha\rho(o,y) + \beta = l(\rho(o,y)) \quad \text{mit} \quad l \in L$$

so ist

$$U^* := \{u_i \mid u_i(x,y) := \rho(x,y) + ((\alpha+1)^i - 1)(\rho(o,y) + \tfrac{\beta}{\alpha}) \quad \wedge \quad i \in \mathbb{N}_0\}$$

ein lineares metrisches Feld, das mit dem Ringoid-Rasterringoid $R \times \mathcal{R}R$ verträglich ist.

Es gilt speziell: $\bigwedge_{i \in \mathbb{N}_0} u_i(x, \sqcup y) \leq u_{i+1}(x,y)$

Bew.:

(MF1) Mit $i = o$ ist $\rho \in U^*$ und (R,ρ) ist nach Voraussetzung metrischer Raum.

(MF2) $$u_i(x,z) = \rho(x,z) + 1^{(i)}(\rho(o,z)) \leq$$

$$\overset{(M3)}{\leq} \rho(x,y) + \rho(y,z) + 1^{(i)}(\rho(o,z)) =$$

$$= \rho(x,y) + u_i(y,z)$$

Lemma:

Es gilt $u_i(x,\Box y) \leq u_{i+1}(x,y)$

Bew.:

Es ist $u_i(x,\Box y) = \rho(x,\Box y) + (\alpha+1)^i - 1)(\rho(o,\Box y) + \frac{\beta}{\alpha})$

Nun ist nach (R6')

$$\rho(x,\Box y) \leq \rho(x,y) + \rho(y,\Box y) \leq \rho(x,y) + \alpha\rho(o,y) + \beta$$

und

$$\rho(o,\Box y) \leq \rho(o,y) + \rho(y,\Box y) \leq (\alpha+1)\rho(o,y) + \beta .$$

Somit folgt weiter:

$$u_i(x, \Box y) \leq \rho(x,y) + \alpha\rho(o,y) + \beta +$$

$$((\alpha+1)^i-1)\{((\alpha+1)\rho(o,y)+\beta) + \frac{\beta}{\alpha}\} =$$

$$= \rho(x,y) + \alpha\rho(o,y) + \beta +$$

$$(\alpha+1)^{i+1}\rho(o,y) - \alpha\rho(o,y) - \rho(o,y) +$$

$$\frac{(\alpha+1)^i-1}{\alpha}(\alpha+1)\beta + \beta$$

$$= \rho(x,y) + ((\alpha+1)^{i+1} - 1)\rho(o,y) +$$

$$\left(\frac{(\alpha+1)^i-1}{\alpha}(\alpha+1) + 1\right)\beta$$

$$= \rho(x,y) + \left((\alpha+1)^{i+1} - 1\right)\rho(o,y) +$$

$$\frac{(\alpha+1)^{i+1} - \alpha - 1 + \alpha}{\alpha}\beta$$

$$= \rho(x,y) + ((\alpha+1)^{i+1} - 1)\left(\rho(o,y) + \frac{\beta}{\alpha}\right)$$

$$= u_{i+1}(x,y) \quad .$$

(MF3) $u_i(a+x, a \boxed{+} y) \overset{(V)}{=} u_i(a+x, \Box(a+y))$

$$\overset{(\text{Lemma})}{\leq} u_{i+1}(a+x, a+y) =$$

$$= \rho(a+x, a+y) + ((\alpha+1)^i-1)\left(\rho(o,a+y)+\frac{\beta}{\alpha}\right)$$

$$= \rho(x,y) + \left((\alpha+1)^i-1\right)\left(\rho(o,a)+\rho(o,y)+\frac{\beta}{\alpha}\right)$$

Da $(\alpha+1)^{i+1} \geq o$ folgt weiter:

$$\leq \rho(x,y) + \big((\alpha+1)^{i+1} - 1\big)\big(\rho(o,a) + \tfrac{\beta}{\alpha}\big)$$

$$+ \big((\alpha+1)^{i+1} - 1\big)\big(\rho(o,y) + \tfrac{\beta}{\alpha}\big)$$

$$- \big((\alpha+1)^{i+1} - 1\big)\tfrac{\beta}{\alpha} =$$

$$= u_{i+1}(a,a) + u_{i+1}(x,y) - u_{i+1}(o,o)$$

(MF4) $u_i(\alpha\cdot x, a \boxdot y) \overset{(V)}{=} u_i(a\cdot x, \square(a\cdot y)$

$$\overset{(\text{Lemma})}{\leq} u_{i+1}(a\cdot x, a\cdot y)$$

$$= \rho(\alpha\cdot x, a\cdot y) + \big((\alpha+1)^{i+1} - 1\big)\big(\rho(o,a\cdot y) + \tfrac{\beta}{\alpha}\big)$$

$$= \rho(a,o)\rho(x,y) + \big((\alpha+1)^{i+1} - 1\big)\big(\rho(o,a)\,\rho(o,y) + \tfrac{\beta}{\alpha}\big)$$

$$= \rho(a,o)\big(\rho(x,y) + \big((\alpha+1)^{i+1} - 1\big)\rho(o,y)\big) +$$

$$\big((\alpha+1)^{i+1} - 1\big)\tfrac{\beta}{\alpha}$$

$$= \rho(\alpha,o)\big(u_{i+1}(x,y) - u_{i+1}(o,o)\big) + u_{i+1}(o,o)$$

(MF5) gilt nach Definition von U^* .

Bemerkung:

Gilt für eine Rundung $\square$ bezüglich einer Prämetrik p :

(R6) $p(o, \square\, y) \leq \alpha\, p(o,y) + B$, so genügt diese Rundung auch

(R6') $\rho(y, \square\, y) \leq \alpha\rho(o,y) + \beta$, denn es ist

$p(y, \Box y) \subseteq (\mathfrak{A} - 1)p(o,y) + B$ und somit

$$\rho(y, \Box y) = \|p(y, \Box y)\| \leq \|(\mathfrak{A} - 1)p(o,y) + B\|$$
$$\leq \|\mathfrak{A} - 1\| \; \|p(o,y)\| + \|B\|$$
$$= \alpha \; \rho(o,y) + \beta$$

mit $\alpha := \|\mathfrak{A} - 1\|$ und $\beta := |B|$.

Auf Grund des Verlustes des Vorzeichens und der Inklusionseigenschaft ist es in metrischen Feldern nicht mehr möglich, Rückwärtsanalysen von Rundungsfehlern vorzunehmen.

Bei vielen Problemen jedoch ist eine Vorwärtsanalyse möglich, so daß in diesen Fällen mit Hilfe von metrischen Feldern eine Fehlerabschätzung formal hergeleitet werden kann.

An demselben einfachen Beispiel wie in Beispiel 2.20 im Abschnitt 2.4.1.3 möge die Arbeitsweise des metrischen Feldes gezeigt werden.

Es gelten dieselben Voraussetzungen wie in Beispiel 2.20 mit $\Theta = \Omega = o$ und es sei das nach Satz 3.5 konstruierte metrische Feld $U^* := \{u_i \mid u_i(x,y) := |x-y| + ((\alpha+1)^i - 1)|y| \wedge i \in N_o\}$ gegeben mit $\alpha := \|\tilde{\mathfrak{A}} - 1\| = 2^{-\tau}$.

Es ist

$$\left| \sum_{\nu=1}^{n} x_\nu - \boxed{\Sigma}_{\mu=1}^{n} x_\mu \right| = u_o\left(\sum_{\nu=1}^{n-1} x_\nu + x_n,\ \boxed{\Sigma}_{\mu=1}^{n-1} x_\mu \boxplus x_n\right) \leq$$

(MF3)

$$\leq u_1\left(\sum_{\nu=1}^{n-2} x_\nu + x_{n-1},\ \boxed{\Sigma}_{\mu=1}^{n-2} x_\mu \boxplus x_{n-1}\right) + u_1(x_n, x_n) \leq$$

(MF3)

$$\leq u_2\left(\sum_{\nu=1}^{n-2} x_\nu,\ \boxed{\Sigma}_{\mu=1}^{n-2} x_\mu\right) + u_2(x_{n-1}, x_{n-1}) + u_1(x_n, x_n) \leq$$

$$\leq \sum_{\nu=2}^{n} u_{n+1-\nu}(x_\nu, x_\nu) + u_{n-1}(x_1, x_1)$$

Mit $\alpha = 2^{-\tau}$ ist $u_i(x,x) = ((1+2^{-\tau})^i - 1)|x| = \eta_i|x|$ wie in Beispiel 2.20 und in [Wilkinson 69].

Sind die Summanden nicht exakt darstellbar und müssen erst auf das Raster $\mathcal{R}_b^\tau\mathbb{R}^*$ gerundet werden, (Rundungsfehler bei Datenübertragung auf Rechenanlagen,) so läßt sich auch dieser Rundungsfehler ohne Mühe miteinbeziehen.

Es gilt:

$$\left| \sum_{\nu=1}^{n} x_\nu - \boxed{\Sigma}_{\mu=1}^{n} (\Box\, x_\mu) \right| =$$

$$= u_o\left(\sum_{\nu=1}^{n-1} x_\nu + x_n,\ \boxed{\Sigma}_{\mu=1}^{n-1} (\Box\, x_\mu) \boxplus (\Box\, x_n)\right) \leq$$

(MF3)

$$\leq u_1\left(\sum_{\nu=1}^{n-1} x_\nu,\ \boxed{\Sigma}_{\mu=1}^{n-1} \Box\, x_\mu\right) + u_1(x_n, \Box\, x_n) \leq$$

(Satz 3.5)

$\leq u_1(\; . \; . \; . \; . \; . \; . \; . \;) + u_2(x_n, x_n) \leq$

(MF3)

$\leq$

.

.

.

$$\leq \sum_{\nu=2}^{n} u_{n+2-\nu}(x_\nu, x_\nu) + u_n(x_1, x_1) \;,$$

d. h. der Index erhöht sich um 1.

Betrachtet man die Größe

$$u_i(x,x) = ((1 + \alpha)^i - 1)|x| \leq |x| e^{i\alpha} =: S_i(x),$$

die offenbar eine durch die i-fach geschachtelte Rundung bewirkte Streuung des Punktes x darstellt, so erkennt man, daß die Summanden der beiden oben durchgeführten Fehlerabschätzungen jeweils mit dem Index exponentiell und im Betrag von x linear wächst.

Am Wachstumsverhalten von Index und Argument der einzelnen Summanden erkennt man sofort, daß der Rundungsfehler $F = \sum_{\nu=1}^{n} S_{n+1-\nu}(x_\nu)$ offenbar dann (wie bekannt) am geringsten ist, wenn von links nach rechts mit betragsmäßig wachsender Folge der Summanden die Summe $\sum_{\nu=1}^{n} x_\nu$ berechnet wird.

III. ERGÄNZUNG UND ERWEITERUNG DER INTERVALLRECHNUNG

Verschiedene Anwendungsbereiche der Intervallrechnung zeigen, daß eine Erweiterung des Raumes der Intervallzahlen von Vorteil sein kann. Folgende Hinweise mögen andeuten, in welcher Weise eine solche Erweiterung benötigt wird.

Es sei das reelle Iterationsverfahren $x=ax+b$ mit $a,b,x \in \mathbb{R} \subseteq \mathbb{IR}$, $|a|<1$ mit Fixpunkt $\hat{x} \in \mathbb{R}$ vorgelegt. Zur numerischen Behandlung dieses Problems sei im Intervallraster $\mathcal{R}\mathbb{IR}^* \cup \{\emptyset\}$ (etwa einer Rechenanlage) eine Einschließung $Y^* \supseteq \hat{x}$ gesucht. Bisher ist es üblich (unter Zugrundelegung des Rasters ($\mathcal{R}\mathbb{IR} \cup \{\emptyset\}$, $\oplus$, $\otimes$, $\sqcap$, $\sqcup$, $\subseteq$), wobei $\sqcap$, $\sqcup$ die Infimums- und Supremumsoperationen bezüglich $\subseteq$ sind), ausgehend von einem $Y_0 \supseteq \hat{x}$ die in Satz I,3.3 a) beschriebene Folge $Y_{i+1} = (a \otimes Y_i \oplus b) \sqcap Y_i$ zu berechnen. Sie konvergiert gegen ein $Y^* \supseteq \hat{x}$. Möchte man nun die Aussage von I,3.3 c ausnutzen, so könnte man mit $Z_0 \subseteq Y^*$ die Iterationsfolge

(*) $Z_{i+1} = ((a \otimes Z_i \oplus b) \sqcup Z_i) \sqcap Y^*$

bilden. Die Analogie, ein $Z_0 \subseteq \hat{x} \in \mathbb{R}$ (entsprechend $w_0 \leq \hat{x}$ nach Satz I,3.3 (c)) zu wählen, scheitert in dem vorliegenden Beispiel daran, daß $\hat{x} \in \mathbb{R}$ ein Atom von $\mathbb{IR}^* \cup \{\emptyset\}$ bezüglich $\subseteq$ ist. Nur für die leere Menge $\emptyset$ gilt:

$$\bigwedge_{\xi \in \mathbb{R}} \emptyset \subseteq \xi .$$

Definiert man notwendigerweise $\emptyset * X = \emptyset + X = \emptyset$ für $X \in \mathbb{IR}^* \cup \{\emptyset\}$, so verletzt der Startwert $x_0 = \emptyset$ i. a. die

Voraussetzung 2. von Satz I,3.3. (Es ist i.a. $f(\emptyset) = \emptyset \neq \hat{x} = f(\hat{x})$ kein erlaubter Startwert). Somit besteht in $\mathcal{R}\,\mathbb{IR}^* \cup \{\emptyset\}$ nicht die Möglichkeit, die Iteration (*) mit einem $Z_o \subseteq \hat{x}$ zu starten.

$\mathbb{IR}$ muß daher so erweitert werden, daß nichttriviale verallgemeinerte Intervalle A mit $A \subseteq \xi \in \mathbb{R}$ existieren, wobei gleichzeitig die Voraussetzungen von Satz I,3.3 (z.B. die Einschließungseigenschaft der arithmetischen Funktionen) erhalten bleiben.

Ein weiterer Anlaß zur Erweiterung der Intervallrechnung geben die Newtonverfahren, die (i.a. bei mehreren Nullstellen im betrachteten Ausgangsintervall) zu einer Division durch Intervalle A mit $0 \subseteq A$ führen, die in $\mathbb{IR}^* \cup \{\emptyset\}$ nicht sinnvoll (d. h. ohne Widerspruch) zu definieren ist.

In Abschnitt 2.4 ist die Fortsetzung einer solchen Division im erweiterten Intervallraum angegeben.
Ein bemerkenswertes "Nebenprodukt" ist die höhere algebraische Abgeschlossenheit, die die Erweiterung der Intervallrechnung zeitigt. Damit ist eine Verallgemeinerung zu der in [Ortolf 69] beschriebenen "Verallgemeinerung der Intervallarithmetik" gegeben.

Um den verschiedenen Bezeichnungen in der Literatur aus dem Wege zu gehen, möge für diese Arbeit festgelegt werden, daß $\mathbb{IR}$ den Intervallraum nach Definition I,2.1 bezeichne (also $\emptyset \notin \mathbb{IR}$). Der Übergang zum vollständigen Verband findet in Abschnitt 2.4 statt. Es zeigt sich, daß im erweiterten Intervallraum die leere Menge nicht benötigt wird.

1. Zusammenstellung der grundlegenden Eigenschaften des Divisionsringoides $(\mathbb{IR},N,+,-,*,/,\subseteq)$ und eine Darstellung dieses Raumes

Bemerkung: Die hier aufgezählten Sätze sind im wesentlichen den Arbeiten von [Kulisch 69] und [Kulisch 72] entnommen, wo auch die entsprechenden Herleitungen und Beweise nachzuschlagen sind.

Nach [Kulisch 72] , ist $(\mathbb{IR},N,+,-,*,/,\subseteq)$ ein durchweg isoton-geordnetes Divisionsringoid. In etwas anderer Zusammenstellung besitzt es folgende algebraische Eigenschaften:

Satz 1.1:

$(\mathbb{IR},N,+,*,/)$ besitzt folgende Eigenschaften:

(a) o) $(\mathbb{IR},+)$ ist abgeschlossen bezüglich +

1) kommutativ: $A+B=B+A$

2) assoziativ: $(A+B)+C=A+(B+C)$

3) Es existiert genau ein Neutralelement o mit

$o+A=A+o=A$

(b) o) $(\mathbb{IR},*)$ ist abgeschlossen bezüglich $*$

1) kommutativ: $A*B=B*A$

2) assoziativ: $(A*B)*C=A*(B*C)$

3) Es existiert genau ein Neutralelement 1 mit

$1*A=A*1=A$

(c) o) $(\mathbb{IR},-)$ ist abgeschlossen bezüglich -

1) darstellbar: $A-B=A+(-1)*B$

(d) o) $(\mathbb{IR},N,/)$ ist abgeschlossen bezüglich / mit Ausnahme der Intervalle $N=\{A \mid o \in A \in \mathbb{IR}\}$ für den Nenner.

1) darstellbar: $A/B=A*(1/B)$

2) $(-1)/B=1/(-B)$

(e) Nullteilereigenschaft:

$$\bigwedge_{A\in\mathbb{IR}} o*A=A*o=o \quad \text{und} \quad A*B=o \iff A=o \ \vee\ B=o$$

(f) Vorzeichenregel: $(-A)*B=-(A*B)$

Eine Eigenschaft, die später im Zusammenhang mit der Berechnung von numerischen Einschließungen von Ergebnissen trotz Rundungserscheinungen auf Rechenanlagen eine wichtige Rolle spielt ist die Einschließungs- oder Inklusions- oder Teilmengeneigenschaft.

<u>Satz 1.2 (Teilmengeneigenschaft)</u>

Für alle Verknüpfungen $o \in \{+,-,*,/\}$ gilt:

$$\bigwedge_{A_1,A_2,B_1,B_2\in\mathbb{IR}} (A_1 \subseteq B_1 \ \wedge\ A_2 \subseteq B_2 \Longrightarrow A_1 o A_2 \subseteq B_1 o B_2)$$

Diese Eigenschaft überträgt sich auf arithmetische Funktionen.

Zum Beweis obiger Aussagen siehe [Kulisch 69].

Eine äquivalente Aussage zu obigem Satz ist

Lemma 1.3:

Gilt für alle Verknüpfungen $o \in \{+,-,*,/\}$

$$\bigwedge_{A,B,C \in I\mathbb{R}} (A \subseteq B \Longrightarrow AoC \subseteq BoC),$$ so gilt die Teilmengen-

eigenschaft nach Satz 1.2.

Bew.:

Sei zusätzlich ein $D \in I\mathbb{R}$ mit $C \subseteq D$ gegeben, so gilt nach Voraussetzung: $A \circ D \subseteq B \circ D$ und $A \circ C \subseteq A \circ D$ und schließlich $A \circ C \subseteq A \circ D \subseteq B \circ D$, d.h. also mit $A \subseteq B \wedge C \subseteq D$ folgt $A \circ C \subseteq B \circ D$.

Satz 1.4 (Stetigkeitseigenschaft)

Für alle Verknüpfungen $o \in \{+,-,*,/\}$ und für beliebiges C gilt:

$$\bigwedge_{\varepsilon > 0} \bigvee_{\delta_\varepsilon > 0} \begin{array}{l} |\lambda(A)-\lambda(B)| < \delta_\varepsilon \Longrightarrow |\lambda(CoA)-\lambda(CoB)| < \varepsilon \\ |\rho(A)-\rho(B)| < \delta_\varepsilon \Longrightarrow |\rho(CoA)-\rho(CoB)| < \varepsilon \end{array}$$

Bei der Division darf natürlich weder A noch $B \in N$ sein.

Bew.:

Nach Definition 2.5 in Kapitel I gilt für $o \in \{+,-,*,/\}$ die Darstellung:

$$AoB = [\lambda Ao\lambda B \sqcap \lambda Ao\rho B \sqcap \rho Ao\lambda B \sqcap \rho Ao\rho B,$$
$$\lambda Ao\lambda B \sqcup \lambda Ao\rho B \sqcup \rho Ao\rho B \sqcup \rho Ao\rho B].$$

Nun sind $\lambda : \mathbb{IR} \to \mathbb{R}$ und $\rho : \mathbb{IR} \to \mathbb{R}$ stetige Abbildungen und $\sqcap$, $\sqcup$, $+$, $-$, $*$ und $/$ stetige Verknüpfungen von $\mathbb{R} \times \mathbb{R} \to \mathbb{R}$, so daß z. B. ohne Beschränkung der Allgemeinheit für die linken Grenzen gilt:

$$\text{Mit } \lambda(C \circ A) = (\lambda C \circ \lambda A) \sqcap (\lambda C \circ \rho A) \sqcap (\rho C \circ \lambda A) \sqcap (\rho C \circ \rho A)$$

$$=: \phi(A)$$

ist ϕ für festgehaltenes $C \in \mathbb{IR}$ eine stetige Funktion $\mathbb{IR} \to \mathbb{R}$, so daß für genügend kleines $\delta_\varepsilon > 0$ mit $|\lambda(A) - \lambda(B)| < \delta_\varepsilon$ stets $|\phi(A) - \phi(B)| < \varepsilon$ gilt.

Veranschaulichung des Intervallraumes $\mathbb{IR}$

Für das spätere Vorhaben, den Intervallraum $\mathbb{IR}$ zu erweitern und gewisse Analogien zur Theorie der komplexen Zahlen aufzudecken, hat sich folgende Darstellung als brauchbar erwiesen,

Durch die Isomorphie $\iota : \mathbb{IR} \longrightarrow \mathbb{R}^2$ sollen die Intervalle in der $\mathbb{R}^2$-Ebene repräsentiert werden. Es ist:

$$\iota : [a,b] \longrightarrow (\mu,\delta) := ((a+b)/2,\ (b-a)/2\ ) \in \mathbb{R}^2$$

Die erste Komponente μ repräsentiert den Mittelpunkt und die zweite Komponente δ den halben Durchmesser des Intervalles. Speziell sind die reellen Zahlen $[a,a] = a \in \mathbb{R}$ [1] dargestellt als $(a,0)$.

---

[1] Siehe Definition I,2.1 .

Die μ-Achse kann daher analog wie in der komplexen Zahlenebene als reelle Achse $\mathbb{R} \subsetneq \mathbb{IR}$ interpretiert werden und vertikal dazu (in mathematisch positiver Orientierung) die Intervallkomponente, die δ-Achse (mit der Einheit $[-1,1] = (o,1)$), die offensichtlich nur für echte Intervalle $\neq o$ ist. Siehe Abb. 1.1.

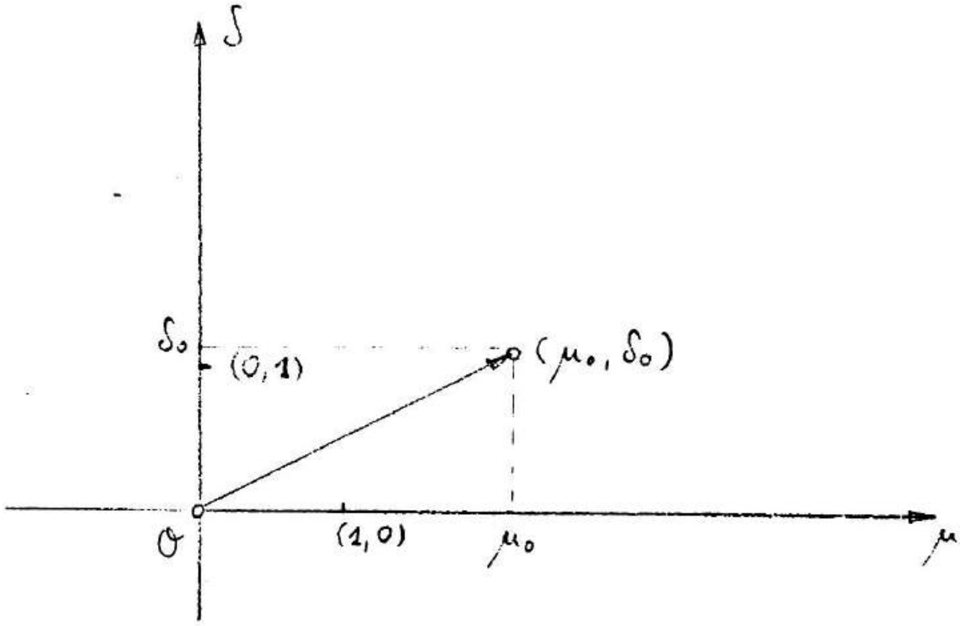


Abb. 1.1

Nach Definition der Intervalle gilt stets, daß für alle $A \in \mathbb{IR}$ $\lambda A \leq \rho A$, d.h. $\rho A - \lambda A \geq o$, also ist der Intervallteil δ nie negativ. Daraus resultiert, daß $\mathbb{IR}$ allein auf der oberen Halbebene einschließlich der reellen Achse repräsentiert wird.

Es wird sich für die Anschauung als praktisch erweisen, jedem Punkt $A = (\mu, \delta) \in \mathbb{IR}$ einen Kegel zuzuordnen mit der Spitze in A und den Mantellinien parallel zu den Winkelhalbierenden $\mu = \pm\, \delta$. Siehe Abb. 1.2.

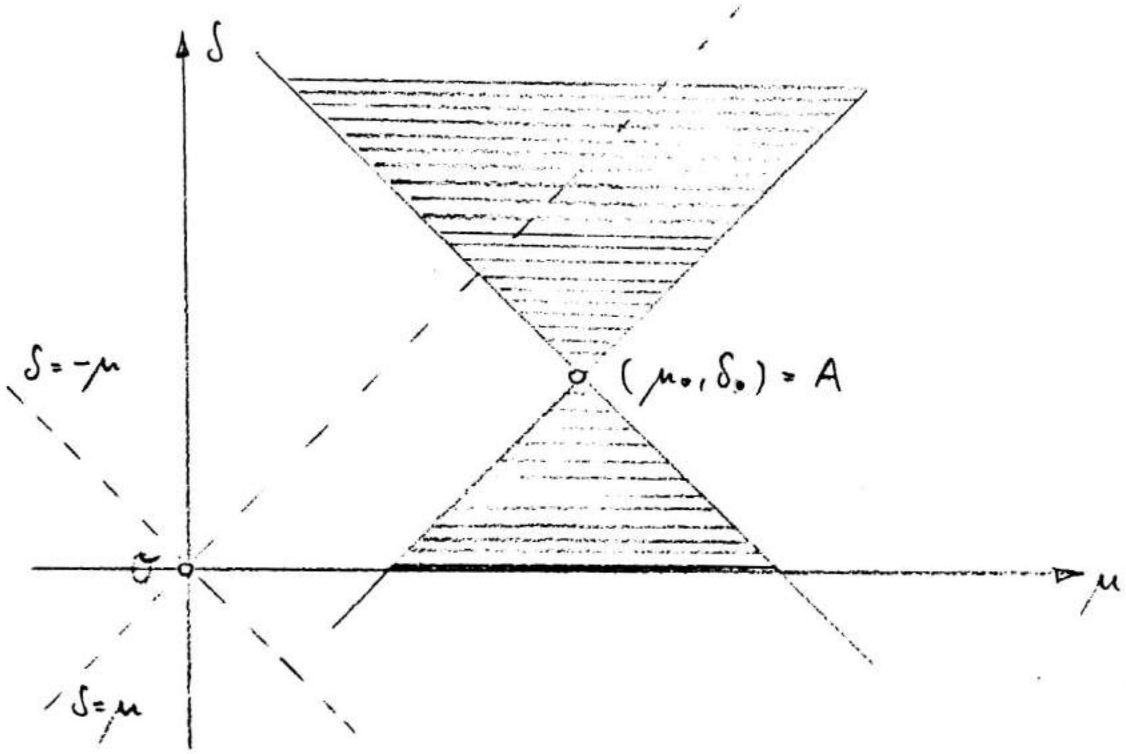


Abb. 1.2

Dieser Kegel charakterisiert einen <u>inneren</u> und einen <u>äußeren</u> Bereich, was im erweiterten Intervallraum jeweils besondere Bedeutungen hat.

Der innere Bereich ist die Menge aller Intervalle X mit $X \subsetneq A$ oder $A \subsetneq X$. Folgerichtig stellt der Schnitt des inneren Kegelbereiches von A mit der μ-Achse gerade die Punktmenge des Intervalles A dar, denn es gilt trivialerweise für alle $\xi \in \mathbb{R}$ mit $\xi \in A$ auch $\xi \subseteq A$ und damit ist ξ Punkt des inneren Kegelbereiches und umgekehrt, ist $\xi \subseteq A$ und $\xi \in \mathbb{R}$ so ist $\xi \in A$.

Um einige Aussagen auch in der μ-δ-Ebene formulieren zu können, folgender Satz:

Satz 1.5:

Sei $A=[a,b]=(\mu_1,\delta_1)$ und $B=[c,d]=(\mu_2,\delta_2)$, so gilt

$$\bigwedge_{A,B \in \mathbb{I}\mathbb{R}} (\ A \subseteq B \iff c \leq a \wedge b \leq d \iff |\mu_2-\mu_1| \leq \delta_2-\delta_1 \ )$$

Bew.: Es gilt die Umkehrabbildung:

$a=\mu_1-\delta_1$ $\qquad$ $b=\mu_1+\delta_1$

$c=\mu_2-\delta_2$ $\qquad$ $d=\mu_2+\delta_2$ $\qquad$ und so folgt

einmal mit $0 \leq a-c = (\mu_1-\mu_2) - (\delta_1-\delta_2)$

$$(*)\ \mu_2-\mu_1 \leq \delta_2-\delta_1$$

und zum anderen mit $0 \leq d-b = (\mu_2-\mu_1) + (\delta_2-\delta_1)$ folgt

$$(**)\ \mu_1-\mu_2 \leq \delta_2-\delta_1 \quad .$$

Aus $(*)$ und $(**)$ folgt schließlich $|\mu_1-\mu_2| \leq \delta_2-\delta_1$.

Umgekehrt, sei $|\mu_2-\mu_1|<\delta_2-\delta_1$ vorgelegt, so folgt mit

$\mu_2-\mu_1 = 1/2\ (c+d-(a+b))$ und $\delta_2-\delta_1 = 1/2\ (d-c-(b-a))$

im einen Fall: $\mu_1-\mu_2 \leq \delta_2 - \delta_1 \Longrightarrow$

$-(c+d)+a+b \leq d-c-b+a \Longrightarrow$

$2b \leq 2d$ und schließlich $b \leq d$;

und im anderen Fall: $\mu_2-\mu_1 \leq \delta_2-\delta_1 \Longrightarrow$

$c+d-(a+b) \leq d-c-b+a \Longrightarrow$

$2c \leq 2a$ und schließlich $c \leq a$.

2. Eine algebraische Erweiterung des Intervallraumes

2.1 Die formale Einführung nichtregulärer Intervalle

Betrachtet man die zu $\mathbb{IR}$ isomorphe Darstellung in $\mathbb{R}^2$ aus Abschnitt 1, so ist, wie bereits erwähnt, der halbe Durchmesser δ der regulären Intervalle stets nicht negativ. Von dieser Sicht bietet sich folgende Erweiterung an: Die formale Einführung der entsprechenden Punkte (μ,δ) aus der unteren Halbebene mit negativem Intervallteil $\delta < o$.

<u>Definition 2.1:</u> Ist $A=[a,b] \in \mathbb{IR}$ mit $a,b \in \mathbb{R}$ und $a\leq b$, so ist die <u>Konjugation</u> $\overline{A} = \overline{[a,b]} := [b,a]$ bzw. $\overline{(\mu,\delta)} := (\mu,-\delta)$ die formale Vertauschung der Intervallgrenzen.

<u>Definition 2.2:</u> Die Menge $\overline{\mathbb{IR}} := \{\overline{A} \mid A \in \mathbb{IR}\}$ ist die Menge der <u>nichtregulären Intervalle</u>.
Die Menge $\mathbb{H} := \overline{\mathbb{IR}} \cup \mathbb{IR}$ ist der <u>erweiterte Intervallraum</u>.

Als nächstes sollen die Konjugation, die Operatoren λ und ρ und die Äquivalenzrelation $\equiv$ auf ganz $\mathbb{H}$ fortgesetzt werden.

<u>Definition 2.3:</u>

a) (Konjugation) Siehe Abb. 2.1

$$\overline{[a,b]} := [b,a]$$

$$[a,b] \in \mathbb{H}$$

b) (Fortsetzung der Operatoren λ und ρ).

$$\bigwedge_{[a,b] \in \mathbb{H}} (\lambda([a,b]) := a \quad \text{und} \quad \rho([a,b]) := b)$$

c) (Fortsetzung der Äquivalenzrelation =)

$$\bigwedge_{A,B \in \mathbb{H}} A = B \iff \lambda A = \lambda B \quad \wedge \quad \rho A = \rho B$$

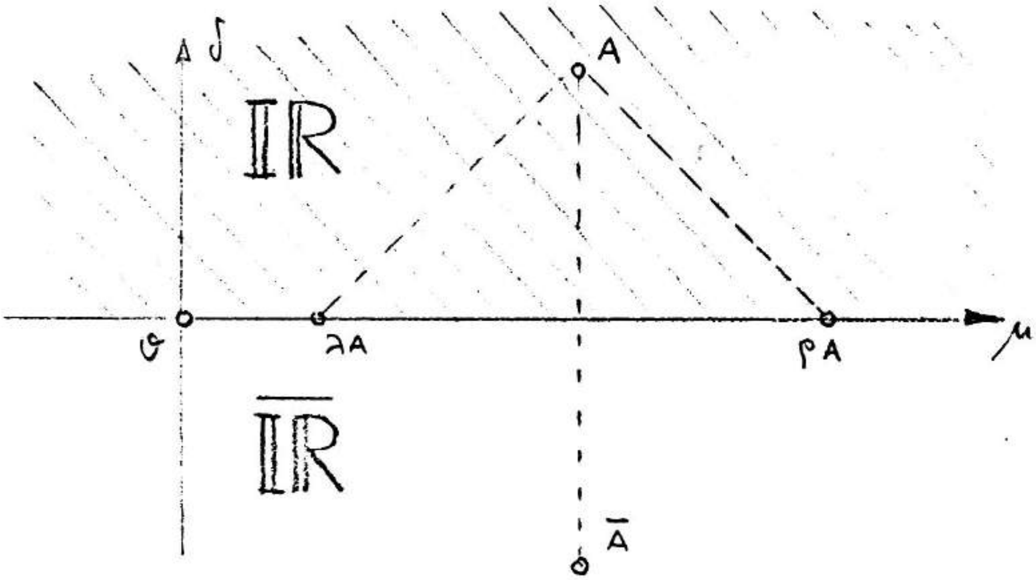


Abb. 2.1

Bemerkung: Aus $\overline{(\mu,\delta)} = (\mu, -\delta)$ folgt unmittelbar, daß die Durchmesser- und Mittelpunktsfunktion aus Definition 2.7 des Abschnitts I formal in ganz $\mathbb{H}$ gelten.

Ferner stellt die Konjugation analog wie in der komplexen Zahlenebene die Spiegelung von $\mathbb{H}$ an der reellen μ-Achse dar. Vergl. Abb. 2.1. Speziell soll künftig in Übereinstimmumg mit Def. I,2.1 vereinbart werden, daß die Punktintervalle $[a,a] = a$ identifiziert werden.

2.2 Fortsetzung der Ordnungsrelation $\subseteq$ auf ganz $\mathbb{H}$

$(\mathbb{I}\mathbb{R}, \subseteq)$ ist halbgeordnet und es wäre zunächst gleichgültig, wie die nichtregulären Intervalle mit $\subseteq$ eingeordnet werden. Im Hinblick jedoch auf die Absicht, inklusionsmonotone Folgen über Punktintervalle hinaus stetig fortzusetzen, sollte die wesentliche Eigenschaft des Satzes 1.6 formal in ganz $\mathbb{H}$ gelten und in $\overline{\mathbb{I}\mathbb{R}} \cap \mathbb{I}\mathbb{R} = \mathbb{R}$ stetig anschließen.

Das wird am einfachsten dadurch erreicht, daß die Definition 2.3 aus Abschnitt I formal auf ganz $\mathbb{H}$ beibehalten wird.

Definition und Satz 2.5:

$(H, \subseteq)$ ist halbgeordnet mit der Halbordnung $\subseteq$[1] gemäß folgender Vorschrift:

Sei $A = (\mu_1, \delta_1)$ und $B = (\mu_2, \delta_2)$.

$$\bigwedge_{A,B \in \mathbb{H}} \quad A \subseteq B \;:\Longleftrightarrow\; |\mu_1 - \mu_2| \le |\delta_2 - \delta_1| \Longleftrightarrow \lambda B \le \lambda A \quad \wedge \quad \rho A \le \rho B$$

Die Halbordnungseigenschaften werden wie folgt erfüllt:

(a) reflexiv $\quad A \subseteq A \quad$ da $\quad \lambda A \le \lambda A \quad \wedge \quad \rho A \le \rho A \quad$ ist,

(b) antisymmetrisch

$$\begin{aligned} A \subseteq B \quad \wedge \quad B \subseteq A &\Longleftrightarrow \lambda B \le \lambda A \wedge \rho A \le \rho B \wedge \lambda A \le \lambda B \wedge \rho B \le \rho A \\ &\Longleftrightarrow \lambda A = \lambda B \quad \wedge \quad \rho A = \rho B \\ &\Longleftrightarrow A = B \end{aligned}$$

---

[1] Eigentlich ist $\subseteq$ in $\overline{\mathbb{I}\mathbb{R}}$ nicht mehr identisch mit der mengentheoretischen Inklusion. Man müßte also genauer $\subseteq_H$ schreiben. Da keine Verwechslungen möglich sind, soll aber das Zeichen $\subseteq$ beibehalten werden.

(c) transitiv $A \subseteq B \quad \wedge \quad B \subseteq C \iff$

$$\lambda C \leq \lambda B \;\wedge\; \lambda B \leq \lambda A \;\wedge\; \rho A \leq \rho B \;\wedge\; \rho B \leq \rho C \Longrightarrow$$

$$\lambda C \leq \lambda A \;\wedge\; \rho A \leq \rho C \Longrightarrow \qquad A \subseteq C$$

Für besondere Zwecke ist es nützlich, folgende schärfere Halbordnung zu besitzen:

$$A \Subset B :\iff A \subseteq B \;\wedge\; \lambda A \neq \lambda B \;\wedge\; \lambda A \neq \rho B \;\wedge\; \rho A \neq \lambda B \;\wedge\; \rho A \neq \rho B$$

Aufgrund der formalen Gleichheit der Ordnungsrelation $\subseteq$ auf $\mathbb{IR}$ und $\mathbb{H}$ bleibt natürlich die Eigenschaft von Satz 1.5 auch auf ganz $\mathbb{H}$ gültig, denn im Beweis dieses Satzes wird nirgends von der Voraussetzung Gebrauch gemacht, daß der Durchmesser nicht negativ in $\mathbb{IR}$ ist.

Satz 2.6:

a) $$\bigwedge_{A,B \in \mathbb{H}} A \subseteq B \iff \overline{A} \supseteq \overline{B}$$

b) $$A \subseteq B \Longrightarrow \bigvee_{x \in \mathbb{R}} (x \subseteq A \subseteq B) \;\vee\; (A \subseteq x \subseteq B) \;\vee\; (A \subseteq B \subseteq x)$$

Bew.: a) $A \subseteq B \iff \lambda B \leq \lambda A \;\wedge\; \rho A \leq \rho B \iff [\rho A, \lambda A] \supseteq [\rho B, \lambda B]$

b) Fall 1: $A \in \mathbb{IR} \Longrightarrow \bigvee_{x \in A} x \subseteq A \subseteq B$

Fall 2: $A \in \overline{\mathbb{IR}} \;\wedge\; B \in \mathbb{IR} \;\wedge\; A \subseteq B \Longrightarrow$

$$\lambda B \leq \lambda A \;\wedge\; \rho A \leq \rho B \;\wedge\; \rho A \leq \lambda A \Longrightarrow$$

$$\bigvee_{\lambda B \leq x \leq \rho B} \rho A \leq x \leq \lambda A \Longrightarrow A \subseteq x \subseteq B$$

Fall 3: $A \in \overline{\mathbb{IR}} \wedge B \in \overline{\mathbb{IR}} \wedge A \subseteq B \Longrightarrow$

$$\overline{A} \supseteq \overline{B} \quad \text{nach Fall 1} \quad \bigvee_{y \in \overline{B}} \quad y \subseteq \overline{B} \subseteq \overline{A}$$

und damit ist $A \subseteq B \subseteq y$.

Folgerung aus diesem Satz ist ein wichtiges Hilfsmittel beim Aufsuchen von monotonen und antitonen Einschließungen.

(a) Ist $A \in \mathbb{IR}$, so gilt $\bigvee_{\xi \in \mathbb{IR}} \xi \subseteq A$ und daraus folgt

$\overline{A} \subseteq \overline{\xi} = \xi \subseteq A$, d.h. mit $\xi \subseteq A$ ist stets $\overline{A} \subseteq \xi$.

(b) Verallgemeinert gilt schließlich: Ist $A \in \mathbb{IR}$ so ist stets $\overline{A} \subseteq A$ und damit gilt für jedes B mit $A \subseteq B$, daß stets $\overline{B} \subseteq \overline{A} \subseteq A \subseteq B$, also $\overline{B} \subseteq A \subseteq B$.

2.3 Fortsetzung der algebraischen Struktur von $\mathbb{IR}$ auf ganz $\mathbb{H}$

Die einzelnen algebraischen Verknüpfungen + - × / werden einzeln auf ganz $\mathbb{H}$ fortgesetzt unter folgenden Konsistenzbedingungen:

(K1) Es sollen die algebraischen Eigenschaften aus Satz 1.1 in ganz $\mathbb{H}$ erhalten bleiben

(K2) Es soll die Einschließungseigenschaft bezüglich $\subseteq$ in $\mathbb{H}$ nach Satz 1.2 erhalten bleiben

(K3) Es soll das Stetigkeitsprinzip nach Satz 1.4 auch in $\mathbb{H}$ gelten.

<u>Bemerkung:</u> In dieser Form und Zusammenstellung gereichen die gestellten Konsistenzbedingungen wahrscheinlich nicht zu einer <u>eindeutigen</u> Fortsetzung. Welche minimalen Konsistenzbedingungen, die dann als Axiome aufzufassen sind, zu einer eindeutigen Fortsetzung genügen, soll hier nicht weiter untersucht werden.

Für diese Arbeit und für die Praxis mag es genügen, irgendeine Fortsetzung solcherart zu finden, daß die obige Konsistenzbedingungen erfüllt sind und damit für numerische Probleme anwendbar ist. Siehe Abschnitt 5 und 6.

Wie sich im weiteren Verlauf dieses Abschnittes ergeben wird, bietet sich diese algebraische Erweiterung auch dadurch an, daß man axiomatisch die Existenz von Inversen bezüglich + * fordert. Diese Methode ist bekannt als "Symmetrisierung eines assoziativen und kommutativen Verknüpfungsgesetzes" [Bourbaki]. Diese Symmetrisierung ist nicht eindeutig, aber alle so erhaltenen Gruppen sind isomorph zueinander. Das Ergebnis der folgenden Abschnitte ist nun, daß die hier (K1)(K2) und (K3) genügende Erweiterung von $\mathbb{R}$ tatsächlich eine Symmetrisierung von $+$ und $*$ darstellt.

Nachdem man dieses a posteriori Ergebnis kennt, hätte man auch eine gesuchte Erweiterung dadurch erhalten können, indem man a priori die Symmetrisierung fordert und unter allen Symmetrisierungen diejenigen aussucht, die nur noch (K2) und (K3) genügen.

2.3.1 Fortsetzung der Addition $+$

Satz 2.7:

Eine mögliche Fortsetzung der Addition in $\mathbb{H}$ ist:

$$\bigwedge_{A,B \in \mathbb{H}} \quad A + B := [\lambda A + \lambda B, \rho A + \rho B]$$

Bew.: (K1) ist erfüllt mit

(a) (o) $(\mathbb{H},+)$ ist abgeschlossen

(1) kommutativ:

$$A+B=[\lambda A+\lambda B,\rho A+\rho B]=[\lambda B+\lambda A,\rho B+\rho A]=B+A$$

(2) assoziativ:

$$\begin{aligned} A+(B+C) &= [\lambda A+(\lambda B+\lambda C),\rho A+(\rho B+\rho C)] \\ &= [(\lambda A+\lambda B)+\lambda C,(\rho A+\rho B)+\rho C] \\ &= (A+B)+C \end{aligned}$$

(3) $o=[o,o] \in \mathbb{H}$ ist Neutralelement der Addition:

$$o+A=[o+\lambda A,o+\rho A]=A+o=A$$

(K2) ist erfüllt mit: Sei $A,B,C \in \mathbb{H}$ und sei $A \subseteq B$, also $(*)$ $\lambda B \leq \lambda A \wedge \rho A \leq \rho B$, so ist für beliebiges $C \in \mathbb{H}$:

$A+C=[\lambda A+\lambda C,\rho A+\rho C]$ und $B+C=[\lambda B+\lambda C,\rho B+\rho C]$.

Nun gilt aber wegen $(*)$:

$\lambda B+\lambda C \leq \lambda A+\lambda C \wedge \rho A+\rho C \leq \rho B+\rho C$, und damit

$A+C=[\lambda A+\lambda C,\rho A+\rho C] \subseteq [\lambda B+\lambda C,\rho B+\rho C] = B+C$

womit die Voraussetzung von Lemma 1.3 erfüllt ist.

(K3) ist erfüllt, da das Ergebnisintervall stetige Funktionen der Intervallgrenzen sind, genauer:

$A+B=[\phi(\lambda A,\lambda B),\psi(\rho A,\rho B)]$ wobei mit

$\phi(x,y)=\psi(x,y)=x+y$ stetige Funktionen über $\mathbb{R} \times \mathbb{R}$ vorliegen.

2.3.2 Fortsetzung der Subtraktion und des Negativoperators -

Satz 2.8:

(1) Die Subtraktion kann mit

$$\bigwedge_{A,B \in \mathbb{H}} A-B:=A+(-B)$$ fortgesetzt werden, wenn dazu

der Negativoperator passend fortgesetzt wird:

(2) $$\bigwedge_{B \in \mathbb{H}} -B := [-\rho B, -\lambda B]$$

Bew.: (1) Da die Subtraktion auf Addition und Negativoperator zurückgeführt ist und die Addition bereits konsistent fortgesetzt ist, bleibt nur noch zu zeigen, daß (2) eine konsistente Fortsetzung ist.

(2) Der Negativoperator ist von $(\mathbb{H},+)$ unabhängig. Jedoch gilt nach Satz 1.1,(f), (Vorzeichenregel), daß $(-1)*A=-A$ ist. (K1) ist also erfüllt, wenn das Produkt konsistent fortgesetzt ist und nach obiger Festsetzung $(-1)*A=-A$ gilt. Entsprechend können die Konsistenzbedingungen (K2) und (K3) auf die konsistente Fortsetzung der Multiplikation zurückgeführt werden.

Bemerkung: In der μ-δ-Ebene stellt der Negativoperator die Spiegelung der $\mathbb{H}$-Ebene an der δ-Achse dar. (Abb. 2.2). Es ist mit $A \in \mathbb{H}$ und

$A=(\mu,\delta)=((\lambda A+\rho A)/2,(\rho A-\lambda A)/2)$

$-A=((-\rho A-\lambda A)/2,(-\lambda A+\rho A)/2)=(-\mu,\delta)$, d.h. es ist $-(\mu,\delta) = (-\mu,\delta)$.

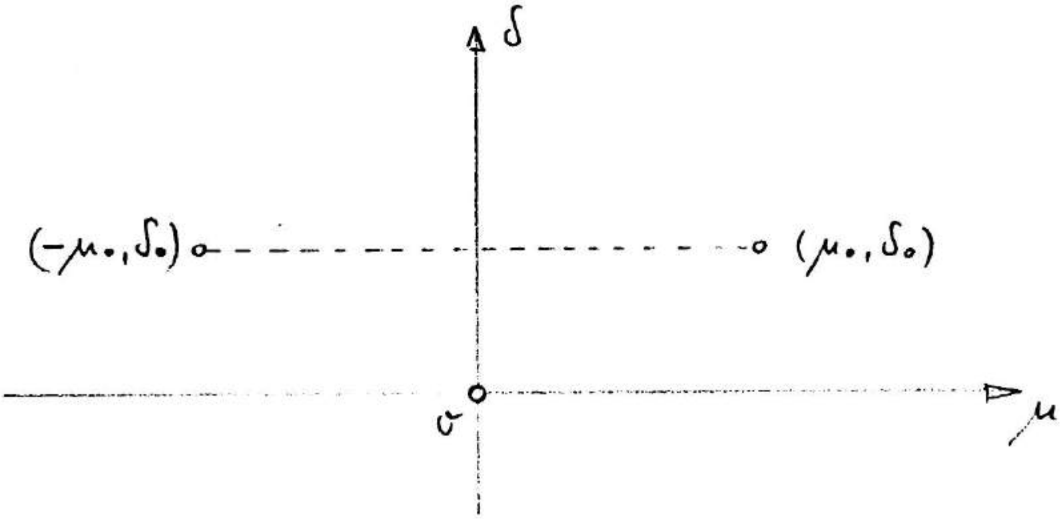


Abb. 2.2

Bemerkenswertes Ergebnis dieser Erweiterung des Intervallraumes ist:

Satz 2.9:

$(\mathbb{H},+)$ ist eine kommutative Gruppe.

Bew.: Nach (K1) ist $(\mathbb{H},+)$ eine **kommutative** Halbgruppe mit Neutralelement $o=[o,o]$. Es bleibt der Nachweis der inversen Elemente. Tatsächlich gilt:

$$\bigwedge_{A\in\mathbb{H}} A+(-\overline{A})=[\lambda A-\lambda A,\rho A-\rho A]=[o,o]=o$$

$-\overline{A}$ ist also das Inverse zu A.

Lemma 2.10:

Es gilt, wie leicht zu verifizieren ist:

$$\overline{A+B} = \overline{A}+\overline{B} \quad \text{und} \quad \overline{A-B} = \overline{A}-\overline{B} \quad .$$

2.3.3 Fortsetzung der Multiplikation $*$ und Einführung des hyperbolischen Ringes

Hierzu erweist sich die Zerlegung von $\mathbb{H}$ in die Sektoren S^+, $\overline{S^+}$, T, $\overline{T}$, S^-, $\overline{S^-}$, L^+, $\overline{L^+}$, L^-, $\overline{L^-}$, S_r und S_l als zweckmäßig. Siehe Abb. 2.3.

Definition 2.11: $S^+ := \{A \in \mathbb{I}\mathbb{R} \mid 0 \leq \lambda A \leq \rho A\}$, $S^- := \{A \in \mathbb{I}\mathbb{R} \mid \lambda A \leq \rho A \leq 0\}$

$T := \{A \mid 0 \in A\}$

$S_r := S^+ \cup \overline{S^+}$ $\qquad$ $S_l := S^- \cup \overline{S^-}$

$L^+ := S^+ \cap T$ $\qquad$ $\overline{L^+} := \overline{S^+} \cap \overline{T}$

$L^- := S^- \cap T$ $\qquad$ $\overline{L^-} := \overline{S^-} \cap \overline{T}$

$L := L^+ \cup \overline{L^+} \cup L^- \cup \overline{L^-}$

Die entsprechenden überstrichenen Größen sind einfach die an der μ-Achse gespiegelten, d. h. konjugierten Größen.

Es gelten speziell folgende Beziehungen:

$\mathbb{H} = S_r \cup S_l \cup T \cup \overline{T}$

$\mathbb{R}^+ = S^+ \cap \overline{S^+}$ $\qquad$ $\mathbb{R}^- = S^- \cap \overline{S^-}$ $\qquad$ $0 = T \cap \overline{T}$

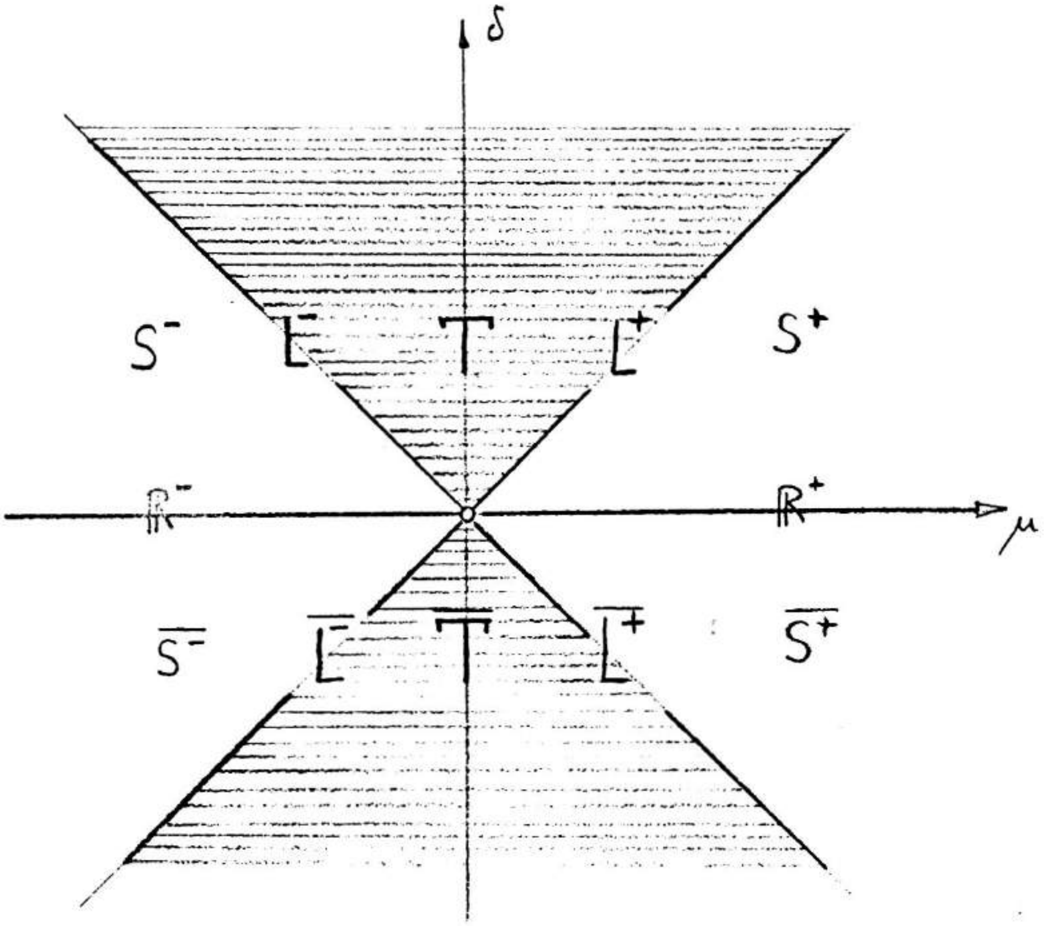


Abb. 2.3

Satz 2.12:

Die Multiplikation $\ast$ ist gemäß der Verknüpfungstabelle Tabelle 2.4 konsistent fortgesetzt.

$\lambda A \lambda B = \lambda A \cdot \lambda B$ ist das Produkt der reellen Zahlen λA und λB usw..

$A\in$ \ $B\in$	S_r	T	S_l	$\overline{T}$
S_r	$[\lambda A\lambda B,\rho A\rho B]$	$[\rho A\lambda B,\rho A\rho B]$	$[\rho A\lambda B,\lambda A\rho B]$	$[\lambda A\lambda B,\lambda A\rho B]$
T	$[\lambda A\rho B,\rho A\rho B]$	$[\min(\lambda A\rho B,\rho A\lambda B),\max(\lambda A\lambda B,\rho A\rho B)]$	$[\rho A\lambda B,\lambda A\lambda B]$	$\circ$
S_l	$[\lambda A\rho B,\rho A\lambda B]$	$[\lambda A\rho B,\lambda A\lambda B]$	$[\rho A\rho B,\lambda A\lambda B]$	$[\rho A\rho B,\rho A\lambda B]$
$\overline{T}$	$[\lambda A\lambda B,\rho A\lambda B]$	$\circ$	$[\rho A\rho B,\lambda A\rho B]$	$[\max(\lambda A\lambda B,\rho A\rho B),\min(\lambda A\rho B,\rho A\lambda B)]$

Tabelle 2.4

Beispiele:

$$[1,2] * [-3,-2]=[2\cdot(-3),2\cdot(-2)]=[-6,-4]$$

$$[-6,+4] * [5,-8]=\circ$$

Zur Beweisführung und auch für andere Zwecke wird es vorteilhaft sein, die Verknüpfung $*$ im hyperbolischen Ring $(\mathbb{H},+,\ )$ darzustellen. (Das Produktzeichen wird im Gegensatz zur Intervallverknüpfung $*$ nicht geschrieben.)

Der hyperbolische Ring ist ein Analogon zum Körper der komplexen Zahlen und trägt eine zum Teil analoge Struktur. Zur Motivation dieser Behauptung und zur Begründung des Attributes hyperbolisch folgende Eigenschaften, die jedoch nicht näher begründet oder untersucht werden sollen:

- Das Produkt zweier komplexer Zahlen stellt eine elliptische Drehbewegung, das hyperbolische Produkt zweier Intervalle eine hyperbolische Drehbewegung dar.

- Faßt man $\mathbb{H}$ als Vektorraum des $\mathbb{R}^2$ auf und definiert folgende Bilinearform: $\phi(A,B):=m(A\overline{B})=m((\mu_1,\delta_1)(\mu_2,\delta_2))=\mu_1\mu_2-\delta_1\delta_2$,

  (m ist Mittelpunktsoperator nach Definition I,2.5,e)),

  so definiert nach [Greub 67] diese Bilinearform ein inneres Produkt vom Index 1 und $\mathbb{H}$ ist dann der 2-dimensionale Pseudoeuklidische Vektorraum. Die durch diese Bilinearform definierte indefinite quadratische Funktion $\psi(A):=\phi(A,A)=m(A\overline{A})=\mu^2-\delta^2$ ergibt nach Greub folgende Aufgliederung:

$A \in S_r \cup S_l$ ist space-like, da $\psi(A) > o$

$A \in T \cup \overline{T}$ ist time-like, da $\psi(A) < o$

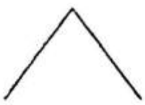
$A \in L^+ \cup \overline{L}^+ \cup L^- \cup \overline{L^-}$ ist light-like, da $\psi(A) = o$

Dieser innere Strukturzusammenhang ist der Anlaß, die einzelnen Sektoren der Abb. 2.3 so zu bezeichnen.

Alle diese Eigenschaften werden jedoch in dieser Arbeit nicht benötigt, außer der Tatsache, daß ($\mathbb{H}$,+,) ein Ring ist und gewisse Verträglichkeitseigenschaften aufweist.

Definition und Satz 2.13:

Der hyperbolische Ring ($\mathbb{H}$,+,) ist der erweiterte Intervallraum $\mathbb{H}$ mit zwei Verknüpfungen, die wie folgt definiert sind:



$$A + B := [\lambda A + \lambda B, \rho A + \rho B]$$

$$A\ B := [\lambda A \lambda B, \rho A \rho B]$$

$A, B \in \mathbb{H}$

Sie genügen den Ringeigenschaften.

Bew.: a) Die hyperbolische Addition ist per definitionem identisch mit der Intervalladdition. Nach Satz 2.9 ist also ($\mathbb{H}$,+) eine Gruppe.

b) (o) ($\mathbb{H}$,) ist abgeschlossen

(1) kommutativ: $AB = [\lambda A \lambda B, \rho A \rho B] = BA$

(2) assoziativ: $(AB)C = [(\lambda A \lambda B)\lambda C, (\rho A \rho B)\rho C]$

$$= [\lambda A(\lambda B \lambda C), \rho A(\rho B \rho C)]$$

$$= A(BC)$$

(3) $1 = [1,1]$ ist Neutralelement mit $A1 = 1A = A$

(4) $\bigwedge_{A \in H \setminus L}$ $\quad 1 /_h A := [1/\lambda A, 1/\rho A]$ ist inverses Element

zu A, d.h. $A /_h A = [\lambda A/\lambda A, \rho A/\rho A] = [1,1] = 1$

Bemerkung: Die hyperbolische Division $/_h$ unterscheidet sich wesentlich von der Intervalldivision / , denn es gilt offenbar $A /_h B = A/\overline{B}$ für $B \notin N = T \cup \overline{T}$

c) Das Distributivgesetz wird wie folgt erfüllt:

Mit $A,B,C \in \mathbb{H}$ ist

$$A(B+C) = [\lambda A(\lambda B + \lambda C), \rho A(\rho B + \rho C)]$$

$$= [\lambda A \lambda B + \lambda A \lambda C, \rho A \rho B + \rho A \rho C]$$

$$= AB + AC$$

<u>Satz 2.14:</u> Darstellung und Verträglichkeitseigenschaften der Operatoren λ und ρ in ($\mathbb{H}$,+,).

(1) $\lambda(A) = [0,1]\overline{A} + [1,0]A$; $\rho(A) = [0,1]A + [1,0]\overline{A}$

(2) $\rho(AB) = \rho(A)\rho(B)$; $\lambda(AB) = \lambda(A)\lambda(B)$

Speziell gilt für $\bigwedge_{\xi \in \mathbb{R}}$ $\rho(\xi A) = \xi\rho(A) \;\backslash\; \lambda(\xi A) = \xi\lambda(A)$

(3) $\rho(\overline{A}) = \lambda(A)$ und $\rho(-A) = -\lambda(A)$ (Intervall-Minus)

(4) $\rho(1/A) = 1/\lambda(A)$

Bew.: Zu (1) genügt der Beweis für λ wegen (3).

$$[o,1]\overline{A}+[1,o]A=[o,1][\rho A,\lambda A]+[1,o]|\lambda A,\rho A|$$
$$=[o,\lambda A]+[\lambda A,o]$$
$$=[\lambda A,\lambda A] = \lambda(A)$$

(2) ist eine unmittelbare Folge aus der Definition 2.12 des hyperbolischen Produktes.

(3) $\rho(\overline{A}) = \rho([\rho(A),\lambda(A)]) = \lambda(A)$

$\rho(-A) = \rho([-\rho(A),-\lambda(A)]) = -\lambda(A)$

Satz 2.15: Konjugations- und Vorzeichenregel

(1) $\overline{AB} = \overline{A}\ \overline{B}$

(2) $(-A)B = -(A\overline{B})$ und $(-_h A)B = -_h(AB)$

Bew.: (1) ergibt sich aus:

$$\overline{[\lambda A,\rho A][\lambda B,\rho B]} = [\rho A\rho B,\lambda A\lambda B] = \overline{A}\ \overline{B}$$

(2) Entsprechend folgt für das Intervall-Minus:

$$(-A)B = [-\rho A,-\lambda A][\lambda B,\rho B] = [-\ A\ B,\ -\ A\ B]$$
$$= -[\lambda A\rho B,\rho A\lambda B] = -(A\overline{B})$$

und die Vorzeichenregel für das hyperbolische Minus gilt allgemein in Ringen.

Bemerkung: Aus Satz 2.9 folgt, daß das Additionsinverse zu A $-\overline{A}$ ist, also gilt die Beziehung zwischen Intervall-Minus und hyperbolischen Minus:

$$-\overline{A} = -_h A \quad .$$

Stellt man Summe und hyperbolisches Produkt in den μ-δ-Koordinaten dar, so zeigen sich interessante Analogien zu den komplexen Zahlen:

(a) $(\mu_1,\delta_1) + (\mu_2,\delta_2) = (\mu_1+\mu_2,\delta_1+\delta_2)$

(b) $(\mu_1,\delta_1)\ (\mu_2,\delta_2) = (\mu_1\mu_2+\delta_1\delta_2,\mu_1\delta_2+\mu_2\delta_1)$

(c) $1/_{h}(\mu,\delta) = \dfrac{(\mu,\delta)}{\mu^2 - \delta^2}$ für $\mu^2 \neq \delta^2$

(d) $\mu^2 - \delta^2 = (\mu,\delta)\overline{(\mu,\delta)}$

Stellt man die Intervalle $A \in \mathbb{H}$ dar über der Orthonormalbasis (μ_o,δ_o) mit $\mu_o = (1,o) = [1,1]$ und $\delta_o = (o,1) = [-1,1]$ und $A = a\mu_o + b\delta_o$, so erhält man durch Setzung:

$$\mu_o\mu_o = \mu_o$$
$$\delta_o\delta_o = \mu_o$$
$$\mu_o\delta_o = \delta_o$$

den hyperbolischen Ring.

Benutzt man die in den Sätzen 1.13, 1.14 und 1.15 aufgezeigten Eigenschaften des hyperbolischen Ringes $(\mathbb{H},+,\ )$ zur Darstellung der Tabelle 2.4, so treten nun in Tabelle 2.5 einige bemerkenswerte innere Strukturen zu Tage. Die Identität der Tabellen 2.4 und 2.5 läßt sich leicht anhand der Verknüpfungsregeln im hyperbolischen Ring nachweisen.

$A\in$ \ $B\in$	S_r	T	S_l	$\overline{T}$
S_r	AB	$\rho(A)B$	$\overline{A}\ B$	$\lambda(A)B$
T	$A\rho(B)$	$[\min(\lambda A\rho B,\rho A\lambda B),$ $\max(\lambda A\lambda B,\rho A\rho B)]$	$\overline{A}\lambda(B)$	o
S_l	$A\ \overline{B}$	$\lambda(A)\overline{B}$	$\overline{A}\ \overline{B}$	$\rho(A)\overline{B}$
$\overline{T}$	$A\lambda(B)$	o	$\overline{A}\rho(B)$	$[\max(\lambda A\lambda B,\rho A\rho B),$ $\min(\lambda A\rho B,\rho A\lambda B)]$

Tabelle 2.5

Bemerkung: Zum Nachweis der Konsistenzbedingungen wäre die Verknüpfungstabelle 2.5 noch viel zu aufwendig, denn man müßte sämtliche 16 Fälle und z. B. beim Assoziativgesetz sogar 64 Fälle durchprüfen. In den folgenden Schritten wird deshalb die Verknüpfungstabelle immer mehr vereinfacht und dafür weitere Eigenschaften des Produktes aufgesucht, an denen (K1),(K2) und (K3) leicht nachweisbar sind.

Seien nach dem n-ten Schritt die Eigenschaften $(E_1),(E_2),\dots(E_n)$ und eine Verknüpfungstabelle T_n vorgelegt, so dürfen sich diese Eigenschaften nicht widersprechen und es muß stets möglich sein, die Tabelle 2.5 vermöge der Tabelle T_n und den Eigenschaften $(E_1),\dots(E_n)$ zu rekonstruieren.

Dies wird im folgenden dadurch gesichert, daß beim n+1-ten Schritt nur eine solche zusätzliche Eigenschaft (E_{n+1}) zugelassen wird, die

(I) $(E_1)\ldots(E_n)$ nicht widerspricht

und eine Tabelle T_{n+1} so angegeben wird,

(II) daß aus dieser Tabelle zusammen mit (E_{n+1}) die Ausgangstabelle T_n stets reproduzierbar ist.

Ist T_n schließlich genügend einfach, so werden an dieser und an den Eigenschaften $(E_1)\ldots(E_n)$ die Konsistenzbedingungen nachgewiesen werden. Vermöge der oben geforderten Bedingung gelten die Konsistenzbedingungen dann auch in Tabelle 2.5.

Satz 2.16:

$(H,*)$ ist nach Tabelle 2.5 kommutativ.

Bew.: Die Tabelle 2.5 ist symmetrisch bezüglich der Hauptdiagonalen, d.h. $A*B=B*A$.

Satz 2.17:

Die Tabelle 2.5 kann mit Hilfe der Eigenschaft

(E_1) $(H,*)$ ist kommutativ und der

(T_1) Tabelle 2.6

vollständig beschrieben werden.

Bew.: Sei $A * B$ vorgelegt, aber $A * B$ sei nicht aus Tabelle 2.6 ablesbar, so ist stets $B * A$ aus dieser Tabelle ablesbar. Vermöge (E_1) gilt aber $A * B = B * A$, so daß $A * B$ berechenbar ist.

$A \in$ \ $B \in$	S_r	T	S_l	$\overline{T}$
S_r	AB	$\rho(A)B$	$\overline{A}\ \ B$	$(A)B$
T		$[\min(\lambda A\rho B, \rho A\lambda B), \max(\lambda A\lambda B, \rho A\rho B)]$	$\overline{A}\lambda(B)$	$\circ$
S_l			$\overline{A}\ \ \overline{B}$	$\rho(A)\overline{B}$
$\overline{T}$				$[\max(\lambda A\lambda B, \rho A\rho B), \min(\lambda A\rho B, \rho A\lambda B)]$

Tabelle 2.6

Lemma 2.19:

Der Konjugations- und Negativoperator bilden folgende Sektoren aufeinander ab:

(1) $\overline{A} \in S_r \ \wedge\ -A \in S_l$, $\qquad \overline{A} \in S_l \ \wedge\ -A \in S_r$

$A \in S_r \qquad\qquad\qquad\qquad A \in S_l$

(2) $\overline{A} \in \overline{T} \ \wedge\ -A \in T$, $\qquad \overline{A} \in T \ \wedge\ -A \in \overline{T}$

$A \in \overline{T} \qquad\qquad\qquad\qquad A \in \overline{T}$

<u>Bew.:</u> Diese Aussagen sind aus der Definition 2.11 bzw. aus Abbildung 2.3 und der Spiegelungseigenschaft der Operatoren unmittelbar ersichtlich.

Zur weiteren Vereinfachung der Tabelle 2.6 folgender

<u>Satz 2.20:</u> (Konjugationsregel)

In Tabelle 2.6 gilt $\bigwedge_{A,B \in \mathbb{H}} \overline{(A*B)} = \overline{A} * \overline{B}$

<u>Bew.:</u> Durch Anwendung des Lemmas 2.19, Satz 2.14((3):$\rho(\overline{A})=\lambda(A)$) und Satz 2.15 ($\overline{AB}=\overline{A}\ \overline{B}$) lassen sich die einzelnen Fälle leicht realisieren.

Zur Demonstration zwei Fälle:

Sei $A \in S_r \wedge B \in S_r$, so ist $\overline{(A*B)} \overset{\text{nach Tabelle 2.6}}{=} \overline{AB}$

$= \overline{A}\ \overline{B} \overset{\text{nach Lemma 2.19 u. Tab. 2.6}}{=} \overline{A} * \overline{B}$

Sei $A \in S_r \wedge B \in T$, so ist

$\overline{(A*B)} \overset{\text{nach Tab. 2.6}}{=} \overline{(A)B} \overset{\text{nach Satz 2.15}}{=} \overline{(A)}\overline{B}$

$\overset{\text{nach Satz 1.14}}{=} (\overline{A})\overline{B} \overset{\text{nach Satz 2.15 u. Tab. 2.6}}{=} \overline{A} * \overline{B}$

usw.

Satz 2.21:

Vermöge von Satz 2.20 kann das Produkt folgendermaßen charakterisiert werden:

(E_1) $(\mathbb{H}, *)$ ist kommutativ

(E_2) $$\bigwedge_{A,B \in \mathbb{H}} \overline{A * B} = \overline{A} * \overline{B}$$

(T_2) Tabelle 2.8

$A \in$ \\ $B \in$	S_r	T	S_l
S_r	AB	$\rho(A)B$	$\overline{A}B$
T		$[\min(\lambda A \rho B, \rho A \lambda B), \max(\lambda A \lambda B, \rho A \rho B)]$	$\overline{A}\lambda(B)$
S_l			$\overline{A}\,\overline{B}$
$\overline{T}$		o	

Tabelle 2.8

Bew.: (E_2) widerspricht (E_1) nicht, da sie unabhängig sind. So bleibt zu zeigen, daß Tabelle 2.6 aus (T_2) und (E_2) reproduzierbar ist. Es fehlen die Verknüpfungen für beliebiges $A \in \mathbb{H}$ und $B \in \overline{T}$. Da $\overline{A} \in \mathbb{H}$ und $\overline{B} \in T$ ist mit (E_2) $A * B = \overline{\overline{A} * \overline{B}}$ und $\overline{A} * \overline{B}$ ist aus Tabelle 2.8 ablesbar.

Satz 2.22:

In Tabelle 2.8 gilt die Vorzeichenregel:

$$\bigwedge_{A,B \in \mathbb{H}} (-A) * B = -(A * B)$$

Bew.: Diese Aussage gilt in Tabelle 2.8 wie sich mit Hilfe des Lemmas 2.19, Satz 2.14 ($\rho(-A)=-\lambda(A)$) und Satz 2.15 ($(-A)B=-AB$) zeigen läßt. Zur Demonstration sei der Fall mit $A \in T \wedge B \in T$ herausgegriffen: Nach Lemma 2.19 ist dann $-A \in T$

und damit

$$
\begin{aligned}
(-A) * B &\overset{\text{nach Tabelle 2.8}}{=} [\min(\lambda(-A)\rho(B), \rho(-A)\lambda B)), \\
&\qquad \max(\lambda(-A)\lambda(B), \rho(-A)\rho(B))] \\
&\overset{\text{Satz 2.14}}{=} [\min(-\rho(A)\rho(B), -\lambda(A)\lambda(B)), \\
&\qquad \max(-\rho(A)\lambda(B), -\lambda(A)\rho(B))] \\
&= [-\max(\rho(A)\rho(B), \lambda(A)\lambda(B)), \\
&\qquad -\min(\rho(A)\lambda(B), \lambda(A)\rho(B))] \\
&\overset{\text{Tabelle 2.8}}{=} -(A * B)
\end{aligned}
$$

Satz 2.23:

Damit läßt sich für das Produkt folgende Beschreibung geben:

(E_1) $(H, *)$ ist kommutativ

(E_2) $\bigwedge_{A,B \in \mathbb{H}} \overline{(A * B)} = \overline{A} * \overline{B}$

(E_3) $\bigwedge_{A,B \in \mathbb{H}} (-A) * B = -(A * B)$

(T_3) Tabelle 2.9

$A \in$ \ $B \in$	S_r	T	$\overline{T}$
S_r	AB	$\rho(A)B$	
T		$[\min(\lambda A \rho B, \rho A \lambda B), \max(\lambda A \lambda B, \rho A \rho B)]$	o

Tabelle 2.9

<u>Bew.:</u> (E_3) widerspricht nicht den Eigenschaften (E_1) und (E_2) da sie unabhängig sind. Tabelle 2.8 ist aus Tabelle 2.9 mit (E_3) ergänzbar:

Sei $A \in \mathbb{H}$ und $B \in S_1$, so ist nach Lemma 2.19 $-B \in S_r$ und vermöge (E_3) ist dann $A * B = -(A * (-B))$ aus Tabelle 2.9 bestimmbar.

Nach diesen Vorbereitungen soll nun endlich der Beweis von Satz 2.12 nachgeholt werden. Nach der in Anschluß an Tabelle 2.5 gemachten Bemerkung genügt es, die Konsistenzbedingungen an (E_1), (E_2), (E_3) und (T_3) aus Satz 2.23 nachzuweisen.

<u>Bew.:</u>

(K1) verträgt sich mit (E_1) und ist von (E_2) und (E_3) unabhängig.

(K1) verträgt sich mit Tabelle 2.9 wie folgt:

(Die Ziffern beziehen sich auf Satz 1.1 (b).)

(o) klar, da höchstens wieder verallgemeinerte Intervalle auftreten

(1) aus (E_1) folgt (1)

(2) Zum Beweis der Assoziativität wird hauptsächlich benutzt, daß der hyperbolische Ring $(\mathbb{H},+,\ )$ assoziativ und das Intervallprodukt kommutativ ist. Folgende Fälle treten auf:

2.1 $A,B,C \in S_r$ so sind sämtlich auftretende Produkte $A * B * C = ABC$ mit der hyperbolischen Form identisch und damit assoziativ.

2.2 $A \in S_r$, $B \in S_r$, $C \in T$

$$(A * B) * C = \rho(AB)C \quad \text{nach Satz 2.14}$$
$$= \rho(A)\rho(B)C = \rho(A)(\rho(B)C) = \rho(A)(B * C) = A * (B * C)$$
$$\text{da } B * C \in T .$$

2.3 $A \in S_r$, $B \in T$, $C \in T$

$$(A * B) * C = (\rho(A)B) * C \quad \text{da} \quad \rho(A)B \in T \quad \text{folgt:}$$
$$= [\min(\rho(\rho(A)B)\lambda(C), \lambda(\rho(A)B)\rho(C)),$$
$$\max(\lambda(\rho(A)B)\lambda(C), \rho(\rho(A)B)\rho(C))]$$
$$\text{da } \rho(A) \geq o \text{ und nach Satz 2.14}$$
$$= \rho(A)[\min(\rho B\lambda C, \lambda B\rho C), \max(\lambda B\lambda C, \rho B\rho C)]$$
$$= \rho(A)(B * C) \quad \text{da} \quad B * C \in T$$
$$= A * (B * C)$$

2.4 Sind $A,B,C \in T$, so liegen alle Verknüpfungen in $\mathbb{R}$ und nach Satz 1.1 gilt in $(\mathbb{R}, *)$ das Assoziativgesetz.

2.5 In den Fällen, in denen ein Faktor in T und ein zweiter Faktor in $\overline{T}$ liegt, ist das Ergebnis stets o, was dem Assoziativgesetz nicht widerspricht.

(3) $1 \in S_r$ ist Neutralelement, denn nach Tabelle 2.9 ist mit $A:=1$

$1 * B = 1B = B$ und $1 * B = \rho(1)B = 1B = B$

$B \in S_r$ $\qquad$ $B \in T$

Die Eindeutigkeit ergibt sich wie üblich mit der Annahme, daß $Y \in H$ ein weiteres sei, so liefert $Y * 1 = Y = Y * 1 = 1$.

(K2) (K2) ist unabhängig von (E_1), (E_2) und (E_3) und Tabelle 2.9 erfüllt (K2), wie mit Hilfe der folgenden Lemmata gezeigt werden kann.

<u>Lemma:</u> Die Teilmengeneigenschaft ist transitiv, d.h. mit $A \subseteq B \wedge B \subseteq C$ und mit beliebigem $D \in \mathbb{H}$ gelte $A * D \subseteq B * D \wedge B * D \subseteq C * D$ so ist mit $A \subseteq C$ auch $A * D \subseteq C * D$.

<u>Bew.:</u> Dies ist eine unmittelbare Folgerung aus der Transitivität von $\subseteq$.

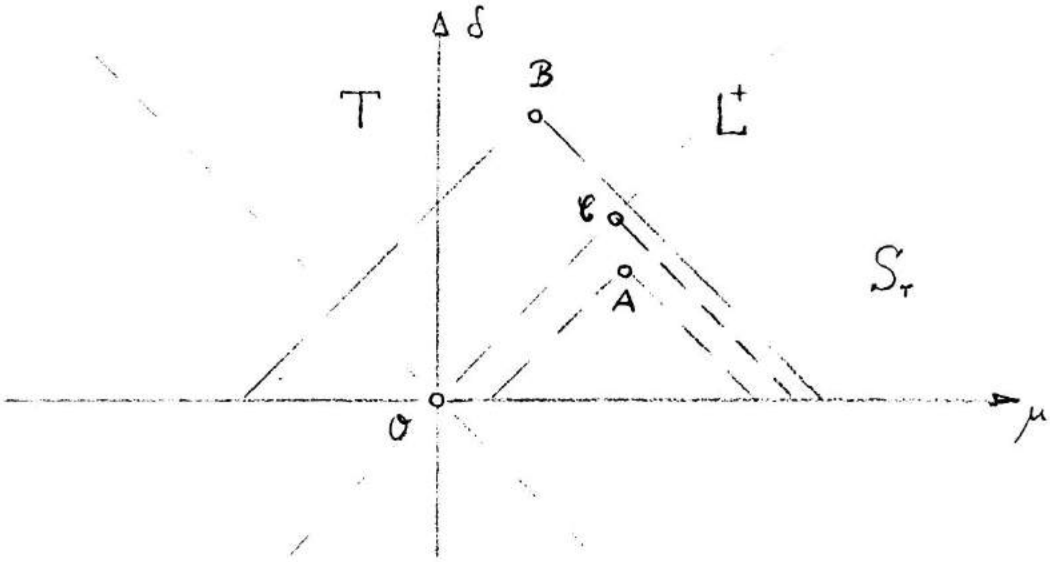


Abb. 2.4

<u>Lemma:</u> Ist $A \in S_r \wedge B \in T$ und $A \subsetneq B$ so gibt es stets ein $C \in L^+$ mit $A \subseteq C \subsetneq B$.

<u>Bew.:</u> In der μ-δ-Ebene liegt z. B. ein solches C gerade auf dem Schnittpunkt des Kegelmantels von B mit der Winkelhalbierenden L^+ . Vergl. auch Abbildung 2.4. Es ist mit $C = [o, \rho B] \wedge A \in S_r \wedge A \subsetneq B$ stets $o \leq \lambda A \wedge \rho A \leq \rho B$ und $\lambda B \leq o \wedge \rho B \leq \rho B$.

Der Fall $A \in T$ und $B \in S_r$ und $A \subsetneq B$ kann nie auftreten, da $A \supsetneq o \not\subseteq B$.

Sei im folgenden stets $A \subsetneq B$ so gilt:

(1) $A \in S_r$, $B \in S_r$, $C \in S_r$ dann ist

$A * C = AC = [\lambda A \lambda C, \rho A \rho C]$

$B * C = BC = [\lambda B \lambda C, \rho B \rho C]$ und mit $\lambda B \leq \lambda A \wedge \rho A \leq \rho B \wedge o \leq \lambda C \leq \rho C$

folgt $\lambda B \lambda C \leq \lambda A \lambda C \wedge \rho A \rho C \leq \rho B \rho C \Longrightarrow A * C \subseteq B * C$.

(2) $A \in S_r$, $B \in S_r$, $C \in T$ dann ist

$A * C = \rho(A)C$ und $B * C = \rho(B)C$ und mit $o \leq \rho(A) \leq \rho(B)$ folgt

$\rho(A)C \subseteq \rho(B)C$

(3) $A \in S_r$, $B \in S_r$, $C \in \overline{T}$ läßt sich zurückführen auf Fall (2) durch Konjugation, denn es ist dann $\overline{A} \in S_r$, $\overline{B} \in S_r$, $\overline{C} \in T$ und wegen $\overline{A} \supseteq \overline{B}$ folgt $\overline{A} * \overline{C} \supseteq \overline{B} * \overline{C}$ und damit $A * C \subseteq B * C$

(4) $A \in T$, $B \in T$, $C \in S_r$ so folgt

$A * C = A\rho C = A * \rho(C)$

$B * C = B\rho C = B * \rho(C)$ und dieser Fall ist mit $\rho(C) \in \mathbb{R}^+$ zurückgeführt auf $(\mathbb{IR}, *)$ wo (K2) bereits gilt.

(5) $A, B, C \in T$ ist mit $T \subseteq \mathbb{IR}$ ebenfalls ganz in $(\mathbb{IR}, *)$ wo stets (K2) gilt.

(6) $A \in T$, $B \in T$, $C \in \overline{T}$ ergibt einfach $A * C = B * C = o$.

(7) Ist $A \in S_r \wedge B \in T$, so gibt es nach dem letzten Lemma ein $C \in L^+ = S_r \cap T$ mit $A \subseteq C \wedge C \subseteq B$, und nach den Fällen (1),(2) und (3) ist somit $A * D \subseteq C * D$ und nach (4),(5) und (6) ist $C * D \subseteq B * D$ und somit gilt nach dem ersten Lemma: $A * D \subseteq B * D$.

Der Nachweis der Stetigkeit verläuft wie folgt:

(K3) ist verträglich mit (E_1), (E_2) und (E_3), da kommutieren und spiegeln an den Achsen μ und δ stetige Operationen sind.

Innerhalb der Sektoren ist das Ergebnis des Produktes $A * B = [\phi(\lambda A, \rho A, \lambda B, \rho B), \psi(\lambda A, \rho A, \lambda B, \rho B)]$ wobei ψ und ϕ selbst stetige Funktionen in allen Argumenten sind ($\min(x,y)$ und $\max(x,y)$ sind stetige Funktionen).

Der Nachweis reduziert sich damit auf die Nahtstellen nach Tabelle 2.9: (1) L^+ und

(2) $\{o\} = T \cap \overline{T}$.

<u>Zu (1):</u> Es ergibt sich mit $A,B \in L^+$, also $A=[o,x]$, $B=[o,y]$:

$$(*) \quad A * B = AB = \rho(A)B = [\min(\lambda A \rho B, \rho A \lambda B), \max(\lambda A \lambda B, \rho A \rho B)] = [o,xy],$$

d.h. je nach Annäherung aus dem betreffenden Sektor trifft eine der drei Produktdarstellungen der Gleichung (*) zu und diese sind auf L^+ identisch.

<u>Zu (2):</u> Mit $A \in T$ und $B \in \overline{T}$ ist stets $A * B = o$ was ja mit beliebiger Annäherung von A oder B gegen o stets o bleibt.

<u>Bemerkung 2.24:</u>

(1) Wie Tabelle 2.5 zeigt, ist das Intervallprodukt nicht nullteilerfrei. Es gilt im Gegensatz zu Satz 1.1(e)

$$\bigwedge_{A \in T} \bigwedge_{B \in \overline{T}} A * B = [o,o] = o \quad \text{auch für } A \neq o \wedge B \neq o.$$

(2) Das Inklusionsdistributivgesetz in $\mathbb{IR}$ gilt nicht mehr in ganz $\mathbb{H}$. Es gilt vielmehr:

$$\bigwedge_{A,B,C \in \mathbb{IR}} A * (B+C) \subseteq A * B + A * C$$

$$\bigwedge_{A,B,C \in \overline{\mathbb{IR}}} A * (B+C) \supseteq A * B + A * C$$ (eine unmittelbare Folge des Satzes 2.20)

2.3.4 Fortsetzung der Division /

Satz 2.25:

Die Division ist mit $N:=T\cup\overline{T}$ und

$$\bigwedge_{\substack{A\in \mathbb{H}\\ B\in \mathbb{H}\setminus N}} A/B:=A*1/B \quad\wedge\quad \bigwedge_{B\in \mathbb{H}\setminus N} 1/B:=[1/\rho B,1/\lambda B]$$

konsistent fortgesetzt.

Bew.: (K1) klar nach Satz 1.1 d) Punkt o) und 1), da die Darstellungen identisch sind. Die Divisionsringoid-eigenschaften sind ebenfalls erfüllt mit:

(D7) $A/1=A*1/1=A*1=A$

(D8) $o/A=o*1/A=o$ für $A\notin N$

(D9) mit $x:=-1$ ist $(-1)*(A/B)=(-1)*A*1/B$

$$=((-1)*A)*1/B$$
$$=A*(-1)/B$$
$$=A*1/((-1)*B)$$

(K2) und (K3) gelten bereits für das Produkt, so daß diese nur noch für die Inversion 1/B nachzuweisen sind.

(K2) $A\subseteq B$ d.h. $\lambda B\leq\lambda A\wedge\rho A\leq\rho B$ so folgt, da die Vorzeichen in $\mathbb{H}\setminus N$ jeweils gleich sind $1/\lambda B\geq 1/\lambda A\wedge 1/\rho A\geq 1/\rho B$ also ist $1/A\subseteq 1/B$.

(K3) Die Stetigkeit ergibt sich aus der Stetigkeit der Funktion $1/x$ für $x \in R$ und $x \neq o$.

Satz 2.26:

(1) Mit $N = T \cup \overline{T}$ ist $(\mathbb{H}\setminus N, *)$ eine kommutative Halbgruppe mit 1 und Inversen in $\mathbb{H}\setminus N$.

(2) $$\bigwedge_{A,B \in H\setminus N} 1/(A * B) = 1/A * 1/B$$

(3) $$\bigwedge_{A,B \in \mathbb{H}\setminus N} 1/(A/B) = B/A$$

(4) Für $A \in \mathbb{H}\setminus N$ gilt:

$$1/A = \frac{A}{A * \overline{A}} = \frac{A}{\lambda(A)\rho(A)}$$

(5) $$d(1/A) = \frac{d(A)}{A * \overline{A}} = \frac{d(A)}{\lambda(A)\rho(A)}$$

Bew.:

(1) Es bleibt wegen der Gültigkeit von (K1) nur noch den Nachweis zu erbringen, daß alle Elemente $A \in \mathbb{H}\setminus N$ stets ein Inverses besitzen. Tatsächlich gilt:

$$\bigwedge_{A \in H\setminus N} A * \overline{1/A} = 1 \text{ , denn es ist } \overline{1/A} = [1/\rho A, 1/\lambda A]$$

und da mit $A \in S_r$ stets $\overline{1/A} \in S_r$ bzw. mit $A \in S_l$ stets $\overline{1/A} \in S_l$ ist, folgt in jedem Falle $A * \overline{1/A} = [1,1] = 1$

(2) Aus $1/(A * B) = X$ folgt nach (1)

$1 = X * \overline{(A * B)} = X * \overline{A} * \overline{B} \Longrightarrow$

$X * \overline{A} = 1/B$ und schließlich $X = 1/B * 1/A$

(3) $X = 1/(1/B)$ ergibt $X * \overline{1/B} = 1 \Longrightarrow X = B$

für $B \in \mathbb{H} \setminus N$. Es ist somit $1/(1/B) = B$.

Für $1/(A/B) = 1/(A * 1/B)$ folgt wegen (2)

$= 1/A * 1/(1/B)$ und nach der Vorüberlegung:

$= 1/A * B = B/A$

(4) $1/A = [1/\rho(A),\ 1/\lambda A] = \left[\dfrac{\lambda(A)}{\rho(A)\lambda(A)},\ \dfrac{\rho(A)}{\rho(A)\lambda(A)}\right]$,

da mit $A \in \mathbb{H} \setminus N$ stets $\rho(A)\lambda(A) \geq o$ ist :

$= A * 1/\rho(A)\lambda(A) = A * 1/(A \times \overline{A})$

(5) $d(1/A) = 1/\lambda(A) - 1/\rho(A) = \dfrac{\rho(A)-\lambda(A)}{\lambda(A)\rho(A)} = \dfrac{d(A)}{A * \overline{A}}$

für $A \in \mathbb{H} \setminus N$.

<u>Lemma 2.27:</u>

In $\mathbb{IR}$ gilt die allgemeine Darstellung der Intervallverknüpfung nach [Kulisch 69]:

Sei zu gegebenen $A, B \in \mathbb{IR}$ und für eine beliebige Verknüpfung

$\circ \in \{+,-,*,/\}$: $P(\circ, A, B) = P := \{\lambda A \circ \lambda B, \rho A \circ \lambda B, \lambda A \circ \rho B, \rho A \circ \rho B\}$ so gilt

$A \circ B = [\min(P), \max(P)]$

$A, B \in \mathbb{IR}$

Eine entsprechende Formel gilt in $\overline{\mathbb{IR}}$:

$$\bigwedge_{A,B \in \overline{\mathbb{IR}}} A \circ B = [\max(P), \min(P)]$$

Bew.: Anhand von Satz 2.14 läßt sich leicht einsehen, daß $P(\circ,\overline{A},\overline{B}) = P(\circ,A,B)$ ist. Benutzt man die Eigenschaft (E_2) aus Satz 2.23, dann ist $\overline{A},\overline{B} \in \mathbb{IR}$ und da P invariant bezüglich Konjugationen ist, ergibt sich

$$A \circ B = \overline{\overline{A} \circ \overline{B}} = \overline{[\min(P), \max(P)]} = [\max(P), \min(P)] \ .$$

2.4 Vervollständigung von $\mathbb{H}$

Wie Abschnitt I,4. zeigt, muß der zugrundegelegte Raum ein vollständiger Verband sein, um ohne Schwierigkeiten auf die entsprechenden endlichen Raster übergehen zu können. In Abschnitt 6 wird eine Klasse von endlichen Intervallrastern $\mathbb{RH}$ eingeführt, die leicht auf Rechenanlagen realisierbar sind.

Auch hier tritt der Fall ein, daß bei Hinzunahme der uneigentlichen Zahlen p und $-p$ zu $\mathbb{R}$ der zugehörige Intervallraum $\mathbb{H}^* := \mathbb{IR}^* \cup \overline{\mathbb{IR}^*}$ einige globale Eigenschaften einbüßt. Jedoch bedeutet das auch hier keine wesentliche Einschränkung, denn es wird $(\mathbb{H}^* \setminus \{0\}, +, -, *, /, \subseteq)$ zu einem vollständigen, durchweg isoton geordneten Divisionsringoid.

Satz 2.28:

Legt man den Vollverband $\mathbb{R}^* = \mathbb{R} \cup \{-p,p\}$ aus Abschnitt I,4. zugrunde, so ist $\mathbb{H}^* := \mathbb{IR}^* \cup \overline{\mathbb{IR}^*}$ ebenfalls ein Vollverband bezüglich $\subseteq$.

Bew.: $(\mathbb{H}^*, \subseteq)$ ist nach Konstruktion und nach Definition 2.5 ordnungsisomorph [1] zu $(\mathbb{R}^* \times \mathbb{R}^*, \leq)$ und somit ein Vollverband. Es gilt insbesondere
$i(\mathbb{H}^*) = [-p,p] =: P \in \mathbb{H}^*$ und $o(\mathbb{H}^*) = [p,-p] = \overline{P} \in \mathbb{H}^*$.

Definition und Satz 2.29:

$(\mathbb{H}^*, \{o\}, +, -, *, /, \subseteq)$ ist gemäß der Verknüpfungstabelle 2.10 und den folgenden Vereinbarungen ein durchweg isoton geordnetes Divisionsringoid mit den ausgezeichneten Elementen $\{-1,o,1\}$.

a) Die Addition und Subtraktion gelten formal nach Satz 2.7 und Satz 2.8 unter Verwendung der Tabelle 4.1 aus Kapitel I.

b) Die Multiplikation gilt formal nach Satz 2.12 und Tabelle 4.2 aus Kapitel I.

c) Mit $L := L^+ \cup \overline{L^+} \cup L^- \cup \overline{L^-}$ und $S := S_r \cup S_l$ ist die Inversion in $\mathbb{H}^* \setminus L$ erklärt:

$$\bigwedge_{A \in \mathbb{H}^* \setminus L} 1/A := [1/\rho A, 1/\lambda A]$$

d) Für $A \in L \setminus \{o\}$ gelten die beiden Fälle:

$1/[o,\rho A] = [1/\rho(A), \; p \cdot \mathrm{sign}(\rho(A))]$

$1/[\lambda A,o] = [p \cdot \mathrm{sign}(\lambda(A)), \; 1/\lambda(A)]$

und allgemein die Division nach Tabelle 2.10 . Für das auftretende Produkt gilt b) und damit die Tabelle 4.2 aus Kapitel I. So ist z. B. $o * p = o$ usw. .

---
1) Genaueres siehe Abschnitt 2.6

$B \in$ / $A \in$	$S \setminus \{o\}$	$T \setminus L$	$\overline{T} \setminus L$
S	$A * 1/B$	$A * 1/B$	$A * 1/B$
T	$A * 1/B$	$[-p,p]$	$A * 1/B$
$\overline{T}$	$A * 1/B$	$A * 1/B$	$[p,-p]$

Tabelle 2.10

Bew.: (D1), (D2), (D3), (D4) und (D5) behalten ihre Gültigkeit aufgrund der Tatsache, daß sie bereits in $\mathbb{H}$ gelten und sich unmittelbar auf $\mathbb{H}^*$ übertragen.

(D6) klar, da durch Hinzunahme von p und -p keine weiteren Elemente x mit $x*x=1$ entstehen.

(D7) Da $1 \in S \setminus \{o\}$ gilt die erste Spalte aus Tabelle 2.10, also $A/1 = A * 1/1 = A$.

(D8) Da $o \in T$ und nach Voraussetzung $A \notin N = T \cup \overline{T}$, ist $o/A = o * 1/A = o$.

(D9) $-1 \in S$ und nach Voraussetzung $B \notin N$, also ist

$$(-1) * (A/B) = (-1) * A * 1/B = ((-1) * A) * 1/B$$
$$= A * (-1)/B = A * 1/((-1) * B)$$

(OD6) Die Einschließungseigenschaft bzw. die Durchweg-Isotonie gilt für die Operationen + - * in $\mathbb{H}$ und für / in $\mathbb{H} \setminus N$ aufgrund der Konsistenzforderung (K2) bei der Konstruktion von $\mathbb{H}$. Bei allen geführten Nachweisen von (K2) können auch ohne weiteres die Grenzelemente $-p$ und $+p$ bzw. $[-p,p]$ und $[p,-p]$ gesetzt werden, so daß (OD6) auch in $\mathbb{H}^*$ gilt.

Eine Ausnahme bildet der Nullraum $N = T \cup \overline{T}$. Dort gilt für Divisoren $A \in N \setminus \{o\}$ keine Inklusionsisotonie im Sinne von (OD6) mehr. Dies stellt natürlich keinen Widerspruch zu (K2) dar, da ja für die Fortsetzung von $\mathbb{H}$ die Elemente $A \in N$ als Divisoren ausgeschlossen wurden (vergl. Abschnitt 2.3.4). Erst die in diesem Abschnitt vorgenommene Vervollständigung von $\mathbb{H}$ zu $\mathbb{H}^*$ ermöglichte die Fortsetzung der Division auf Divisoren aus $N \setminus \{o\}$.
Bemerkenswerterweise erlaubt gerade die Verletzung der Inklusionsisotonie, wie sie im folgenden dargestellt wird, eine geschlossene Darstellung des Newtonverfahrens für mehrere Nullstellen.
Zur Beschreibung der Ordnungsstruktur bei der Abbildung $X \to 1/X$ in $N \setminus \{o\}$ erweist sich folgende Halbordnungsrelation als nützlich.

<u>Definition und Satz 2.30:</u>

a) $A \prec B \Leftrightarrow \lambda A \sqcup \rho A \leq \rho B \sqcap \lambda B$ ist eine Ordnungsrelation 2. Art, d. h. sie ist nicht reflexiv, aber asymmetrisch und transitiv.

d. h. es gilt:

(01') $A \lessdot B \Longrightarrow B \not\lessdot A$ und

(02') $A \lessdot B \wedge B \lessdot C \Longrightarrow$

$\lambda A \sqcup \rho A \leq \lambda B \sqcap \rho B \leq \lambda B \sqcup \rho B \leq \lambda C \sqcap \rho C$

$\Longrightarrow A \lessdot C$

b) Gelegentlich wird auch die in [Moore 69] und [Kulisch 72] eingeführte Ordnungsrelation benutzt:

$A \underline{\ll} B \Longleftrightarrow \lambda A \leq \lambda B \wedge \rho A \leq \rho B$

<u>Lemma 2.32:</u>

Es gelten folgende Eigenschaften:

a) $A \lessdot B \Longleftrightarrow \overline{A} \lessdot \overline{B}$, $\quad A \underline{\ll} B \Longleftrightarrow \overline{A} \underline{\ll} \overline{B}$

b) $A \lessdot B \Longleftrightarrow -A \gtrdot -B$, $\quad A \underline{\ll} B \Longleftrightarrow -A \underline{\gg} -B$

c) $A \lessdot B \Longrightarrow A \underline{\ll} B$

d) $A \underline{\ll} B \wedge A \underline{\ll} \overline{B} \Longleftrightarrow A \lessdot B$

e) $A \underline{\ll} B \Longrightarrow A + C \underline{\ll} B + C$

<u>Bew.:</u> Die Behauptungen a), b) und c) sind einfache Folgerungen aus Definition 2.30

d) ergibt sich aus folgender Äquivalenzkette:

$A \underline{\ll} B \wedge A \underline{\ll} \overline{B} \Longleftrightarrow$

$\lambda A \leq \lambda B \wedge \rho A \leq \rho B \wedge \lambda A \leq \rho B \wedge \rho A \leq \lambda B \Longleftrightarrow$

$\lambda A \leq \lambda B \sqcap \rho B \wedge \rho A \leq \lambda B \sqcap \rho B \Longleftrightarrow$

$\lambda A \sqcup \rho A \leq \lambda B \sqcap \rho B \Longleftrightarrow A \lessdot B$

e) $A \underline{\ll} B \Longleftrightarrow \lambda A \leq \lambda B \wedge \rho A \leq \rho B \Longrightarrow$

$\lambda A + \lambda C \leq \lambda B + \lambda C \wedge \rho A + \rho C \leq \rho B + \rho C \Longrightarrow$

$A + C \underline{\ll} B + C$.

<u>Satz 2.33:</u>

In $N\setminus\{0\} = T \cup \overline{T}\setminus\{0\}$ gelten folgende Ordnungsbeziehungen bei der Inversion $X \to 1/X$. Die verschiedenen Fälle sind in der folgenden Tabelle 2.11 zusammengefaßt. (A und B sind jeweils in der entsprechenden Zeile und Spalte Elemente aus $S_r\setminus\{0\}$, $T\setminus\{0\}$, $S_l\setminus\{0\}$ und $\overline{T}\setminus\{0\}$. Zu jedem der möglichen Fälle $A \subsetneq B$ oder $A \supsetneq B$ gibt die entsprechende Tabellenstelle an, ob $1/A \supsetneq 1/B$ oder $1/A \subsetneq 1/B$ oder $1/A \underline{<<} 1/B$ oder $1/A \underline{>>} 1/B$ gilt. Mit Ausnahme der Fälle in der Hauptdiagonalen der Tabelle ist nur höchstens einer der Fälle $A \subsetneq B$ oder $B \subsetneq A$ möglich. <u>In der Hauptdiagonalen jedoch steht zuoberst stets der Fall für $A \supsetneq B$ und darunter der Fall $A \subsetneq B$).</u>

A∈ \ B∈	$S_r\setminus\{0\}$	$T\setminus\{0\}$	$S_l\setminus\{0\}$	$\overline{T}\setminus\{0\}$
$S_r\setminus\{0\}$	$1/A \supsetneq 1/B$ $1/A \subsetneq 1/B$	$1/A \underline{>>} 1/B$		$1/A \underline{>>} 1/B$
$T\setminus\{0\}$	$1/A \underline{<<} 1/B$	$1/A \supsetneq 1/B$ $1/A \subsetneq 1/B$	$1/A \underline{>>} 1/B$	$1/A \subsetneq 1/B$
$S_l\setminus\{0\}$		$1/A \underline{<<} 1/B$	$1/A \supsetneq 1/B$ $1/A \subsetneq 1/B$	$1/A \underline{<<} 1/B$
$\overline{T}\setminus\{0\}$	$1/A \underline{<<} 1/B$	$1/A \supsetneq 1/B$	$1/A \underline{>>} 1/B$	$1/A \supsetneq 1/B$ $1/A \subsetneq 1/B$

Tabelle 2.11

<u>Bew.:</u> Mit Hilfe der Eigenschaften

$A \subseteq B \Longleftrightarrow -A \subseteq -B \Longleftrightarrow \overline{A} \supseteq \overline{B}$ nach Satz 2.6 und 2.8 und

$A \underline{\ll} B \Longleftrightarrow -A \underline{\gg} -B \Longleftrightarrow \overline{A} \underline{\ll} \overline{B}$ nach Lemma 2.32

und mit Hilfe von Lemma 2.19 läßt sich Tabelle 2.11 reduzieren auf eine Minimaltabelle Tabelle 2.11'.

$A \in$ \ $B \in$	$S_r \setminus \{o\}$	$T \setminus \{o\}$	$\overline{T} \setminus \{o\}$
$S_r \setminus \{o\}$	$1/A \supseteq 1/B$		$1/A \underline{\gg} 1/B$
$T \setminus \{o\}$		$1/A \supseteq 1/B$	$1/A \subseteq 1/B$

Tabelle 2.11'

Diese Tabelle ist nur angelegt für den Fall $A \supseteq B$!

Diese verbleibenden vier Fälle beweisen sich wie folgt:

Fall 1: $A \in S_r \setminus L \quad B \in S_r \setminus L$ folgt schon aus Satz 2.25 wegen (K2).

Fall 2: $A, B \in T \setminus \{o\}$, so gilt:

$$A \supseteq B \Longleftrightarrow \lambda A \le \lambda B \wedge \rho B \le \rho A$$
$$\Longleftrightarrow 1/\lambda B \le 1/\lambda A \wedge 1/\rho A \le 1/\rho B$$
$$\Longleftrightarrow \lambda(1/A) \le \lambda(1/B) \wedge \rho(1/B) \le \rho(1/A)$$
$$\Longleftrightarrow 1/A \supseteq 1/B$$

Fall 3: $A \in T \setminus \{o\} \quad B \in \overline{T} \setminus \{o\}$, so gilt:

$$A \supseteq B \Longleftrightarrow \lambda A \le \lambda B \wedge \rho B \le \rho A$$
$$\text{da} \quad \lambda A \le o \le \lambda B \wedge \rho B \le o \le \rho A$$
$$\Longleftrightarrow 1/\lambda A \le 1/\lambda B \wedge 1/\rho B \le 1/\rho A$$
$$\Longleftrightarrow 1/A \subseteq 1/B$$

Bew.: Fall 4: $A \in S \setminus \{o\}$ $B \in \overline{T} \setminus L$ ergibt:

$$A \supseteq B \Longleftrightarrow \lambda A \leq \lambda B \wedge \rho B \leq \rho A$$

da $\rho B \leq o \leq \rho A$

$$\Longleftrightarrow 1/\lambda B \leq 1/\lambda A \wedge 1/\rho B \leq 1/\rho A$$

$$\Longleftrightarrow \rho(1/B) \leq \rho(1/A) \wedge \lambda(1/B) \leq \lambda(1/A)$$

$$\Longleftrightarrow 1/B \underline{\ll} 1/A$$

womit Satz 2.33 bewiesen ist.

Der folgende Satz ist zur Formulierung des modifizierten Newtonverfahrens von Bedeutung.

Satz 2.34:

a) Sei $A \in T \setminus L \wedge B \in S_r \setminus L$ bzw. $B \in S_l \setminus L$, so gilt $A \supseteq B \wedge A \supseteq \overline{B} \Longleftrightarrow 1/A \prec 1/B$ bzw. $1/A \succ 1/B$

b) entsprechend sei $A \in \overline{T} \setminus L \wedge B \in S_r \setminus L$ bzw. $B \in S_l \setminus L$, so gilt: $A \subseteq B \wedge A \subseteq \overline{B} \Longleftrightarrow 1/A \prec 1/B$ bzw. $1/A \succ 1/B$

c) $A, B \in \mathbb{IR}$, $\widetilde{A}, \widetilde{B} \in \mathbb{IR}$, und ist $A \subseteq \widetilde{A}$ $B \subseteq \widetilde{B}$, so gilt:

(c.1) $B \notin T \wedge \widetilde{B} \notin T \Longrightarrow A/B \subseteq \widetilde{A}/\widetilde{B}$

(c.2) $B \in \mathbb{IR} \setminus T \wedge \widetilde{B} \in T \setminus \{o\} \wedge \widetilde{A} \in T$ $A/B \subseteq \widetilde{A}/\widetilde{B}$

(c.3) $B \in \mathbb{IR} \setminus T \wedge \hat{B} \in T \setminus \{o\} \wedge A \in \mathbb{IR} \setminus T \wedge \widetilde{A} \in \mathbb{IR} \setminus T \Longrightarrow$

für $A/B \in S_r$: $A/B \succ \widetilde{A}/\widetilde{B}$

für $A/B \in S_l$: $A/B \prec \widetilde{A}/\hat{B}$

(c.4) $B \in T \setminus \{o\} \wedge \widetilde{B} \in T \setminus \{o\} \Longrightarrow A/B \subseteq \widetilde{A}/\widetilde{B}$

Bew.: Nach Tabelle 2.10 folgt im Falle a) o.B.d.A für $B \in S_r \setminus L$:

$$A \supseteq B \iff 1/A \ll 1/B \quad \text{und}$$

$$A \supseteq \bar{B} \iff 1/A \ll 1/\bar{B} \iff 1/A \ll \overline{1/B}$$

und nach Lemma 2.32 d) folgt daher:

$$A \supseteq B \wedge A \supseteq \bar{B} \iff 1/A \ll 1/B \wedge 1/A \ll \overline{1/B}$$

$$\iff 1/A < 1/B$$

entsprechend beweisen sich die restlichen Fälle von a) und b).

c) (c.1) gilt nach Satz 2.12 und der Eigenschaft, daß $A/B = A * 1/B$ ist.

(c.2) Es ist $A/B \in \mathbb{IR}$ und somit ist mit

$$\tilde{A}/\tilde{B} = [-p,p] = P \qquad A/B \subseteq P = \tilde{A}/\tilde{B}\ .$$

(c.3) $A \subseteq \tilde{A} \in S_r \implies A \in S_r$ und damit ist

$$A/B = [\lambda A/\rho B,\ \rho A/\lambda B] \quad \text{und}$$

$$\tilde{A}/\tilde{B} = \tilde{A} * (1/\tilde{B}) = \lambda\tilde{A}(1/\tilde{B})$$

Nun ist nach a) $1/\tilde{B} < 1/B$ und da

$$o \leq \lambda\tilde{A} \leq \lambda A \implies$$

$$\lambda\tilde{A}(1/\tilde{B}) < \lambda A(1/\tilde{B}) < \lambda A(1/B)$$

$$\tilde{A}/\tilde{B} < [\lambda A/\rho B, \lambda A/\lambda B] \ll [\lambda A/\rho B, \rho A/\lambda B]$$

da $\lambda A \leq \rho A \wedge \lambda B > o$.

Die restlichen Fälle lassen sich mit Hilfe der Negativoperation und Lemma 2.32 auf diesen Fall (c.3) zurückführen.

(c.4) $B, \tilde{B} \in T \setminus \{o\}$ und $A, \tilde{A} \in T$ so gilt nach Satz 2.33 $A * 1/B \subseteq A * 1/\tilde{B} \subseteq \tilde{A} * 1/\tilde{B}$.

Für den Fall $A \notin T$, $\tilde{A} \in T$ ist

$A * 1/B \subseteq P = \tilde{A}/\tilde{B}$ und für den Fall

$A, \tilde{A} \in T$ ist $A/B = P \subseteq P = \tilde{A}/\tilde{B}$.

Zur Veranschaulichung der Ordnungsbeziehungen bei der Inversion in $N \setminus \{o\}$ sind in Abbildung 2.12 und der folgenden Zusammenstellung die wesentlichsten Fälle für ein festes $A \in T \setminus \{o\}$ aufgeführt:

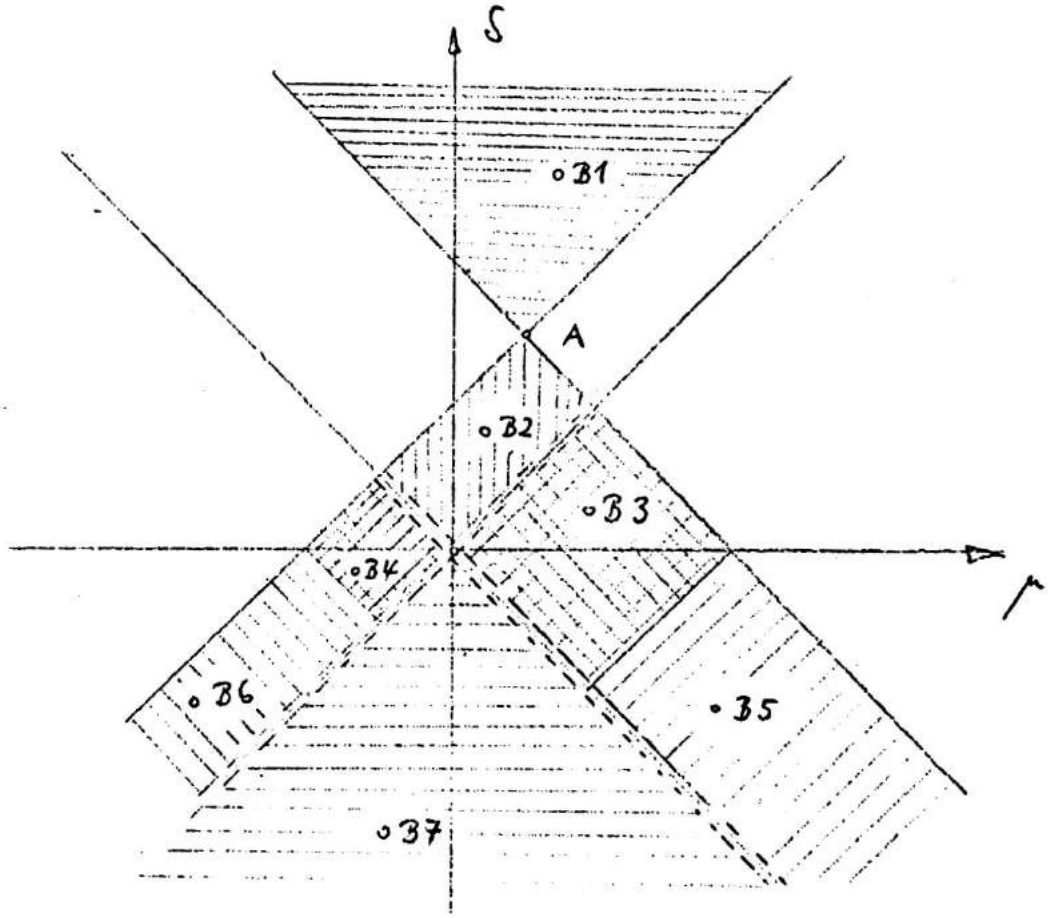


Abb. 2.12

Abbildung 2.12 weist folgende 7 Fälle auf:

$B1 \supseteq A \supseteq B2 \Longleftrightarrow 1/B1 \supseteq 1/A \supseteq 1/B2$

$B3 \subseteq A \Longleftrightarrow 1/B3 > 1/A$

$B4 \subseteq A \Longleftrightarrow 1/B4 < 1/A$

$B5 \subseteq A \Longleftrightarrow 1/B5 \underline{\gg} 1/A$

$B6 \subsetneq A \Longleftrightarrow 1/B6 \underline{\ll} 1/A$

$B7 \subsetneq A \Longleftrightarrow 1/B7 \supseteq 1/A$

Es ist bemerkenswert, daß die Inversion je nach Konstellation die drei verschiedenen Ordnungsrelationen $\subsetneq$, $\underline{\ll}$ und $<$ ineinanderüberführt. Natürlich kann aus Tabelle 2.10 auch die Umkehrung abgelesen werden, etwa:

$A \underline{\ll} B$ für $A \in T\setminus\{o\}$ und $B \in S_r\setminus\{o\}$, so ist dies äquivalent mit $1/A \subseteq 1/B$, da $1/A \in \overline{T}\setminus\{o\} \wedge 1/B \in S_r\setminus\{o\}$.

An Satz 2.26 anschließend kann nun im Hinblick auf die erweiterte Division in $\mathbb{H}^*\setminus\{o\}$ folgendes ergänzt werden:

<u>Satz 2.35:</u>

a) $\bigwedge_{\substack{A \in \mathbb{H}^*\setminus N \\ B \in \mathbb{H}^*\setminus\{o\}}} 1/(A * B) = 1/A * 1/B$

b) $\bigwedge_{A \in \mathbb{H}^*\setminus\{o\}} 1/(1/A) = A$

c) $\bigwedge_{\substack{A \in \mathbb{H}^*\setminus\{o\} \\ B \in \mathbb{H}^*\setminus N}} (\, 1/(A/B) = B/A \,,\; 1/(B/A) = A/B \,)$

Bew.:

a) Für $A,B \in \mathbb{H} \setminus N$ ist die Aussage bereits in Satz 2.26 gezeigt. Es bleiben damit die vier Fälle:

$A \in S_r \setminus L \wedge B \in T \setminus \{0\}$

$A \in S_r \setminus L \wedge B \in \overline{T} \setminus \{0\}$

$A \in S_l \setminus L \wedge B \in T \setminus \{0\}$

$A \in S_l \setminus L \wedge B \in \overline{T} \{0\}$.

Unter Verwendung der Eigenschaften (E_2) und (E_3) aus Satz 2.23 und Lemma 2.19 lassen sich alle vier Fälle auf einen reduzieren, z. B. auf den ersten Fall.

Für diesen gilt:
Mit $A \in S_r \setminus L \wedge B \in T \setminus \{0\}$ gilt
$1/A \in S_r \wedge 1/B \in \overline{T}$ und nach Tabelle 2.5:

$$1/A * 1/B = \lambda(1/A)1/B = (1/\rho(A))(1/B) = 1/\rho(A)B$$
$$= 1/(A * B)$$

b) Sei $A = [\lambda A, \rho A]$, so ist
$1/A = |1/\rho A,\ 1/\lambda A|$ und schließlich
$1/(1/A) = [1/(1/\lambda A),\ 1/(1/\rho A)] = [\lambda A, \rho A] = A$

c) Mit $A \in \mathbb{H} \setminus \{0\}$ ist nach b):
$A = 1/(1/A) \in \mathbb{H} \setminus \{0\}$ und nach a) somit:
$A/B = A * 1/B = 1/(1/A) * 1/B =$
$= 1/(1/A * B) = 1/(B/A)$;
entsprechend folgt $B/A = 1/(A/B)$.

Es ist seltsam, daß die Fälle a) und c) nicht allgemein für $A,B \in \mathbb{H}^* \setminus \{o\}$ erweitert werden können. Daß dem so ist, beweist folgendes Gegenbeispiel:

Sei $A = [-1,2]$ und $B = [-4,5]$, so ist

$1/(A * B) = 1/[-8,10] = [1/10,-1/8]$, aber

$1/A * 1/B = [1/2,-1] * [1/5,-1/4] = [+1/4,-1/5]$

Bemerkung 2.36:

1. Bei der Division bzw. bei der Inversion ist bemerkenswert, daß in $\mathbb{H}$ die Inversion nur für Elemente aus $S_r \cup S_l \setminus L$ definierbar war. In $\mathbb{H}^*$ jedoch kann die Inversion auch für Elemente aus $T \cup \overline{T} \setminus \{o\}$ sinnvoll definiert werden. Eine Motivierung und anschauliche Bedeutung dieser Inversion zeigt Abschnitt 5 beim Newton-Verfahren. Vergleiche auch hierzu den Abschnitt 2.5.

2. Im Gegensatz zu den Inversen in $S_r \cup S_l \setminus L$ sind die '' Inversen '' in $T \cup \overline{T} \setminus \{o\}$ keine eigentlichen Inverse mehr, denn es gilt:

   $A/\overline{A}=1$ für $A \in S \setminus L$, aber $A/\overline{A}=[\min(\lambda A/\rho A,\rho A/\lambda A),1]$ für $A \in T \setminus \{o\}$

   $A/\overline{A}=[1,\min(\lambda A/\rho A,\rho A/\lambda A)]$ für $A \in \overline{T} \setminus \{o\}$

   d.h., betrachtet man die Funktion $\upsilon(A):=A/\overline{A}$ so ist der Bildbereich $\upsilon(\mathbb{H}^*) = \{1\} \cup \{[\kappa,1] \mid -p \le \kappa \le 1\} \cup \{[1,\kappa] \mid -p \le \kappa \le 1\}$.
   Der Bildbereich ist in Abbildung 2.13 angedeutet.

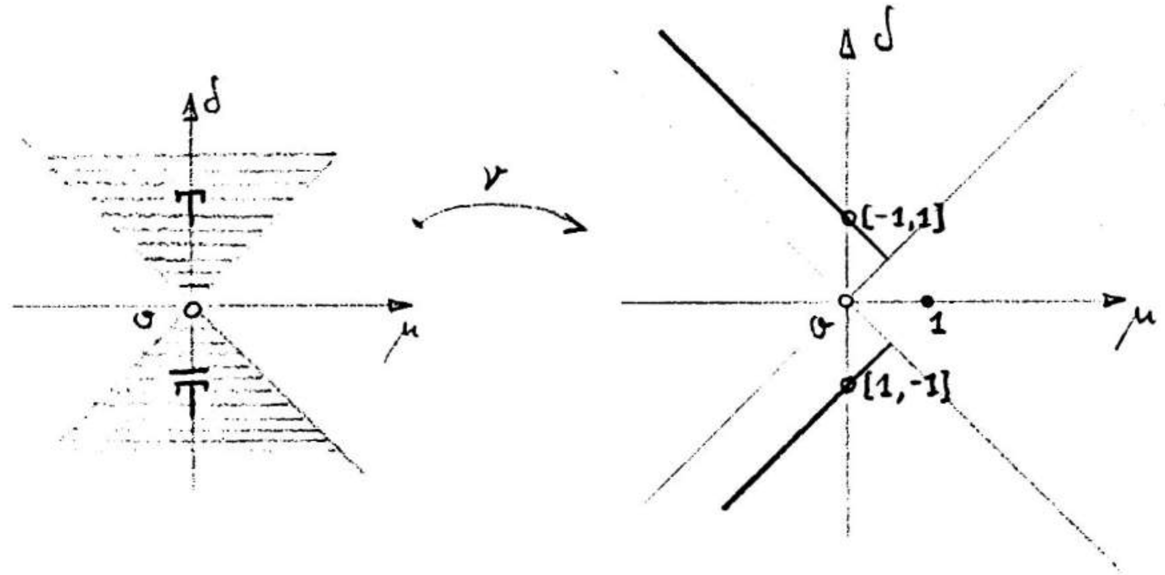


Abb. 2,13

2.5 Interpretation der nichtregulären Intervalle $\overline{II\mathbb{R}}$

Wie der Abschnitt 6 dieses Kapitels zeigen wird, erfüllt der formal erweiterte Intervallraum $\mathbb{H}^*$ die anfänglich gestellte Forderung, antitone Einschließungen von Fixpunkten zu berechnen. Unbefriedigend bleibt aber bisher, daß man mit Größen rechnet, die keinen Sinn zu haben scheinen. Die Inversion gibt uns einen Schlüssel, wie zu einer Interpretation der nichtregulären Intervalle möglicherweise zu gelangen ist.

Die algebraischen Operationen + - * führen nicht aus den regulären Intervallen $II\mathbb{R}$ hinaus, d.h. ($II\mathbb{R}$,+,-,*) ist abgeschlossen. Anders jedoch die Division bzw. die Inversion. Bildet man z. B. $1/[-2,3]$ so erhält man nach Satz 2.29: $1/[-2,3]=[1/3,-1/2]$, also ein Element aus $\overline{II\mathbb{R}}$, ein nichtreguläres Intervall.

Die Inversion ist zunächst definiert in $\mathbb{IR} \setminus T$. Für ein A mit $o \notin A$ ist $1/A = h(\{1/a \mid a \in A\})$. Da die Funktion $1/a$ stetig für $a \neq o$ ist und A einfach zusammenhängend, so ist auch die Menge $\{1/a \mid a \in A\}$ einfach zusammenhängend. Es ist deshalb $1/A = \{1/a \mid a \in A\}$.

Definiert man aber auf dieselbe Weise die Inversion für $A \in T$, also mit $o \in A$, so zeigt Abbildung 2.14 , daß kein einfach zusammenhängendes sondern zwei getrennte Ergebnisintervalle $(-\infty, 1/\lambda A]$ und $[1/\rho A, \infty)$ auftreten[1] .

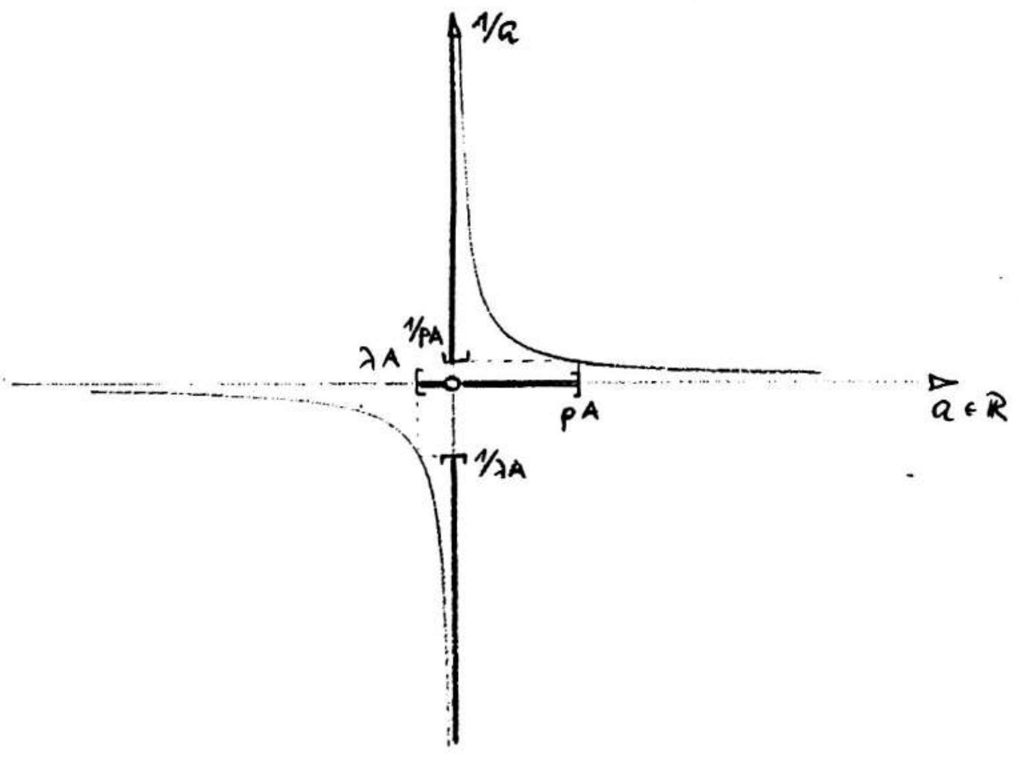


Abb. 2.14

---

[1] Die runden Klammern können in IH^* durch eckige Klammern und $+\infty$ und $-\infty$ durch $+p$ und $-p$ ersetzt werden.

Nach Satz 2.29 jedoch ergibt sich: $1/A=[1/\rho A,1/\lambda A] \in \overline{T} \subsetneq \overline{IR}$, also lasch ausgedrückt gerade das ''einfach zusammenhängende Loch'' oder genauer $\overline{1/A} = [1/\lambda A,1/\rho A]$ ist gerade das abgeschlossene Komplement zu dem oben anschaulich berechneten Ergebnis.

Der Konjugation könnte man daher die Bedeutung der Komplementbildung bezüglich der Grundmenge $\mathbb{C}\mathbb{I}\mathbb{R}$ zuordnen (vergl. [Kahan 68]).

Diese Interpretation führt, wie folgende Beispiele zeigen, zu erheblichen Komplikationen bezüglich der algebraischen Struktur:

Sei $A = [2,1] = (-\infty,1] \cup [2,\infty)$ und $B = [0,0.5]$, so ist $A \ast B=\{ab \mid a \in A \wedge b \in B\}=(-\infty,\infty)$, ein Ergebnis, das zwar dem Einschließungsgesetz genügt, aber für Iterationsverfahren ungeeignet ist. Z. B. würde das Iterationsverfahren $A:=A \ast [0,0.5]+1$ mit Fixpunkt $\hat{A} = [1,2]$ nicht einmal konvergieren, mit $A_0 = [2,1] = (-\infty,1] \cup [2,\infty)$, denn es wäre dann $A_1 = (-\infty,\infty)$. Dagegen liefert das Produkt bzw. die Summe nach der formalen Definition das sinnvolle Ergebnis $A_1 = [2,1] \ast [0,0.5] + 1 = [1,1.5]$ und es ist $\hat{A} \supsetneq A_1 \supsetneq A_0$ usw..

Nun ist es allerdings falsch, deshalb obige Interpretation zu verwerfen, weil sie etwa stets zu unbrauchbaren Ergebnissen führt, sondern die Schuld liegt darin, daß die algebraische Verknüpfung falsch fortgesetzt wurde. Würde man nämlich das Produkt in diesen Fällen über das Komplement berechnen, also mit $A = \mathbb{R} \setminus \overline{A}$, $B = \mathbb{R} \setminus \overline{B}$ und

$A \ast B = \mathbb{R} \setminus \{a'b' | a' \in \overline{A} \wedge b' \in \overline{B}\}$ so ist tatsächlich $[2,1] \ast [0,0.5] = \mathbb{R} \setminus ((-\infty,0] \cup [0,5,\infty)) = [0,0.5]$, also dasselbe wie nach Tabelle 2.5. Doch auch diese Form versagt in anderen Fällen, etwa im Beispiel $A = [1,2]$ und $B = [0.5,0] = (-\infty,0] \cup [0.5,\infty)$. Es ist dann $\mathbb{R} \setminus \{a'b' | a' \in \overline{A} \wedge b' \in \overline{B}\} = \mathbb{R} \setminus (-\infty,\infty) = \emptyset$ dafür jedoch stimmt es in der ursprünglichen Form: $\{ab | a \in A \wedge b \in B\} = (-\infty,0] \cup [0.5,\infty) = [0.5,0]$.

Diese verwirrenden Schwierigkeiten zu umgehen ist der Anlaß, $\mathbb{IR}$ formal unter geeigneten Konsistenzbedingungen nach $\mathbb{H}$ algebraisch fortzusetzen.

Die anfangs gegebene Interpretation der nicht regulären Intervalle braucht deshalb nicht verworfen zu werden, sondern wie der Abschnitt 2.8 zeigen wird, entstehen die Widersprüche dadurch, daß die Komplexbildung nach Definition I, 2.5 direkt übertragen wurde, was nicht in dieser Weise erlaubt ist.

2.6 Fortsetzung der Supremums- und Infimumsoperation

Auf Grund der formalen Fortsetzung der Ordnungsrelation von $\mathbb{IR}$ auf $\mathbb{H}$ nach Definition 2.5, ist es naheliegend, zu vermuten, daß sich die entsprechenden Supremums- und Infimumsoperationen in analoger Weise formal fortsetzen.

Diese Vermutung bestätigt sich mit Hilfe folgender Ordnungsisomorphie.

Sei $(\mathbb{R}, \leq)$ der natürlich geordnete Raum der reellen Zahlen, $(\mathbb{R}', \leq')$ der "entgegengesetzt" geordnete Raum der reellen Zahlen, d. h. es gilt:

$\mathbb{R}' = \mathbb{R}$ und mit $a,b \in \mathbb{R}'$ ist $a \leq' b \Longleftrightarrow b \leq a$.

Der Produktraum $(\mathbb{R}' \times \mathbb{R}, \leq)$ sei wie folgt geordnet:

Mit $(a,b), (c,d) \in \mathbb{R}' \times \mathbb{R}$ ist

$(a,b) \leq (c,d) \Longleftrightarrow a \leq' c \;\wedge\; b \leq d$ nach Definition von $(\mathbb{R}', \leq')$: $\Longleftrightarrow c \leq a \;\wedge\; b \leq d$.

$(\mathbb{H}, \subseteq)$ ist dann entsprechend Definition 2.5 ordnungsisomorph zu $(\mathbb{R}' \times \mathbb{R}, \leq)$:

Mit

$$\phi : \mathbb{H} \to \mathbb{R}' \times \mathbb{R}, \text{ d. h. } \bigwedge_{A \in \mathbb{H}} \phi(A) = (\lambda A, \rho A),$$

$$\phi^{-1} : \mathbb{R}' \times \mathbb{R} \to \mathbb{H}, \text{ d. h. } \bigwedge_{(a,b) \in \mathbb{R}' \times \mathbb{R}} \phi^{-1}((a,b)) = [a,b]$$

und

$$\begin{aligned} A \subseteq B &\Longleftrightarrow \phi(A) \leq \phi(B) \\ &\Longleftrightarrow (\lambda A, \rho A) \leq (\lambda B, \rho B) \\ &\Longleftrightarrow \lambda B \leq \lambda A \;\wedge\; \rho A \leq \rho B \end{aligned}$$

erhält man gerade die Definition von 2.5 .

$(\mathbb{R}' \times \mathbb{R}, \leq)$ ist als Produktraum von Verbänden ein Verband und somit ist $(\mathbb{H}, \subseteq)$ ein Verband. Die Infimums- bzw.

Supremumsoperation, die in $\mathbb{H}$ mit $\bar{\sqcap}$ bzw. $\underline{\sqcup}$ bezeichnet werden soll, überträgt sich vermöge des Isomorphismus ϕ wie folgt:

$\sqcap$ bzw. $\sqcup$ bezeichne die Infimums- bzw. Supremumsoperation in $(\mathbb{R}, \leq)$ bzw. (dual vertauscht) in $(\mathbb{R}', \leq')$.

Es gilt:

$$\phi(A \underline{\sqcup} B) = \sup_{(\mathbb{R}' \times \mathbb{R}, \leq')}(\phi(A), \phi(B)) =$$

$$= \sup_{(\mathbb{R}' \times \mathbb{R}, \leq')}((\lambda A, \rho A), (\lambda B, \rho B)) =$$

$$= (\lambda A \sqcap \lambda B, \rho A \sqcup \rho B)$$

d. h.

$$A \underline{\sqcup} B := \phi^{-1}(\phi(A \underline{\sqcup} B)) =$$

$$= \phi^{-1}((\lambda A \sqcap \lambda B, \rho A \sqcup \rho B))$$

$$= [\lambda A \sqcap \lambda B, \rho A \sqcup \rho B].$$

Entsprechend folgt:

$$A \bar{\sqcap} B := [\lambda A \sqcup \lambda B, \rho A \sqcap \rho B]$$

Zusammenfassend gilt daher:

Definition und Satz 2.40:

Es gilt in $(\mathbb{H}, \subseteq)$:

$$\bigwedge_{A,B \in \mathbb{H}} \quad \begin{aligned} A \mathbin{\underline{\cup}} B &::= \sup(A,B) = [\lambda A \sqcap \rho B, \rho A \sqcup \rho B] \\ A \mathbin{\pi} B &::= \sup(A,B) = [\lambda A \sqcup \rho B, \rho A \sqcap \rho B] \end{aligned}$$

Da $\mathbb{IR} \subseteq \mathbb{H}$ auch als Teilmenge $\mathbb{IR} \subseteq \mathbb{PR}$ der Potenzmenge $(\mathbb{PR}, \subseteq)$ aufgefaßt werden kann, ist es in manchen Situationen von Vorteil, eine Beziehung zwischen den mengentheoretischen Operationen $\cap$ und $\cup$ und den Operationen π und $\underline{\cup}$ herzustellen.

Satz 2.41:

In $\mathbb{IR}$ gilt

a) $$\bigwedge_{\substack{A,B \in \mathbb{IR} \\ A \cap B \neq \emptyset}} \quad A \cap B = A \mathbin{\pi} B \quad \wedge \quad A \cup B = A \mathbin{\underline{\cup}} B$$

b) $$\bigwedge_{A,B \in \mathbb{IR}} \quad A \cap B = \emptyset \iff (\lambda A \sqcup \lambda B) > (\rho A \sqcap \rho B)$$

<u>Bew.:</u> Bis auf Vertauschungen von A und B sind folgende 3 Fälle zu unterscheiden:

Fall 1: $\lambda A \leq \rho A < \lambda B \leq B$ so ist $\lambda A \sqcup \lambda B = \lambda B > \rho A = \rho A \sqcap \rho B$, also tritt b) in Kraft, denn tatsächlich ist $A \cap B = \emptyset$.

Fall 2: $\lambda A \leq \lambda B \leq \rho A \leq \rho B$ so ist $A \cap B \neq \emptyset$, also tritt a) in Kraft und es ist $A \cup B = [\lambda A, \rho B]$ und $A \cap B = [\lambda B, \rho A]$.

Fall 3: $\lambda B \leq \lambda A \leq \rho A \leq \rho B$ so ist $A \cap B = A \neq \emptyset$ und es gilt a) mit $A \cup B = B = [\lambda B, \rho B]$ und $A \cap B = A = [\lambda A, \rho A]$.

In den Fällen also, wo a) aus Satz 2.41 zutrifft, lassen sich $\cup$ und $\cap$ auf Supremums- und Infimumsoperationen($\sqcup$, $\sqcap$) der Intervallgrenzen zurückzuführen.

<u>Satz 2.42:</u>

Für alle $A,B,C \in \mathbb{H}$ gelten folgende Eigenschaften:

(1) $\overline{A \underline{\cup} B} = \overline{A} \,\overline{\cap}\, \overline{B}$

(2) $$\bigwedge_{A,B \in \mathbb{IR}} A \cap B \neq \emptyset \Longleftrightarrow (A \,\overline{\cap}\, B = A \cap B \wedge A \underline{\cup} B = A \cup B)$$
$$\Longleftrightarrow A \,\overline{\cap}\, B \in \mathbb{IR}$$

(3) $A \subseteq B \Longleftrightarrow (A \sqcup B = B \wedge A \sqcap B = A) \Longrightarrow A \sqcup B \supseteq A \wedge A \sqcap B \subseteq A$

(4) $(\mathbb{H}^*, \sqcup)$ und $(\mathbb{H}^*, \sqcap)$ sind kommutative Halbgruppen mit den Neutralelementen $[-p,p] = P$ bzw. $[p,-p] = \overline{P}$.

(5) $(\mathbb{H}^*, \sqcup, \sqcap)$ ist distributiv

(6) $A \subseteq B \wedge C \subseteq D \Longrightarrow A \sqcup C \subseteq B \sqcup C$

(7) $\sqcap$ und $\sqcup$ sind stetige Operationen

Bew.:

(1) $\overline{[\lambda A \sqcup \lambda B, \rho A \sqcap \rho B]} = [\rho A \sqcap \rho B, \lambda A \sqcup \lambda B]$ und wegen $\rho \overline{A} = \lambda A$ folgt
$= [\lambda \overline{A} \sqcap \lambda \overline{B}, \rho \overline{A} \sqcup \rho \overline{B}]$

(2) Aus Satz 2.41 a) folgt unmittelbar:
$A \cap B \neq \emptyset \Longrightarrow (A \sqcap B = A \cap B \wedge A \sqcup B = A \cup B)$
Umgekehrt, ist $A \sqcap B = A \cap B$, so ist wegen Satz 2.41 b) dies nur für $A \sqcap B \in \mathbb{IR}$ möglich, also ist $A \cap B \in \mathbb{IR}$, also ist $A \cap B \neq \emptyset$.

(3) $A \subseteq B$, d.h. nach Definition 2.6 es ist $\lambda B \leq \lambda A \wedge \rho A \leq \rho B$, also ist $\lambda A \sqcup \lambda B = \lambda A \wedge \lambda A \sqcap \lambda B = \lambda B \wedge$
$\rho A \sqcap \rho B = \rho A \wedge \rho A \sqcup \rho B = \rho B$, womit die Behauptung
$A \sqcap B = [\lambda A, \rho A] = A$ und $A \sqcup B = [\lambda B, \rho B] = B$ folgt.

(4) Abgeschlossenheit, Kommutativität und Assoziativität von $\sqcup$ und $\sqcap$ ergeben sich aus den entsprechenden Eigenschaften von $\sqcup$ und $\sqcap$ in $(R, \leq)$.

Insbesondere gilt:

Wegen $A \subseteq P$ folgt nach (3) $A \,\overline{\sqcap}\, P = A$ und

wegen $A \supseteq \overline{P}$ folgt nach (3) $A \,\underline{\sqcup}\, \overline{P} = A$.

(5) $(A \,\underline{\sqcup}\, B) \,\overline{\sqcap}\, C = [\lambda A \sqcap \lambda B, \rho A \sqcup \rho B] \,\overline{\sqcap}\, C$

$$= [(\lambda A \sqcap \lambda B) \sqcup \lambda C, (\rho A \sqcup \rho B) \sqcap \rho C]$$

Da nach Definition 110 aus Kapitel I der linear geordnete Verband $(\mathbb{R}^*, \leq)$ in $\sqcup$ und $\sqcap$ distributiv ist, folgt weiter:

$$= [(\lambda A \sqcup \lambda C) \sqcap (\lambda B \sqcup \lambda C), (\rho A \sqcap \rho C) \sqcup (\rho B \sqcap \rho C)]$$

$$= [\lambda A \sqcup \lambda C, \rho A \sqcap \rho C] \,\underline{\sqcup}\, [\lambda B \sqcup \lambda C, \rho B \sqcap \rho C]$$

$$= (A \,\overline{\sqcap}\, C) \,\underline{\sqcup}\, (B \,\overline{\sqcap}\, C).$$

Analog läuft der Beweis für $(A \,\overline{\sqcap}\, B) \,\underline{\sqcup}\, C = (A \,\underline{\sqcup}\, C) \,\overline{\sqcap}\, (B \,\underline{\sqcup}\, C)$

(6) Dies folgt durch sukzessive Anwendung von:

$a \leq b \wedge c \leq d \Rightarrow a \sqcup c \leq b \sqcup d$ usw. auf die Intervallgrenzen.

(7) $\sqcup$ und $\sqcap$ sind stetige Operationen was sich auf $\underline{\sqcup}$ und $\overline{\sqcap}$ unmittelbar überträgt.

<u>Bemerkung 2.43:</u>

Die modifizierte Vereinigungs- und Durchschnittsbildung läßt sich in der μ-δ -Ebene leicht veranschaulichen und konstruieren:

<u>Konstruktion:</u> Die Kanten bzw. Mantellinien der den beiden Intervallen A und B zugeordneten Kegel schneiden sich für $A \nsubseteq B \wedge B \nsubseteq A$ in genau 2 Punkten.

In den Fällen, wo $A \subseteq B$ oder $B \supseteq A$ ist, sind als die beiden Schnittpunkte A und B selbst zu nehmen. $A \sqcup B$ ist dann der Schnittpunkt mit dem größeren δ-Wert und $A \sqcap B$ ist der Schnittpunkt mit dem kleineren δ-Wert (vergl. Abbildung 2.16).

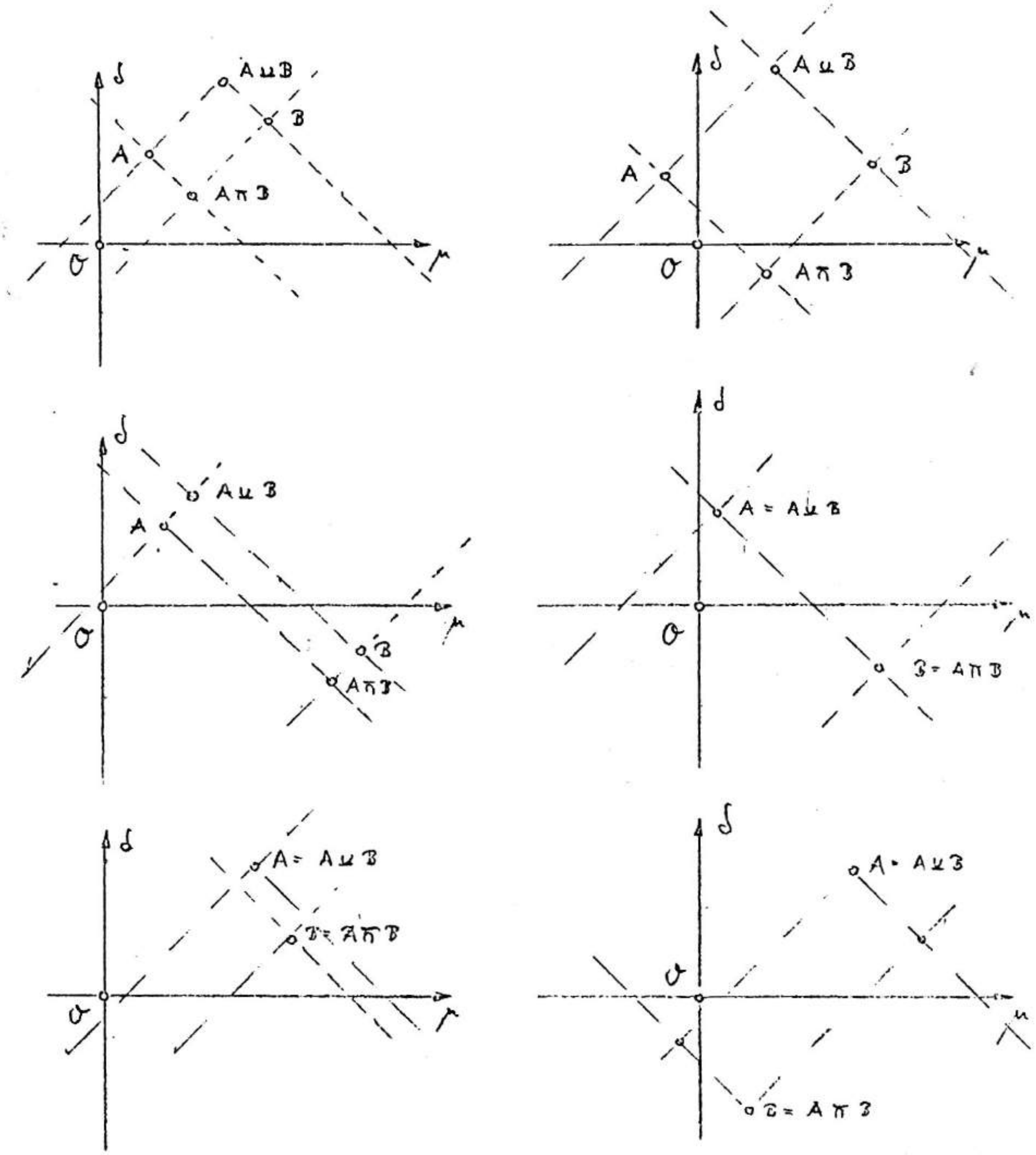


Abb. 2.16

2.7 Verträglichkeitseigenschaften von $\underline{\sqcup}$ und $\overline{\sqcap}$ mit den algebraischen Operationen +, -, $*$, /

Diese Eigenschaften werden im Rahmen dieser Arbeit nicht mehr voll ausgenutzt. Da sie aber eine wichtige Rolle im Zusammenhang von Einschließungsverfahren für analytische Probleme über $\mathbb{IH}$ spielen, sollen sie der Vollständigkeit halber hier aufgeführt werden.

Satz 2.44:

Es gelten folgende Distributivgesetze:

(1) $\bigwedge\limits_{\circ \in \{+,*\}} \; \bigwedge\limits_{A,B,C \in \mathbb{IH}}$ $A\circ(B \,\underline{\sqcup}\, C)=(A\circ B)\,\underline{\sqcup}\,(A\circ C)$ bzw. $A\circ(B \,\overline{\sqcap}\, C)=(A\circ B)\,\overline{\sqcap}\,(A\circ C)$

(2) $\bigwedge\limits_{A,B \in \mathbb{IH}}$ $-(A \,\underline{\sqcup}\, B)=(-A)\,\underline{\sqcup}\,(-B)$ bzw. $-(A \,\overline{\sqcap}\, B)=(-A)\,\overline{\sqcap}\,(-B)$

(3) $\bigwedge\limits_{\substack{A,B \in \mathbb{IH}\setminus\{0\} \\ AB \geqslant 0}}$ $1/(A \,\underline{\sqcup}\, B)=1/A \,\underline{\sqcup}\, 1/B$ bzw, $1/(A \,\overline{\sqcap}\, B)=1/A \,\overline{\sqcap}\, 1/B$

(4) $\bigwedge\limits_{\substack{A,B \in \mathbb{IH}\setminus\{0\} \\ AB \leqslant 0}}$ $1/(A \,\underline{\sqcup}\, B)=1/A \,\overline{\sqcap}\, 1/B$ bzw. $1/(A \,\overline{\sqcap}\, B)=1/A \,\underline{\sqcup}\, 1/B$

Bew.: Da ($\mathbb{IH}, \subseteq$) ein durchweg isoton geordneter Verband ist, könnte man denken, daß Satz 1.11 aus Kapitel I anwendbar wäre. (H, $\leq$) ist jedoch entgegen der dortigen Voraussetzung nicht linear geordnet. Im folgenden Beweisgang wird wesentlich der Satz 1.11 aus Kapitel I verwendet, denn der $\mathbb{IH}$ zugrundegelegte Raum $\mathbb{R}$ ist linear geordneter Verband.

(1) $A+(B \sqcup C)=[\lambda A,\rho A]+[\lambda B\sqcap\lambda C,\rho B\sqcup\rho C]$

$=[\lambda A+(\lambda B\sqcap\lambda C),\rho A+(\rho B\sqcup\rho C)]$ und wegen (OD1) in $(\mathbb{R},\leq)$

folgt $=[(\lambda A+\lambda B)\sqcap(\lambda A+\lambda C),(\rho A+\rho B)\sqcup(\rho A+\rho C)]$

$=(A+B)\sqcup(A+C)$

Zum Beweis des Produktes soll Satz 2.23 herangezogen werden. Dazu muß gezeigt werden, daß sich (E_1), (E_2) und (E_3) mit $\sqcup$ und $\sqcap$ vertragen. Tatsächlich gilt:

$\overline{A * (B \sqcup C)}=\overline{A} * \overline{(B \sqcup C)}=\overline{A} * (\overline{B} \sqcap \overline{C})$ d. h. wegen (E_2) sind $\sqcup$ und $\sqcap$ nicht unabhängig voneinander.

Wegen (2) folgt die Verträglichkeit mit (E_3):

$(-A) * (B \sqcup C)=A * (-(B \sqcup C))=A * ((-B) \sqcup (-C))$

Bei Verwendung von Tabelle 2.9 muß die Distributivität sowohl für $\sqcup$ als auch für $\sqcap$ unabhängig wegen (E_2) nachgewiesen werden.

$A \in S_r$, $B \in S_r$, $C \in S_r$ so ist $B \sqcup C \in S_r$ und es folgt:

$A * (B \sqcup C)=A(B \sqcup C)=[\lambda A(\lambda B\sqcap\lambda C),\rho A(\rho B\sqcup\rho C)]$ und da $\lambda A,\rho A\geq o$ folgt

$=[\lambda A\lambda B\sqcap\lambda A\lambda C,\rho A\rho B\sqcup\rho A\rho C]$

$=AB \sqcup AC=A * B \sqcup A * C$

$A \in S_r$, $B \in S_r$, $C \in T$ so ist $B \sqcup C \in T$ also ist

$A * (B \sqcup C)=\rho(A)(B \sqcup C)$ und da $\rho(A)\geq o$ ist

$=[\rho A\lambda B\sqcap\rho A\lambda C,\rho A\rho B\sqcup\rho A\rho C]$ und da $\lambda C\leq o\leq\lambda B$ $\rho A\geq o$ folgt

$\rho A\lambda B\sqcap\rho A\lambda C=\lambda A\lambda B\sqcap\rho A\lambda C=\rho A\lambda C$ und damit

$=[\lambda A\lambda B\sqcap\rho A\lambda C,\rho A\rho B\sqcup\rho A\rho C]$

$=AB \sqcup \rho(A)C = A * B \sqcup A * C$

$A \in T$, $B \in S_r$, $C \in S_r$ so ist $B \sqcup C \in S_r$ und daher:

$A \ast (B \sqcup C) = A\rho(B \sqcup C = A(\rho B \cup \rho C)$

$= [\lambda A(\rho B \cup \rho C), \rho A(\rho B \cup \rho C)]$ und da $\lambda A \leq o \leq \rho A$ folgt

$= [\lambda A \rho B \sqcap \lambda A \rho C, \rho A \rho B \sqcup \rho A \rho C]$

$= A\rho(B) \sqcup A\rho(C) = A \ast B \sqcup A \ast C$

$A \in T$, $B \in S_r$, $C \in T$ so ist $B \sqcup C \in T$

$A \ast (B \sqcup C) = [\lambda A \rho(B \sqcup C) \sqcap \rho A \lambda(B \sqcup C), \lambda A \lambda(B \sqcup C) \sqcup \rho A \rho(B \sqcup C)]$

$= [\lambda A(\rho B \quad \rho C) \quad \rho A(\lambda B \quad \lambda C), \lambda A(\lambda B \quad \lambda C) \quad \rho A(\rho B \quad \rho C)]$

da $\lambda A \leq o \leq \rho A$ folgt weiter

$= [\lambda A \rho B \sqcap \lambda A \rho C \sqcap \rho A \lambda B \sqcap \rho A \lambda C, \lambda A \lambda B \sqcup \lambda A \lambda C \sqcup \rho A \rho B \sqcup \rho A \rho C]$

und da ferner $\lambda B, \rho B, \lambda C, \rho C \leq o$ gilt stets:

$\lambda A \rho B \sqcap \rho A \lambda B = \lambda A \rho B \wedge \lambda A \rho C \sqcap \rho A \lambda C = \lambda A \rho C \quad \wedge$

$\lambda A \lambda B \sqcup \rho A \rho B = \rho A \rho B \wedge \lambda A \lambda C \sqcup \rho A \rho C = \rho A \rho C$ und damit

$= [\lambda A \rho B \sqcap \lambda A \rho C, \rho A \rho B \sqcup \rho A \rho C]$

$= A\rho(B) \sqcup A\rho(C) = A \ast B \sqcup A \ast C$

Ist dagegen $A \in T$ und $B \sqcup C \in \overline{T}$ so ist trivialerweise $B, C \in \overline{T}$ und deshalb $o = A \ast (B \sqcup C) = A \ast B \sqcup A \ast C = o \sqcup o = o$

Ist $A \in \overline{T}$ und sei o.B.d.A. $B \in H$ und $C \in T$ so ist $B \sqcup C \in T$ und damit $o = A \ast (B \sqcup C) = A \ast B \sqcup A \ast C = X \sqcup o$ mit $X \in \overline{T}$, also ist $X \subsetneq o$ und deshalb $X \sqcup o = o$.

Das war eine Demonstration der wesentlichsten Fälle, denn alle Möglichkeiten wurden hiermit noch nicht erfaßt. Die restlichen Fälle lassen sich aber stets auf ein vorgeführtes Analogon zurückführen.
Entsprechend verlaufen die Beweise für $\sqcap$.

(2) $-[\lambda A \sqcap \lambda B, \rho A \sqcup \rho B] = [-(\rho A \sqcup \rho B),\ -(\lambda A \sqcap \lambda B)]$ und nach Satz 1.11 aus Kapitel I folgt:

$$= [(-\rho A) \sqcap (-\rho B), (-\lambda A) \sqcup (-\lambda B)]$$
$$= [\lambda(-A) \sqcap \lambda(-B), \rho(-A) \sqcup \rho(-B)]$$
$$= (-A) \,\underline{\cup}\, (-B)$$

(3) $A,B \in \mathbb{H} \setminus L \ \wedge\ AB > o \Longleftrightarrow \lambda A \lambda B > o \ \wedge\ \rho A \rho B > o$.

Nun ist $1/(A \,\underline{\cup}\, B) = 1/[\lambda A \sqcap \lambda B, \rho A \sqcup \rho B]$
$$= [1/(\rho A \sqcup \rho B),\ 1/(\lambda A \sqcap \lambda B)]$$

Nach Voraussetzung haben λA und λB wie auch ρA und ρB dasselbe Vorzeichen, so daß nach Satz I,1.16 (d) gilt:

$$= [1/\rho A \sqcap 1/\rho B,\ 1/\lambda A \sqcup 1/\lambda B]$$
$$= 1/A \,\underline{\cup}\, 1/B \ .$$

(4) $A,B \in \mathbb{H} \setminus L \ \wedge\ AB < o \Longleftrightarrow \lambda A \lambda B < o \ \wedge\ \rho A \rho B > o$.

Nun ist $1/(A \,\underline{\cup}\, B) = [1/(\rho A \sqcup \rho B),\ 1/(\lambda A \sqcap \lambda B)]$ und nach Voraussetzung haben λA und λB wie auch ρA und ρB verschiedenes Vorzeichen, so daß nach Satz I,1.11 (d) gilt:

$$= [1/\rho A \sqcup 1/\rho B,\ 1/\lambda A \sqcap 1/\lambda B]$$
$$= 1/A \,\overline{\cap}\, 1/B \ .$$

Daß für $AB \in N = T \cup \overline{T}$ i.a. keine Distributivität gilt, zeigt das Gegenbeispiel:

$1/([2,3] \,\underline{\cup}\, [-5,1]) = 1/[-5,3] = [1/3,-1/5]$ aber

$1/[2,3] \,\underline{\cup}\, 1/[-5,1] = [1/3,1/2] \,\underline{\cup}\, [1,-1/5] = [1/3,1/2]$.

Satz 2.45:

Sei $N = T \cup \overline{T}$, so gilt für beliebige $A,B,C \in H$:

(1) $(A+B) \underline{\cup} C = (A \underline{\cup} (C-\overline{B})) + B$

(2) $(A-B) \underline{\cup} C = (A \underline{\cup} (C+\overline{B})) - B$

(3) $(A-B) \underline{\cup} C = A - (B \underline{\cup} (\overline{A}-C))$

(4) $(A*B) \underline{\cup} C = (A \underline{\cup} (C/\overline{B})) * B \qquad B \notin N$

(5) $(A/B) \underline{\cup} C = (A \underline{\cup} (C\ \overline{B})) \,/\, B \qquad B \notin N$

(6) $(A/B) \underline{\cup} C = A/(B \underline{\cup} (\overline{A}/C))$ sofern $A \notin N \wedge C \notin L \wedge B(\overline{A}/C) > 0$

(7) $(A/B) \underline{\cup} C = A/(B \overline{\cap} (A/C))$ sofern $A \notin N \wedge C \notin L \wedge B(\overline{A}/C) < 0$

Für $\overline{\cap}$ gelten entsprechende Regeln.

Bew.:

(1) Aus Satz 2.44 folgt

$(A \underline{\cup} (C-\overline{B})) + B = A + B \underline{\cup} C - \overline{B} + B = (A + B) \underline{\cup} C$

(2) $(A \underline{\cup} (C+\overline{B})) - B = A - B \underline{\cup} C + \overline{B} - B = (A - B) \underline{\cup} C$

(3) $A - (B \underline{\cup} (\overline{A}-C)) = -((-A) + (B \underline{\cup} (\overline{A}-C))) = -((B-A) \underline{\cup} (\overline{A}-C-A))$

$= -((B-A) \underline{\cup} (-C) = (A - B) \underline{\cup} C$

(4) $(A \underline{\cup} (C/\overline{B})) * B = (A * B) \underline{\cup} (C/\overline{B} * B)$

$= (A * B) \underline{\cup} C$

(5) $(A \underline{\cup} (C * \overline{B}))/B = (A \underline{\cup} (C * \overline{B})) * 1/B = (A/B) \underline{\cup} (C * \overline{B}/B) =$

$= (A/B) \underline{\cup} C$

(6) $A/(B \sqcup \overline{A}/C) = A * 1/(B \sqcup (\overline{A}/C))$ nach Definition 2.29

$= A * (1/B \sqcup 1/(\overline{A}/C))$ nach Satz 2.44(3)

$= A * (1/B \sqcup C/\overline{A})$ nach Satz 2.35(c)

$= A/B \sqcup A * C/\overline{A}$ nach Satz 2.44(1)

$= A/B \sqcup C$ da mit $A \notin N$

$A/\overline{A} = 1$

(7) $A/(B \sqcap \overline{A}/C) = A * 1/(B \sqcap (\overline{A}/C))$

$= A * (1/B \sqcup 1/(\overline{A}/C))$ nach Satz 2.44(4)

$= A/B \sqcup C$

Vermöge der Dualität $A \sqcap B = \overline{\overline{A} \sqcup \overline{B}}$ gelten entsprechende Formeln für $\sqcap$.

Für spezielle Zwecke ist eine besondere Darstellung des Produktes für Intervalle aus N, das nach Tabelle 2.5 offenbar nicht über das hyperbolische Produkt darstellbar ist, von Bedeutung.

Lemma 2.46:

a) $\bigwedge_{A,B \in T}$ $A * B = A\rho(B) \sqcup \overline{A}\lambda(B) = \rho(A)B \sqcup \lambda(A)\overline{B}$

b) $\bigwedge_{A,B \in \overline{T}}$ $A * B = A\lambda(B) \sqcap \overline{A}\rho(B) = \lambda(A)B \sqcap \rho(A)\overline{B}$

Bew.: a) Auf Grund der Identität: $A = \lambda A \sqcup \rho A$ und

$B = \lambda B \sqcup \rho B$

für $A, B \in T$ folgt nach Satz 2.44:

$$A * B = A * (\lambda B \cup \rho B) = A * \lambda B \cup A * \rho B$$

nun ist $\lambda B \leq o$ und $\rho B \geq o$, so daß nach Tabelle 2.5 folgt:

$$A * B = \overline{A}\lambda B \cup A\rho B \text{ .}$$

Auf entsprechende Weise beweisen sich die anderen Fälle.

2.8 Punktweise Berechnung der Intervallverknüpfungen

Wie die widersprüchlichen Beispiele in Abschnitt 2.5 zeigten, darf offenbar die in Definition I, 2.5 angegebene punktweise Berechnung nicht ohne weiteres auf ganz $\mathbb{H}$ übertragen werden. In diesem Abschnitt soll gezeigt werden wie diese Berechnungsweise geeignet modifiziert auf ganz $\mathbb{H}$ erweitert werden kann. Hierzu muß zunächst das Elementsymbol $\in$ auf ganz $\mathbb{H}$ geeignet fortgesetzt werden.

Definition 2.48:

$$\bigwedge_{A \in H} \bigwedge_{a \in \mathbb{R}} a \in' A \Leftrightarrow \begin{cases} a \in A & \text{für} \quad A \in \mathbb{IR} \\ a \in \overline{A} & \text{für} \quad A \in \overline{\mathbb{IR}} \end{cases}$$

$\in'$ soll nur stehen, wenn links von $\in'$ stets ein Element von $\mathbb{R}$ steht. Dieses Symbol deckt sich für $A \in \mathbb{IR}$ mit dem mengentheoretischen Symbol $\in$, hat aber in $\overline{\mathbb{IR}}$ nur formale Bedeutung.

Nach Definition I,2.5 gilt offenbar für $A,B \in \mathbb{IR}$ die Darstellung:

$$A \circ B = \bigcup_{a \in A} \bigcup_{b \in B} (a \circ b) \qquad \text{für} \quad \circ \in \{+,-,*,/\} .$$

Es bietet sich daher an, diese Formel umzuschreiben auf die Verwendung von $\underline{\cup}$ und $\overline{\cap}$, die im Gegensatz zu $\cup$ und $\cap$ in ganz $\mathbb{H}$ Gültigkeit haben. Wie folgende Abschnitte zeigen, ist dies leicht möglich.

<u>Satz 2.49:</u>

Jedes Intervall $A \in \mathbb{H}$ besitzt folgende <u>Punktdarstellung</u>:

$$A = \begin{cases} \underline{\bigcup}_{a \in A} a & \text{für } A \in \mathbb{IR} \\ \overline{\bigcap}_{a \in \overline{A}} & \text{für } A \in \overline{\mathbb{IR}} \end{cases}$$

<u>Bew.:</u> Ist $A \in \mathbb{IR}$, so ist $\underline{\cup} = \cup$ und deckt sich mit der Darstellung in $\mathbb{IR}$.

Ist $A \in \overline{\mathbb{IR}}$, so ist $\overline{A} \in \mathbb{IR}$ und es gilt in $\mathbb{IR}$

$\overline{A} = \underline{\bigcup}_{a \in \overline{A}} a$ und nach Satz 2.42 (1) folgt:

$$A = \overline{\underline{\bigcup}_{a \in \overline{A}} a} = \overline{\bigcap}_{a \in \overline{A}} \overline{a} = \overline{\bigcap}_{a \in \overline{A}} a .$$

Um diese lästige Fallunterscheidung nicht immer schreiben zu müssen, möge folgender ζ-Operator eingeführt werden:

Definition 2.50:

Sei $\phi : \mathbb{R} \to \mathbb{R}$ beliebige reelle Abbildung, so ist der "Zusammensetzoperator" ζ wie folgt definiert:

$$\bigwedge_{\substack{A \in \mathbb{H} \wedge \\ a \in \mathbb{R}}} \zeta(a,A)\phi(a) := \begin{cases} \bigcup\limits_{a \in A} \phi(a) & \text{für } A \in \mathbb{IR} \\ \bigcap\limits_{a \in \overline{A}} \phi(a) & \text{für } A \in \overline{\mathbb{IR}} \end{cases}$$

Ist ϕ stetig, so bildet der ζ-Operator ϕ ab auf eine stetige Intervallfunktion. Speziell gilt nach Satz 2.49:

$A = \zeta(a,A)a$.

Ferner lassen sich nun mit $\sqcap$ und $\sqcup$ bzw. mit Hilfe des ζ-Operators die algebraischen Verknüpfungen in $(\mathbb{H},+,-,*,/)$ in Analogie zu Definition I,2.5 angeben.

Satz 2.51:

(1) Für jede Verknüpfung $\circ \in \{+,*\}$ gilt in $\mathbb{H}$:

$$\bigwedge_{A,B \in \mathbb{H}} A \circ B = \zeta(a,A)\zeta(b,B)(a \circ b)$$

(2) $$\bigwedge_{A \in \mathbb{H}} -A = \zeta(a,A)\,(-a)$$

(3) $$\bigwedge_{A \in \mathbb{H} \setminus N} 1/A = \zeta(a,A)(1/a)$$

eine Sonderrolle spielt natürlich die Inversion in $N \setminus L$:

(4) $$\bigwedge_{-A \in N \setminus L} 1/A = \begin{cases} \overline{\bigcap}_{a \notin A} 1/a & \text{für } A \in \mathbb{IR} \\ \bigcup_{a \notin \overline{A}} 1/a & \text{für } A \in \overline{\mathbb{IR}} \end{cases}$$

Bew.:

(1) Nach Satz 2.49 können für $A \circ B$ nur die vier Fälle auftreten.

$$A \circ B = \begin{cases} \left(\bigcup_{a \in A} a\right) \circ \left(\overline{\bigcap}_{b \in \overline{B}} b\right) & \text{für } A \in \mathbb{IR}, B \in \overline{\mathbb{IR}} \\ \left(\bigcup_{a \in A} a\right) \circ \left(\bigcup_{b \in B} b\right) & \text{für } A \in \mathbb{IR}, B \in \mathbb{IR} \\ \left(\overline{\bigcap}_{a \in \overline{A}} a\right) \circ \left(\overline{\bigcap}_{b \in \overline{B}} b\right) & \text{für } A \in \overline{\mathbb{IR}}, B \in \overline{\mathbb{IR}} \\ \left(\overline{\bigcap}_{a \in \overline{A}} a\right) \circ \left(\bigcup_{b \in B} b\right) & \text{für } A \in \overline{\mathbb{IR}}, B \in \overline{\mathbb{IR}} \end{cases}$$

Durch sukzessive Anwendung der Eigenschaften aus Satz 2.44(1) und auf Grund der Distributivität $\underline{\bigcup}$ und $\overline{\bigcap}$ gewinnt man weiter:

$$A \circ B = \left\{ \begin{array}{l} \bigcup\limits_a \overline{\bigcap\limits_b} (a \circ b) \\[2ex] \bigcup\limits_a \underline{\bigcup\limits_b} (a \circ b) \\[2ex] \overline{\bigcap\limits_a} \overline{\bigcap\limits_b} (a \circ b) \\[2ex] \overline{\bigcap\limits_a} \underline{\bigcup\limits_b} (a \circ b) \end{array} \right\} = \zeta(a,A)\zeta(b,B)(a \circ b)$$

Auf entsprechende Weise gewinnt man (2)(3) und (4) aus den entsprechenden Eigenschaften (2)(3) und (4) des Satzes 2.44.

2.9 Zwei weitere Operationen: Der Links- und Rechtsschnitt

Neben der Infimums- und Supremumsoperation ist es zur Darstellung des Newtonverfahrens nützlich, folgende weitere Verknüpfungen von Intervallen einzuführen:

Definition 2.53:

(1) Linksschnitt $\lhd$:

$$\bigwedge_{A,B \in \mathbb{IH}} A \lhd B := [\lambda A \sqcap \lambda B, \rho A \sqcap \rho B]$$

(2) Rechtsschnitt $\rhd$:

$$\bigwedge_{A,B \in \mathbb{H}} A \rhd B := [\lambda A \sqcup \lambda B, \rho A \sqcup \rho B]$$

Vergl. Abbildung 2.18.

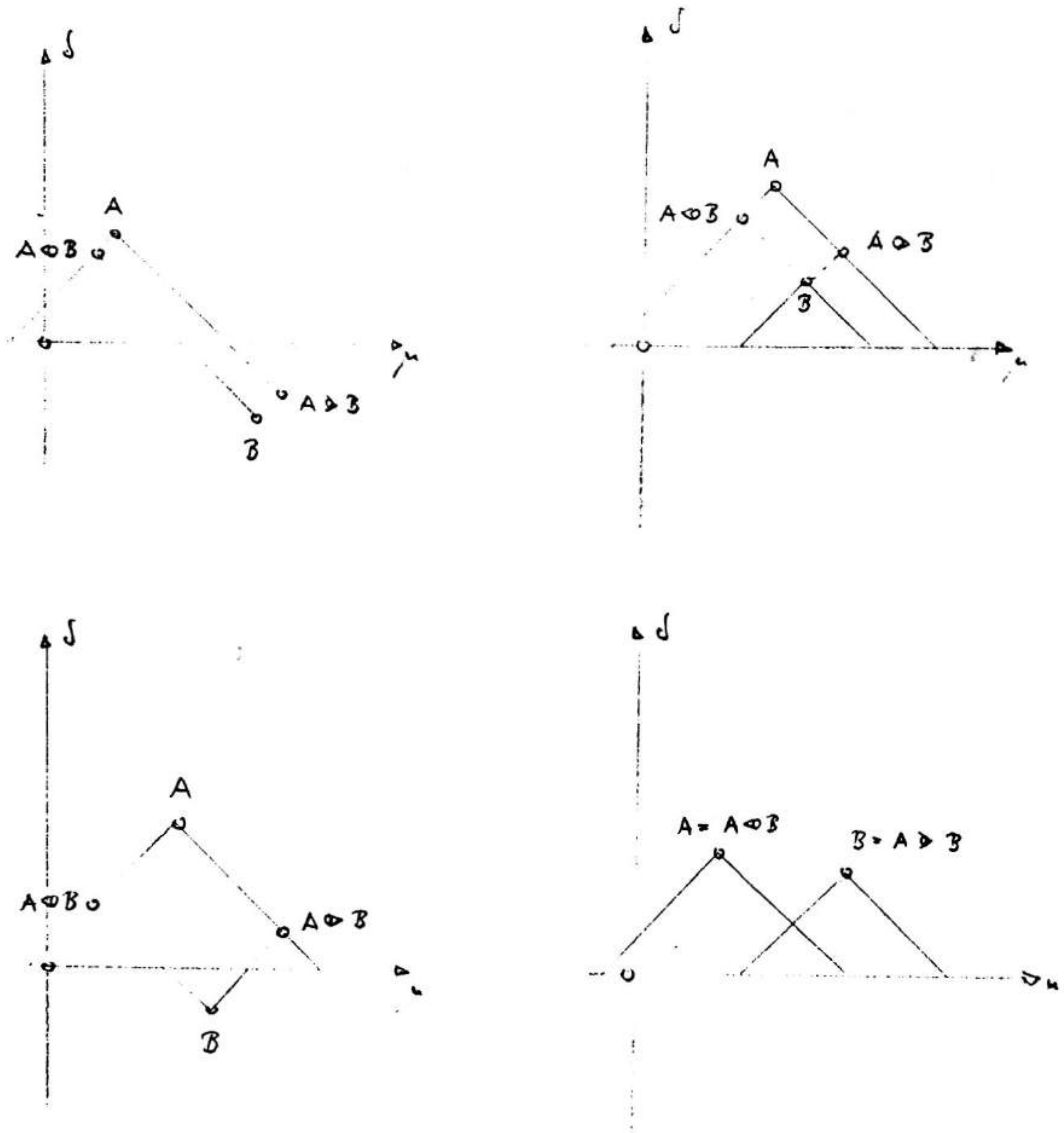


Abb. 2.18

Satz 2.54:

$A,B,C \in \mathbb{H}$ beliebig

(o) $\lambda(A \lessdot B) = \lambda A \sqcap \lambda B, \quad \rho(A \lessdot B) = \rho A \sqcap \rho B$

(1) $A \lessdot A = A \quad , \quad A \lessdot B = B \lessdot A$

(2) $(A \lessdot B) \lessdot C = A \lessdot (B \lessdot C) = (A \lessdot C) \lessdot (B \lessdot C)$

(3) $(A \lessdot B) \gtrdot A = A$

(4) $(A \lessdot B) \gtrdot C = (A \gtrdot C) \lessdot (B \gtrdot C)$

(5) $A \lessdot B \underline{\underline{\ll}} A \quad \wedge \quad A \lessdot B \underline{\underline{\ll}} B$

(6) $A \lessdot B \underline{\underline{\ll}} A \gtrdot B$

Entsprechende Beziehungen gelten auch für $\gtrdot$.

Bew.:

Auf Grund der Eigenschaft (o), die eine unmittelbare Folge der Definition 2.53 darstellt, übertragen sich alle Eigenschaften von $\sqcap$ und $\sqcup$ aus Definition I,1.10 auf $\lessdot$ bzw. $\gtrdot$.

Satz 2.55:

$A,B,C \in \mathbb{H}$ beliebig

(1) (a) $A \in \mathbb{IR} \quad \wedge \quad B \in \overline{\mathbb{IR}} \quad \Rightarrow \quad A \lessdot B \prec A \gtrdot B$

(b) $A \in \mathbb{IR} \;\wedge\; B \in \overline{\mathbb{IR}} \;\wedge\; B \subseteq A \longrightarrow$

$$B \subsetneq A \bowtie B \subsetneq A \;\wedge\; B \subsetneq A \bowtie B \subseteq A$$

(c) $A \bowtie B \in \mathbb{IR} \longrightarrow A \in \mathbb{IR} \;\vee\; B \in \mathbb{IR}$

(2) (a) $(A \bowtie B) \,\bar{\cap}\, C = (A \,\bar{\cap}\, C) \bowtie (B \,\bar{\cap}\, C)$

(b) $(A \,\bar{\cap}\, B) \bowtie C = (A \bowtie C) \,\bar{\cap}\, (B \bowtie C)$

(3) $(A \bowtie B) \,\bar{\cap}\, B = (A \,\bar{\cap}\, B) \bowtie B$

(4) $(A \bowtie B) + C = (A + C) \bowtie (B + C)$

(5) $-(A \bowtie B) = (-A) \bowtie (-B)$

(6) $\overline{(A \bowtie B)} = \overline{A} \bowtie \overline{B}$

(7) $\displaystyle\bigwedge_{\substack{C \in S_r \\ A,B \in \mathbb{IH}}} (A \bowtie B)C = AC \bowtie BC$

(8) $\displaystyle\bigwedge_{\substack{C \in S_r \\ A,B \in H \\ A \supseteq B}} (A \bowtie B) * C = (A * C) \bowtie (B * C)$

(9) $A \subsetneq B \longrightarrow (A \bowtie C) \subsetneq (B \bowtie C)$

(10) $A \underline{\underline{<<}} B \longrightarrow (A \bowtie C) \underline{\underline{<<}} (B \bowtie C)$

(11) $A \prec B \longrightarrow A \,\bar{\cap}\, C \leq (B \,\bar{\cap}\, C) \bowtie C$

Bew.:

(1) a) $A \in \mathbb{IR} \Longrightarrow \lambda A \leq \rho A \quad \wedge \quad B \in \overline{\mathbb{IR}} \Longrightarrow \lambda B \geq \rho B$

Also ist nach Definition 2.30 und Definition 2.53:

$$(\lambda A \sqcap \lambda B) \sqcup (\rho A \sqcap \rho B) \leq (\lambda A \sqcup \lambda B) \sqcap (\rho A \sqcup \rho B)$$

nachzuweisen.

Nach dem Distributivgesetz ist diese Ungleichung äquivalent mit:

$$(\lambda A \sqcup \rho A) \sqcap (\lambda B \sqcup \rho A) \sqcap (\lambda A \sqcup \rho B) \sqcap (\lambda B \sqcup \rho B) \leq$$

$$(\lambda A \sqcap \rho A) \sqcup (\lambda B \sqcap \rho A) \sqcup (\lambda A \sqcap \rho B) \sqcup (\lambda B \sqcap \rho B)$$

Wegen $\lambda A \sqcup \rho A = \lambda A$ ist dies $\Longleftrightarrow$

$$\rho A \sqcap (\lambda B \sqcup \rho A) \sqcap (\lambda A \sqcup \rho B) \sqcap \lambda B \leq \lambda A \sqcup (\lambda B \sqcap \rho A) \sqcup (\lambda A \sqcap \rho B) \sqcup \sqcup \rho B$$

nach dem Absorbtionsgesetz $\Longleftrightarrow$

$$\rho A \sqcap (\lambda A \sqcup \rho B) \sqcap \lambda B \leq \lambda A \sqcup (\lambda B \sqcup \rho A) \sqcup \rho B \Longleftrightarrow$$

$$(\rho A \sqcap \lambda B) \sqcap (\lambda A \sqcup \rho B) \leq (\lambda A \sqcup \rho B) \sqcup (\rho A \sqcap \lambda B)$$

was offenbar wegen $a \sqcap b \leq a \sqcup b$ richtig ist.

(1) (b) o.B.d.A. ist zu zeigen, daß $B \subsetneq A \Leftrightarrow B \subseteq A$ gilt,

d.h.

$$(*) \ \lambda A \leq (\lambda A \sqcap \lambda B) \leq \lambda B \qquad \wedge$$
$$\rho B \leq (\rho A \sqcap \rho B) \leq \rho A$$

Nach Voraussetzung aber ist

$\lambda A \leq \lambda B \quad \wedge \quad \rho B \leq \rho A$, so daß

$\lambda A \sqcap \lambda B = \lambda A \quad \wedge \quad \rho A \sqcap \rho B = \rho B$ folgt,

womit die Bedingung $(*)$ erfüllt ist.

(c) Aus $A \Leftrightarrow B \in \mathbb{IR} \qquad \Longrightarrow$

$\lambda A \sqcap \lambda B \leq \rho A \sqcap \rho B$

es muß $\lambda A \leq \rho A \sqcap \rho B \quad \vee \quad \lambda B \leq \rho A \sqcap \rho B$ sein.

o.B.d.A. folgt aus

$\lambda A \leq \rho A \sqcap \rho B$, daß

$\lambda A \leq \rho A \quad \vee \quad \lambda A \leq \rho B$, d.h.

$A \in \mathbb{IR} \quad \vee \quad (\lambda A \leq \rho B \quad \wedge \quad A \in \overline{\mathbb{IR}})$.

Der Fall $A \in \overline{\mathbb{IR}} \quad \wedge \quad \lambda A \leq \rho B$ führt jedoch zu der Aussage:

$\rho A \leq \lambda A \leq \rho B \qquad \Rightarrow$

$\lambda A \sqcap \lambda B \leq \rho A \qquad \Rightarrow \qquad$ da $\lambda A \geq \rho A$

$\lambda B \leq \rho A \leq \lambda A \leq \rho B$, d.h. $B \in \mathbb{IR}$.

(2)

a) $(A \diamond B)\ \overline{\sqcap}\ C = [(\lambda A \sqcap \lambda B) \sqcup \lambda C, (\rho A \sqcap \rho B) \sqcap \rho C]$

$= [(\lambda A \sqcup \lambda C) \sqcap (\lambda B \sqcap \lambda C), (\rho A \sqcap \rho C) \sqcap (\rho B \sqcap \rho C)]$

$= (A \diamond C)\ \overline{\sqcap}\ (B \diamond C)$

b) $(A\ \overline{\sqcap}\ B) \diamond C = [(\lambda A \sqcup \rho B) \sqcap \lambda C, (\rho A \sqcap \rho B) \sqcap \lambda C]$

$= (A \diamond C)\ \overline{\sqcap}\ (B \diamond C)$ nach derselben

Überlegung wie in (2) a)

(3) $(A \diamond B)\ \overline{\sqcap}\ B = (A\ \overline{\sqcap}\ B) \diamond B$ nach (2) a)

(4) $(A \diamond B) + C = [(\lambda A \sqcap \lambda B) + \lambda C, (\rho A \sqcap \rho B) + \rho C]$

$= [(\lambda A + \lambda C) \sqcap (\lambda B + \lambda C), (\rho A + \rho C) \sqcap (\rho B + \rho C)]$

$= (A + C) \diamond (B + C)$

(5) $-(A \diamond B) = -[\lambda A \sqcap \lambda B, \rho A \sqcap \rho B]$

$= [-\rho A \sqcup -\rho B, -\lambda A \sqcup -\lambda B]$

$= (-A) \diamond (-B)$

(6) $\overline{(A \diamond B)} = \overline{[\lambda A \sqcap \lambda B, \rho A \sqcap \rho B]}$

$= [\rho A \sqcap \rho B, \lambda A \sqcap \lambda B]$

$= [\lambda \overline{A} \sqcap \rho \overline{B}, \rho \overline{A} \sqcap \rho \overline{B}]$

$= \overline{A} \diamond \overline{B}$

(7) $(A \diamond P)C = [(\lambda A \sqcap \lambda B)\lambda C, (\rho A \sqcap \rho B)\rho C] =$

da $\lambda C, \rho C \geq 0$ folgt somit:

$= [\lambda A \lambda C \sqcap \lambda B \lambda C, \rho A \rho C \sqcap \rho B \rho C]$

$= AC \diamond BC$

(8) Es ist

$A \diamond B = [\lambda A \sqcap \lambda B, \rho A \sqcap \rho B]$

wegen $\lambda A \leq \lambda B \wedge \rho B \leq \rho A$

$= [\lambda A, \rho B]$

Somit ist im Falle

a) $A \in S_r \Longrightarrow A \diamond B \in S_r \Longrightarrow$ nach (7)

$(A \diamond B) * C = (A \diamond B)C = AC \diamond BC = A * C \diamond B * C$

b) $A \in S_l \Longrightarrow A \diamond B \in S_l \Longrightarrow$ nach (7)

$(A \diamond B) * C = (A \diamond B)\overline{C} = A\overline{C} \diamond B\overline{C} = A * C \diamond B * C$

c) Mit $A \geq B$ gilt $A * C \geq B * C$, somit ist

$A * C \diamond B * C = [\lambda(A * C), \rho(B * C)]$.

Sei $A \in T$ und

(c.1) $B \in T$, so ist $A \diamond B \in T$ und somit

$$\begin{aligned}
(A \diamond B) * C &= [\lambda A, \rho B] * C \\
&= [\lambda A, \rho B]\rho C \\
&= [\lambda A \rho C, \rho B \rho C] \\
&= [\lambda(A \rho C), \rho(B \rho C)] \\
&= [\lambda(A * C), \rho(B * C)] \\
&= A * C \diamond B * C
\end{aligned}$$

nach obiger Vorüberlegung.

(c.2) $B \in S_r$, so ist $A \infty B \in T$ und somit

$$(A \infty B) * C = \tau$$
$$= [\lambda A\rho C, \rho B\rho C]$$
$$= [\lambda(A * C), \rho(B * C)]$$
$$= A * C \infty B * C$$

(c.3) $B \in S_l$, so ist $A \infty B \in S_l$ und somit

$$(A \infty B) * C = [\lambda A, \rho B]\overline{C}$$
$$= [\lambda A\rho C, \rho B\lambda C]$$
$$= [\lambda(A * C), \rho(B\overline{C})]$$
$$= [\lambda(A * C), \rho(B * C)]$$
$$= A * C \infty B * C \qquad \text{q.e.d.}$$

(9) $A \subseteq B \Longrightarrow$

$\lambda B \leq \lambda A \wedge \rho A \leq \rho B \Longrightarrow$

$\lambda B \sqcap \lambda C \leq \lambda A \sqcap \lambda C \wedge \rho A \sqcap \rho C \leq \rho B \sqcap \rho C \Longrightarrow$

$(A \infty C) \subseteq (B \infty C)$

(10) $A \ll B \Longrightarrow$

$\lambda A \leq \lambda B \wedge \rho A \leq \rho B \Longrightarrow$

$\lambda A \sqcap \lambda C \leq \lambda B \sqcap \lambda C \wedge \rho A \sqcap \rho C \leq \rho B \sqcap \rho C$

$(A \infty C) \ll (B \infty C)$

(11) Trivialerweise ist $\lambda C \leq \lambda A \sqcup \lambda C \Longrightarrow$

$\lambda C = (\lambda B \sqcup \lambda C) \sqcap C\lambda \leq \lambda A \sqcup \lambda C$

Ferner folgt mit $A \in B$, daß $\rho A \leq \rho B$, $\Longrightarrow$

$\rho A \sqcap \rho C \leq \rho B \sqcap \rho C = (\rho B \sqcap \rho C) \sqcap \rho C$,

somit ist

$A \,\overline{\pi}\, C \subseteq (B \,\overline{\pi}\, C) \infty C$.

Mit Hilfe dieser neu eingeführten Operationen läßt sich der bei der Inversion in 2.5 auftretende Komplex $M=\{1/a \mid a \in A \in \mathbb{IR}\}$ darstellen.

Satz 2.56:

Es gilt in $\mathbb{H}^* := \mathbb{IR}^* \cup \overline{\mathbb{IR}^*}$ mit $P = [-p,p]$

für festes $A \in T \subseteq \mathbb{IR}$:

$$M = \{1/a \mid a \in A\} = (1/A \mathrel{<\!\circ} P) \cup (1/A \mathrel{\circ\!>} P)$$

Bew.:

$$\begin{aligned} \text{Es ist } M &= \{1/a \mid a \in A\} = \\ &= \{1/a \mid a \in [\lambda A, o] \quad \vee \quad a \in [o, \rho A]\} = \\ &= \{1/a \mid a \in [\lambda A, o]\} \cup \{1/a \mid a \in [o, \rho A]\} = \\ &= [-p, 1/\lambda A] \cup [1/\rho A, p] \quad \text{nach Definition 2.29.} \end{aligned}$$

Nach Definition II,4.1 ist $-p$ kleinstes bzw. p größtes Element in $\mathbb{R}^*$, so daß weiter folgt:

$$\begin{aligned} M &= [-p \sqcap 1/\rho A, p \sqcap 1/\lambda A] \cup [-p \sqcup 1/\rho A, p \sqcup 1/\lambda A] \\ &= (1/A \mathrel{<\!\circ} P) \cup (1/A \mathrel{\circ\!>} P). \end{aligned}$$

Dies ist eine wichtige Beziehung zur Darstellung des modifizierten Newton-Verfahrens. (Es ist zu beachten, daß damit nach Definition 2.29 $M = \{1/a \mid a \in [o,\alpha]\}$ das Intervall $[1/\alpha, p]$ zugeordnet ist.)

3. Das Vektoid $V_n\mathbb{H}$

Alle Begriffe und Formeln, die in $\mathbb{H}$ gelten, lassen sich im wesentlichen durch komponentenweise Übertragung auf das n-dimensionale Vektoid $V_n\mathbb{H}$ anwenden.

Definition 3.1:

$(V_n\mathbb{H}, +, \cdot, *)$ ist das (in Anlehnung an [Ullrich 72]) definierte Vektoid über dem Ringoid $(\mathbb{H}, +, *)$ mit den ausgezeichneten Elementen $\{-1, o, 1\}$. Entsprechend ist $(M_n\mathbb{H}, +, *)$ die Menge der verallgemeinerten Intervallmatrizen.

Es ist $A \in V_n\mathbb{H}$ mit $A = (A_1, \ldots, A_n) = (A_i)$

und $A_i \in \mathbb{H}$ für $i = 1, \ldots, n$.

Es ist $\mathfrak{A} \in M_n\mathbb{H}$ mit $\mathfrak{A} = (A_{11}, \ldots, A_{1n}, \ldots, A_{2n}) = (A_{ij})$

und $A_{ij} \in \mathbb{H}$ für $i,j = 1, \ldots, n$.

Die Verknüpfungen sind wie folgt definiert:

$$\bigwedge_{A,B \in V_n\mathbb{H}} \quad A + B := (A_i + B_i)$$

$$\bigwedge_{\substack{\lambda \in \mathbb{H} \\ A \in V_n\mathbb{H}}} \quad \lambda \cdot A := (\lambda \cdot A_i)$$

$$\bigwedge_{\substack{\mathfrak{A} \in M_n\mathbb{H} \\ B \in V_n\mathbb{H}}} \mathfrak{A} * B := \Big(\sum_{j=1}^{n} (A_{ij} * B_j)\Big)$$

$$\bigwedge_{\mathfrak{A},\, \mathfrak{B} \in M_n\mathbb{H}} \mathfrak{A} + \mathfrak{B} := \big((A_{ij}\ B_{ij})\big)$$

$$\bigwedge_{\mathfrak{A},\, \mathfrak{B} \in M_n\mathbb{H}} \mathfrak{A} * \mathfrak{B} := \Big(\Big(\sum_{k=1}^{n} (A_{ik} * B_{kj})\Big)\Big)$$

Darüber hinaus sind die wesentlichsten Operationen aus $\mathbb{H}$ auf $V_n\mathbb{H}$ bzw. $M_n\mathbb{H}$ wie folgt fortgesetzt:

(1) $-A := (-A_i)$ bzw. $-\mathfrak{A} = ((-A_{ij}))$

(2) $\overline{A} := (\overline{A}_i)$ bzw. $\overline{\mathfrak{A}} = ((\overline{A}_{ij}))$

(3) $A \sqcup B := (A_i \sqcup B_i)$ bzw. $\mathfrak{A} \sqcup \mathfrak{B} := ((A_{ij} \sqcup B_{ij}))$
(entsprechend für $\sqcap$)

(4) $A \leq B \Longleftrightarrow \bigwedge_{i=1,\dots,n} A_i \subseteq B_i$ bzw. $\mathfrak{A} \leq \mathfrak{B} \Longleftrightarrow \bigwedge_{i,j=1,\dots n} A_{ij} \subseteq B_{ij}$

(5) (In Vorgriff auf Abschnitt 4.):

$\|A\| := (\|A_i\|)$ bzw. $\|\mathfrak{A}\| := ((\|A_{ij}\|))$

(6) $d(A) := (d(A_i))$

Damit gelten alle Eigenschaften des Vektoides nach [Ullrich 72]:

(V1)(V2)(VD1)(VD2)(VD3) und (VD4) .

Darüber hinaus überträgt sich die Existenz der additiven Inversen von $\mathbb{H}$ auf $V_n\mathbb{H}$ und $M_n\mathbb{H}$ und die Existenz von additiven und multiplikativen Inversen auf $M_n\mathbb{H}$.

4. Norm, Prämetrik und Metrik in $\mathbb{H}$

In [Herzberger 69], [Alefeld 68], [Mayer 68] und insbesondere in [Moore 69] sind verschiedene Metriken und die dazugehörigen Konvergenzbegriffe für Intervallfolgen in $\mathbb{IR}$ eingeführt worden. Im erweiterten Intervallraum $\mathbb{H}$ soll eine dieser Metriken beibehalten und auf ganz $\mathbb{H}$ erweitert werden. Dabei stellt es sich heraus, daß $\mathbb{H}$ zu einem normierten Vektorraum ausgebaut werden kann.

4.1 $\mathbb{H}$ als normierter Vektorraum

Nach [Mayer 68] ist $\mathbb{IR}$ ein quasilinearer Raum, in der dann die Norm rückwärts über eine homogene supermetrische Metrik $q(A,B)$ durch Setzung $||A||:=q(A,o)$ definiert ist. Mayer zeigte in der Dissertation [Mayer 68], daß in quasilinearen Räumen aus einer vorgegebenen Norm $p(A)$ durch Differenzbildung (falls sie überhaupt definiert ist) keine Metrik $q(A,B):=p(A-B)$ zu gewinnen ist.

Die in dieser Arbeit durchgeführte Vervollständigung von $\mathbb{IR}$ zu $\mathbb{H}$ liefert einen Vektorraum $\mathbb{H}$, wenn man als Skalarraum nicht die reellen Zahlen sondern den hyperbolischen Ring $(\mathbb{H},+,\ )$ nimmt.

Satz 4.1:

Der erweiterte Intervallraum $(\mathbb{H},+)$ ist über dem hyperbolischen Ring $(\mathbb{H},+,\ )$ ein Vektorraum.

Bew.:

1. Nach Satz 2.9 ist der erweiterte Intervallraum $(\mathbb{H},+)$ eine kommutative Gruppe mit dem Nullelement $[o,o]=o$ und $(\mathbb{H},+,\ )$ ist ein Ring mit 1-Element $[1,1]$.

2. Sind $\alpha,\beta,A,B \in \mathbb{H}$ und sei $\alpha\beta$ das hyperbolische Produkt von α und β, so gilt nach Satz 2.13 und vermöge der Identität der Addition in $(\mathbb{H},+)$ und $(\mathbb{H},+,\ )$:

$$\alpha(A+B) = \alpha A + \alpha B$$

$$(\alpha+\beta)A = \alpha A + \beta A$$

$$\alpha\beta A = \alpha(\beta A)$$

$$1A = A$$

$$oA = o$$

Somit ist $(\mathbb{H},+)$ ein 1-dimensionaler Vektorraum über den Ring $(\mathbb{H},+,\ )$.

Bemerkung:

Man könnte auf diese Hilfskonstruktion verzichten, denn nach Abschnitt 4.2 erhält man ohnehin eine Metrik. Im Hinblick jedoch auf die Tatsache, daß $\ast$ nach Tabelle 2.5 im hyperbolischen Ring darstellbar ist, könnte es beweistechnisch von Vorteil sein, den erwähnten Zusammenhang zu kennen.

Definition und Satz 4.3:

Die in [Moore 69] und [Mayer 68] bereits verwendete Norm läßt sich auf ganz $\mathbb{H}$ fortsetzen. Es ist

$$\bigwedge_{A \in \mathbb{H}} \|A\| := |\lambda A| \sqcup |\rho A| = \max(|\lambda A|, |\rho A|)$$ eine Norm und $(\mathbb{H}, \| \cdot \|)$ ein normierter Vektorraum .

Bew.:

$\mathbb{H}$ ist ein Vektorraum unter Berücksichtigung der oben gemachten Bemerkung. Ferner gilt:

(a) $\|A\| = o \Longleftrightarrow |\lambda A| = |\rho A| = o \Longleftrightarrow A = o$

(b) $$\begin{aligned} \|\alpha A\| &= \| [\lambda(\alpha)\lambda(A), \rho(\alpha)\rho(A)] \| = \\ &= |\lambda(\alpha)|\,|\lambda(A)| \sqcup |\rho(\alpha)|\,|\rho(A)| \leq \\ &\leq (|\lambda(\alpha)| \sqcup |\rho(\alpha)|)\cdot(|\lambda(A)| \sqcup |\rho(A)|) = \|\alpha\| \cdot \|A\| \end{aligned}$$

(c) $$\begin{aligned} \|A+B\| &= |\lambda A + \lambda B| \sqcup |\rho A + \rho B| \leq \\ &\leq |\lambda A| + |\lambda B| \sqcup |\rho A| + |\rho B| \leq \quad \text{nach Satz 1.16 (e) in I,} \\ &\leq (|\lambda A| \sqcup |\rho A|) + (|\lambda B| \sqcup |\rho B|) = \|A\| + \|B\| \end{aligned}$$

Lemma 4.4:

Es gelten ferner folgende Eigenschaften:

(d) $\|\overline{A}\| = \|A\|$

(e) $\|A \veebar B\| \leq \|A\| \sqcup \|B\|$

$\|A \sqcap B\| \leq \|A\| \sqcup \|B\|$

(f) $A \subsetneq B \subsetneq C \Longrightarrow \| B \| \leq \| A \| \sqcup \| C \|$

$\overline{B} \subsetneq A \subsetneq B \quad \wedge \quad B \in \mathbb{IR} \Longrightarrow \| A \| \leq \| B \|$

(g) $\| A \| = \bigsqcup_{a \in' A} |a|$

(h) $\| A * B \| \leq \| A \| \quad \| B \|$

$\| \frac{1}{A} \| = \frac{\| A \|}{A * \overline{A}} = \frac{\| A \|}{\lambda(A)\rho(A)}$ für alle $A \in \mathbb{H} \setminus N$

Bew.:

(d) $|\rho(\overline{A})| \sqcup |\lambda(\overline{A})| = |\lambda(A)| \sqcup |\rho(A)| = \| A \|$

(e) $\| A \underline{\cup} B \| = \| [\lambda A \sqcap \lambda B, \rho A \sqcup \rho B] \| =$

$= |\lambda A \sqcap \lambda B| \sqcup |\rho A \sqcup \rho B| \leq$

$\leq |\lambda A| \sqcup |\lambda B| \sqcup |\rho A| \sqcup |\rho B| = \| A \| \sqcup \| B \|$

Entsprechend folgt:

$\| A \overline{\cap} B \| = \| \overline{A} \underline{\cup} \overline{B} \| \leq \| \overline{A} \| \sqcup \| \overline{B} \| = \| A \| \sqcup \| B \|$

(f) Aus $A \subsetneq B \subsetneq C \Longrightarrow$

$\lambda C \leq \lambda B \leq \lambda A \quad \wedge \quad \rho A \leq \rho B \leq \rho C \Longrightarrow$

$|\lambda B| \leq |\lambda A| \sqcup |\lambda C| \quad \wedge \quad |\rho B| \leq |\rho A| \sqcup |\rho C| \Longrightarrow$

$\| B \| = |\lambda B| \sqcup |\rho B| \leq |\lambda A| \sqcup |\lambda C| \sqcup |\rho A| \sqcup |\rho C| =$

$= \| A \| \sqcup \| B \|$.

Speziell mit $A := \overline{B}$, $B := A$ und $C := B$ folgt die zweite Behauptung.

(h) Aus Satz 2.26 folgt unmittelbar:

$$\left\| \frac{1}{X} \right\| = \left\| \frac{X}{X * \overline{X}} \right\| = \left\| \frac{1}{X * \overline{X}} * X \right\| = \frac{1}{X * \overline{X}} \| X \| \quad \text{da} \quad X * \overline{X} \in \mathbb{R}^+$$

Die zweite Aussage dieses Punktes ergibt sich wie folgt:

Fall 1: Ist mindestens einer der Faktoren A oder B $\notin N = T \cup \overline{T}$, so ist nach Tabelle 2.5 $A * B = \phi\psi$ mit ϕ und $\psi \in \mathbb{H}$, und nach Satz 4.3 (b) ist $\|\phi\psi\| \leq \|\phi\| \, \|\psi\|$. Nach Verknüpfungstabelle 2.5 aber ist ϕ stets gleich einer der Größen: $\lambda A, \rho A$ oder $\overline{A}$. Für die Norm gilt jedoch stets die Abschätzung:
$\|\phi\| \leq \|A\|$. Entsprechend gilt für ψ :
$\|\psi\| \leq \|B\|$. Somit folgt $\|A * B\| \leq \|\psi\| \, \|\psi\| \leq \|A\| \cdot \|B\|$.

Fall 2: Sind $A, B \in T$, so ist nach Lemma 2.44.a:

$$\begin{aligned} \|A * B\| &= \|A\rho(B) \sqcup \overline{A}\lambda(B)\| \leq \quad \text{nach (e)} \\ &\leq \|A\rho(B)\| \sqcup \|\overline{A}\lambda(B)\| \leq \quad \text{nach Satz 4.3 (b)} \\ &\leq \|\cdot A\| \; |\rho(B)| \sqcup \|A\| \cdot |\rho(B)| = \\ &= \|A\| \, (|\rho(B)| \sqcup |\lambda(B)|) = \|A\| \cdot \|B\| \end{aligned}$$

Sind $A, B \in \overline{T}$, so ist $\|A * B\| = \|\overline{\overline{A} * \overline{B}}\| \leq \|A\| \cdot \|B\|$ nach obigem Fall.

Sind o.B.d.A. $A \in T$ und $B \in \overline{T}$, so ist nach Tabelle 2.5:
$\|A * B\| = o \leq \|A\| \cdot \|B\|$.

Bemerkung 4.5:

Die Isonormen in $\mathbb{H}$, d.h. $\{X \mid \|X\| = \text{constans}\}$, sind in $\mathcal{O}$ konzentrische Quadrate mit den Eckpunkten auf den μ-δ-Achsen.

Siehe Abbildung 4.1.

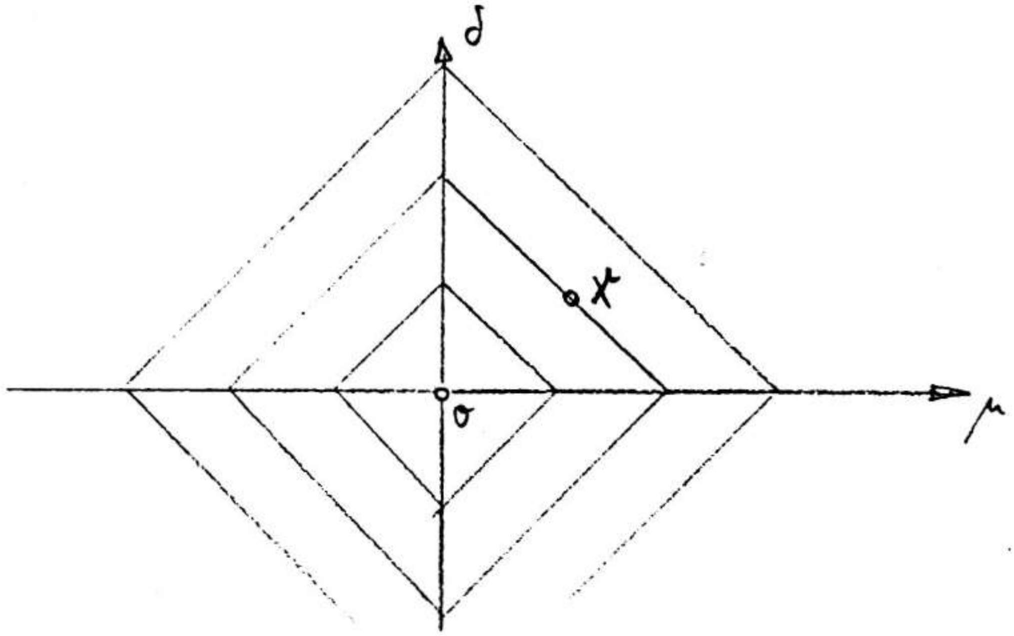


Abb. 4.1

4.2 Eine Prämetrik und Metrik in $\mathbb{H}$ und $V_n\mathbb{H}$

Wie Kapitel II.2 zeigt, ist für Fehlerabschätzungen beim gerundeten Rechnen das Aufstellen eines prämetrischen Feldes notwendig. Nach der Zusammenfassung in Satz II,2.11 bedarf es dazu zunächst einer Prämetrik in $\mathbb{H}$.

Definition und Satz 4.6:

Die Abbildung $p\colon \mathbb{H} \times \mathbb{H} \to \mathbb{H}$ mit

$$\bigwedge_{X,Y \in \mathbb{H}} p(X,Y) := X - \overline{Y}$$ ist eine Prämetrik, die die Eigenschaften

(P1)(P2)(P3)(P4)(P5)(P6)(P7) und (P8) der Definition II.1.3 erfüllt.

Bew.:

(P1) $X - \overline{Y} = o \iff X = Y$

(P2) $X - \overline{Y} + Y - \overline{X} = (X-\overline{X}) + (Y-\overline{Y}) = o + o = o$

(P3) $X - \overline{Z} = (X-\overline{Y}) + (Y-\overline{Z})$

(P4) $(\mathbb{H},\subseteq)$ ist geordnet und es gilt:
$X \subseteq Y \Longrightarrow X - \overline{Z} \subseteq Y - \overline{Z}$ nach Satz 2.29, Eigenschaft (K2).

(P5)
$$\begin{aligned} A + X - \overline{(A+Y)} &= A + X - \overline{A} - \overline{Y} = \\ &= A - \overline{A} + X - \overline{Y} \\ &= X - \overline{Y} \end{aligned}$$

(P6) $A + X - \overline{(B+Y)} = (A-\overline{B}) + (X-\overline{Y})$

(P7) Nach Forderung der Definition II,1.3 muß $\mathbb{H}$ Skalarraum von $\mathbb{H}$ sein, wobei der Zielraum der Prämetrik ein linearer Raum sein soll. Dies ist nach Definition und Satz 4.3 mit dem hyperbolischen Produkt möglich. (P7) braucht daher nur für das hyperbolische Produkt gelten.

$$AX - \overline{AY} = AX -_h AY = A(X -_h Y)$$

$$= A(X - \overline{Y}) \quad \text{(nach Satz 2.15)}$$

(P8) $1 - o = o - (-1) = 1$

Definition 4.7:

Die Prämetrik $P_n : V_n\mathbb{H} \times V_n\mathbb{H} \to V_n\mathbb{H}$ sei komponentenweise definiert durch die Prämetrik $p : \mathbb{H} \times \mathbb{H} \to \mathbb{H}$ mit

$$\mathfrak{A} = (A_1, \ldots, A_n) \quad \mathfrak{L} = (B_1, \ldots, B_n)$$

$$p_n(\mathfrak{A}, \mathfrak{L}) := (p(A_1, B_1), \ldots, p(A_n, B_n))$$

d. h. es gilt

$$p_n(\mathfrak{A}, \mathfrak{L}) = \mathfrak{A} - \overline{\mathfrak{L}} = (A_1 - \overline{B}_1, \ldots, A_n - \overline{B}_n)$$

Definition und Satz 4.9:

Die in Definition 4.3 definierte Norm erzeugt in $(\mathbb{H}, \leq)$ gemäß Satz II,1.4 mit

$$\bigwedge_{A,B \in \mathbb{H}} \acute{q}(A,B) := \|p(A,B)\| = \|A - \overline{B}\|$$ eine

Metrik.

Sie erfüllt (M1)(M2)(M3)(M4)(M5)(M6) und (M8).

(M7) ist ebenfalls bezüglich des hyperbolischen Produktes gültig. Darüber hinaus gilt

$$(M7') \quad q(A \ast B, A \ast C) \leq q(A,o)\, q(B,C) = \|A\|\, q(B,C)$$

und

$$(M7'') \quad q(A/B, A/C) \leq \frac{\|A\|}{\|B \ast C\|}\, q(B,C)$$

<u>Bew.:</u> Alle Aussagen ergeben sich als Folgerung aus Satz II.1.4.

Zum Beweis von (M7') müßten auf Grund der Sektoraufteilung des Produktes 64 Fälle untersucht werden. Zur Vereinfachung des Beweises dient folgendes Lemma:

<u>Lemma 4.10:</u>

Zu zwei beliebigen Punkten $X,Y \in \mathbb{H}$ existiert eine Menge $R(X,Y) \neq \emptyset$ mit der Eigenschaft:

$$\bigwedge_{Z \in R(X,Y)} \quad \|X - \overline{Z}\| + \|Z - \overline{Y}\| = \|X - \overline{Y}\|$$

$R(X,Y)$ ist ein zu den δ-μ-Achsen achsenparalleles von X und Y aufgespanntes Rechteck.

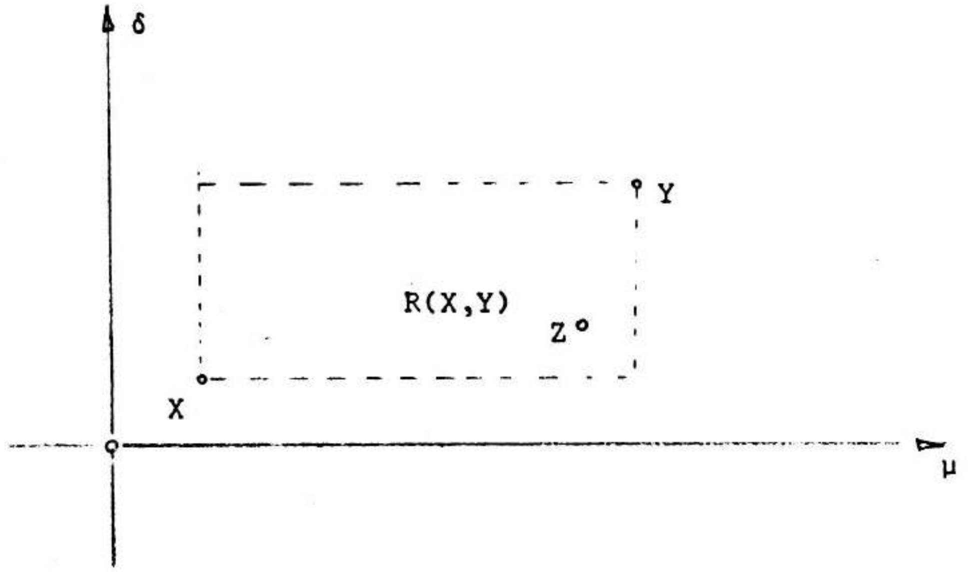


Abb. 4.2

Bew.:

Betrachtet man die nach Bemerkung 4.5 definierten Isonormen von $\|X - \bar{Z}\| = c$ mit Mittelpunkt X und von $\|Z - \bar{Y}\| = \tilde{c}$ mit Mittelpunkt Y, so folgt aus der Bedingung

$$\| X - \bar{Z}\| + \|Z - \bar{Y}\| = c + \tilde{c} = r = \|X - \bar{Y}\| ,$$

daß mit $\tilde{c} = r - c$ ein zum Parameterwert $c \in \mathbb{R}^+$ gehöriger Punkt Z(c) gleichzeitig auf der c-Isonorme um X und auf der (r-c)-Isonorme um Y liegen muß. Da jedoch die Isonormen Quadrate sind, die konzentrisch um X bzw. Y liegen mit ihren Eckpunkten auf den μ-δ-Achsenparallelen durch X und Y , so kann Z(c) nur auf und im Inneren des Rechteckes liegen, das durch diese μ-δ-Achsenparallelen durch X und Y aufgespannt wird. Siehe Abbildung 4.3.

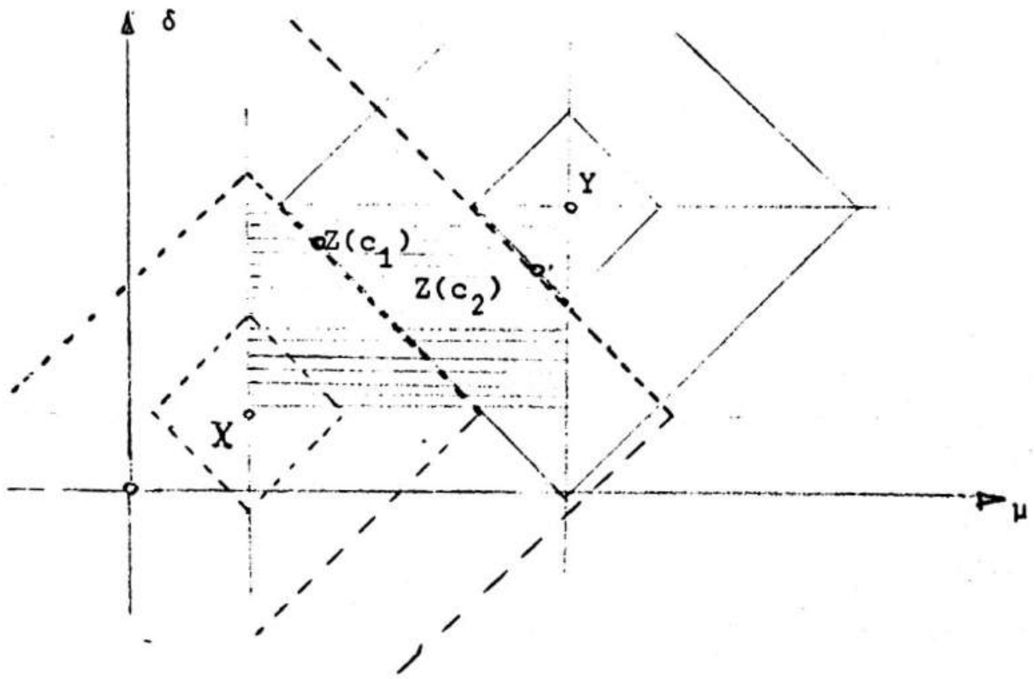


Abb. 4.3

Mit Hilfe der folgenden nun bekannten Eigenschaften und Tatsachen

1) In $\mathbb{IR}$ gilt (M7') [Mayer 68]

2) $\| -A \| = \| A \|$

3) $\| \overline{A} \| = \| A \|$

4) Lemma 4.10

lassen sich die 64 Fälle auf folgende noch zu beweisende 3 Fälle reduzieren:

(I) $A \in S_r \quad \wedge \quad X,Y \in T$

(II) $A \in S_r \quad \wedge \quad X,Y \in S_r$

(III) $A \in T \quad \wedge \quad X,Y \in \overline{T}$

Zum Beweis der drei Fälle:

$$\text{(I)}\quad \|A * X - \overline{A * Y}\| = \|\rho(A)X -_h \rho(A)Y\|$$

$$= \|\rho(A)(X -_h Y)\|$$

$$\overset{\text{Satz 4.3(b)}}{=} \|\rho(A)\| \; \|X - \overline{Y}\|$$

$$\leq \|A\| \; \|X - \overline{Y}\|$$

$$\text{(II)}\quad \|A * X - \overline{A * Y}\| = \|AX -_h AY\|$$

$$= \|A(X -_h Y)\|$$

$$\overset{\text{Satz 4.3(b)}}{=} \|A\| \; \|X - \overline{Y}\|$$

$$\text{(III)}\quad \|A * X - \overline{A * Y}\| = 0 \leq \|A\| \; \|X - \overline{Y}\|$$

Mit 1), 2), 3), 4), (I)(II) und (III) lassen sich die restlichen Fälle leicht beweisen.

Folgende ausgewählte Beispiele mögen dies zeigen:

Sei $A \in S_r \wedge X,Y \in S_1$, so ist

$$\| A * X - \overline{A * Y}\| \overset{2)}{=} \|A * (- X) - \overline{A * (- Y)}\|$$

mit $(- X), (- Y) \in S_r$ und nach (II) folgt somit:

$$\leq \|A\| \; \|(-X) - \overline{(-Y)}\| \overset{2)}{=}$$
$$= \|A\| \; \|X - \overline{Y}\|$$

Sei $A \in S_r$, $X,Y \in \overline{T}$, so ist

$$\| A * X - \overline{A * Y} \| \overset{3)}{=} \| \overline{A} * \overline{X} - \overline{\overline{A} * \overline{Y}} \|$$

mit $\overline{A} \in S_r \wedge \overline{X},\overline{Y} \in T$ und nach (I) somit:

$$\leq \|\overline{A}\| \; \|\overline{X} - \overline{\overline{Y}}\| \overset{3)}{=}$$

$$= \|A\| \; \|X - \overline{Y}\|$$

Sei $A \in S_r \wedge X \in S_r \wedge Y \in S_1$, so folgt

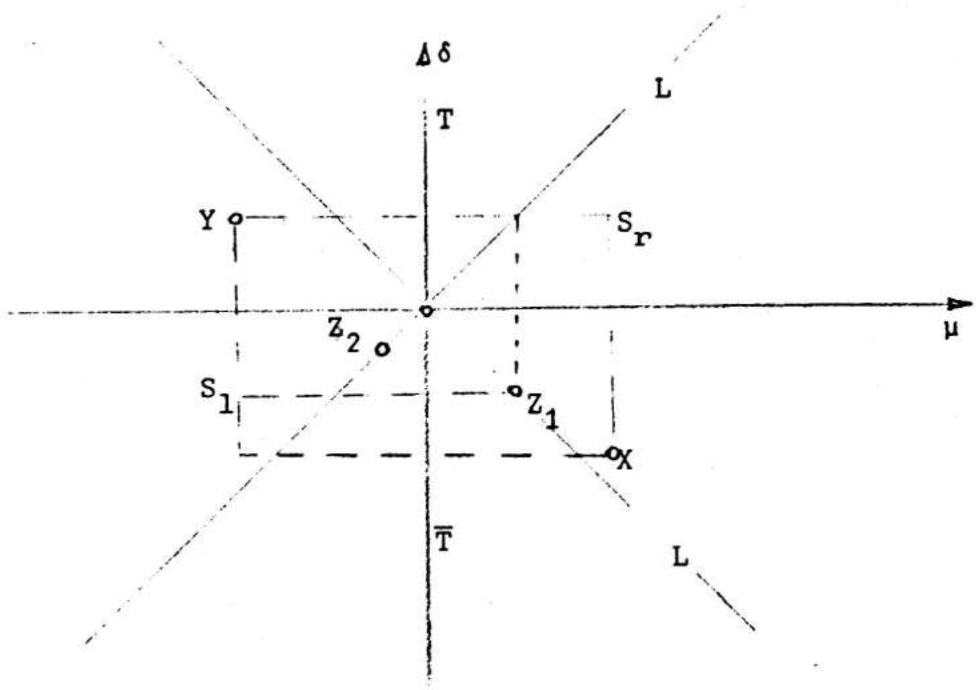


Abb. 4.4

mit Hilfe der Dreiecksungleichung mit geeigneten Z_1, $Z_2 \in R(X,Y)$ und $Z_1 \in \overline{T} \cap S_r$ und $Z_2 \in \overline{T} \cap S_1$ und $Z_2 \in R(Z_1,Y)$:

$$\begin{aligned}\|A * X - \overline{A * Y}\| \leq\ & \|A * X - \overline{A * Z_1}\| + \\ & \|A * Z_1 - \overline{A * Z_2}\| + \\ & \|A * Z_2 - \overline{A * Y}\|\end{aligned}$$

Die drei nun vorliegenden Summanden betreffen nach Konstruktion die Fälle:

$A \in S_r \quad \wedge \quad X, Z_1 \in S_r$

$A \in S_r \quad \wedge \quad Z_1, Z_2 \in \overline{T}$

$A \in S_r \quad \wedge \quad Z_2, Y \in S_1$

für die oben bereits nachgewiesen wurde, daß (M7') gilt.

Damit folgt weiter:

$$\begin{aligned}\|A * X - \overline{A * Y}\| \leq\ & \|A\| \quad \|X - \overline{Z_1}\| + \\ & \|A\| \quad \|Z_1 - \overline{Z_2}\| + \\ & \|A\| \quad \|Z_2 - \overline{Y}\| = \\ =\ & \|A\| \quad (\| X - Z_1\| + \|Z_1 - \overline{Z_2}\| + \|Z_2 - \overline{Y}\|) =\end{aligned}$$

Nach Konstruktion der Z_1 und Z_2 gilt Lemma 4.10 und somit ist

$$\begin{aligned}& = \|A\| \quad (\|X - \overline{Z_2}\| + \|Z_2 - \overline{Y}\|) = \\ & = \|A\| \quad \|X - \overline{Y}\| .\end{aligned}$$

$$\begin{aligned}\text{(M7'')}\, q(A/B, A/C) & = \|A/B - \overline{A/C}\| = \\ & = \|A * 1/B - \overline{A * 1/C}\| = \text{nach (M7')} \\ & \leq \|A\| \cdot \|1/B - \overline{1/C}\| =\end{aligned}$$

und da $B, C \in \mathbb{H} \setminus N$ gilt:

$$= \|A\| \quad \|\overline{C}/(B * C) - \overline{B/(B * C)}\| =$$

$$= \|A\| \quad \|\overline{C} * 1/(B * C) - \overline{\overline{B} * 1/(B * C)}\| \quad \leq$$

nach (M7')

$$\leq \frac{\|A\|}{\|B * C\|} \| \quad \overline{C} - B\|$$

Entsprechend ist mit $\mathfrak{A} = (A_1, \ldots, A_n)$ und $\mathfrak{B} = (B_1, \ldots, B_n)$

$$q(\mathfrak{A},\mathfrak{B}) = (q(A_1,B_1), \ldots, q(A_n,B_n))$$

$$= (\|A_1-\overline{B}_1\|, \ldots, \|A_n - \overline{B}_n\|)$$

eine Pseudometrik in V_nH .

5. Das modifizierte Newton-Verfahren zur Bestimmung der reellen Nullstellen einer reellen Funktion in einem vorgegebenen Intervall X_o .

In [Alefeld 68], [Krawczyk 68] und [Barth 71] ist das modifizierte Newton-Verfahren zur Einschließung der einfachen Nullstellen erschöpfend beschrieben. In diesen Darstellungen und Algorithmen ist jedoch unbefriedigend, daß sie im Falle mehrer einfacher Nullstellen in einem vorgegebenen Intervall X_o , d. h. bei nichtmonotoner Ableitung der Funktion, keine geschlossene Darstellung zulassen.

Mit Hilfe der in Abschnitt 2 definierten Inversion für Intervalle $A \in N \setminus \{o\}$, gelangt man ohne Schwierigkeiten zu einer automatischen Abspaltung weiterer nullstellenverdächtiger Intervalle.

Folgende Sätze sind Grundlage des Satzes 5.4 und geben an, wie die Nullstellenmenge $\nu(L_1,X_o)$ einer quasilinearen Kurve 1. Ordnung vom Grade 1

$$L_1(\tau) = A + B * (\tau - c) : \mathbb{R} \to \mathbb{H}$$

zu berechnen ist.

<u>Definition 5.1:</u>

Die <u>Nullstellenmenge</u> einer stetigen Funktion $F: \mathrm{I\!R} \ni X_o \to \mathrm{I\!H}$ mit der Erzeugenden f und Parameterbereich $\mathfrak{A} \in V_n \mathrm{I\!H}$, d. h. $F(\tau) := \{f(\alpha, \tau) \mid \alpha \in \mathfrak{A}\}$ ist:

$\nu(F, X_o) := \{x \in X_o \mid f(\alpha, x) = o \quad \wedge \quad \alpha \in' \mathfrak{A}\}$.

Im folgenden wird nur die Nullstellenmenge der quasilinearen Kurve 1. Ordnung vom Grade 1 interessieren. Die Überlegung ist die, eine reelle Funktion $f(x)$ lokal in einem vorgegebenen Intervall $X_o \in \mathrm{I\!I\!R}$ durch eine solche quasilineare Kurve einzuschließen, d. h. zu f existieren zwei Intervalle $A, B \in \mathrm{I\!I\!R}$, so daß

$$\bigwedge_{x \in X_o} f(x) \in A + B * (x - c) = L_1(x)$$ gilt. Siehe Abbildung 5.1.

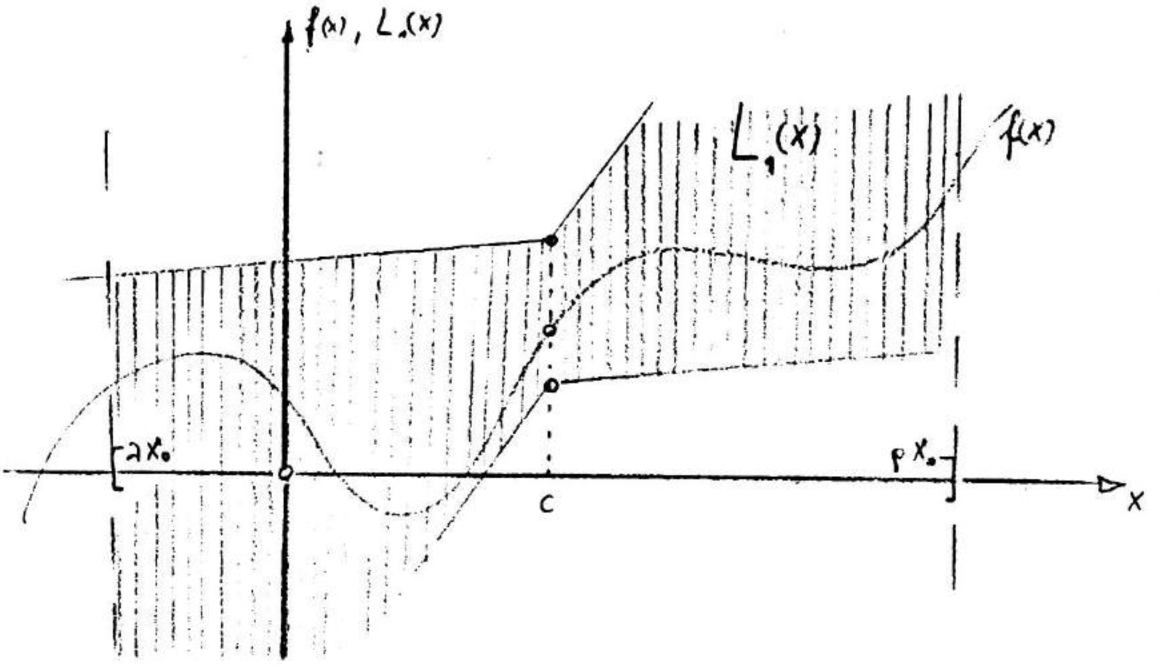


Abb. 5.1

Die Nullstellen von f in X_o werden dann durch die Nullstellenmenge $\nu(L_1,X_o)$ eingeschlossen, denn wie man sieht, gilt speziell:

$$\nu(f,X_o) = \{x \in X_o | f(x)=o\} \subseteq$$

$$\subseteq \{x \in X_o | a+b(x-c)=o \quad \wedge \quad a \in A \quad \wedge \quad b \in B \quad \wedge$$

$$f(x) \in A + B * (x-c)\} .$$

A,B werden sich im einfachsten Falle für $f \in C^1_{Xo}$ aus dem Mittelwertsatz berechnen. Tatsächlich ist von diesem allgemeinen Standpunkt aus gesehen nichts über A und B vorausgesetzt, als daß eben $f \in A + B * (x-c)$ gilt.

Satz 5.2:

Die Nullstellenmenge $\nu(L_1,X_o)$ einer quasilinearen Kurve 1. Ordnung vom Grade 1

$$L_1(\tau) = A + B * (\tau - c) \quad \text{mit} \quad A,B \in \mathbb{IR}, B \not\ni o$$

berechnet sich zu:

Fall 1: Für $B \in \mathbb{IR} \setminus T$ (Abb. 5.2)

$$\nu(L_1(X_o)) := \begin{cases} (c-A/B) \sqcap X_o & \text{falls} \quad (c-A/B) \sqcap X_o \in \mathbb{IR} \\ \\ \emptyset & \text{sonst} \end{cases}$$

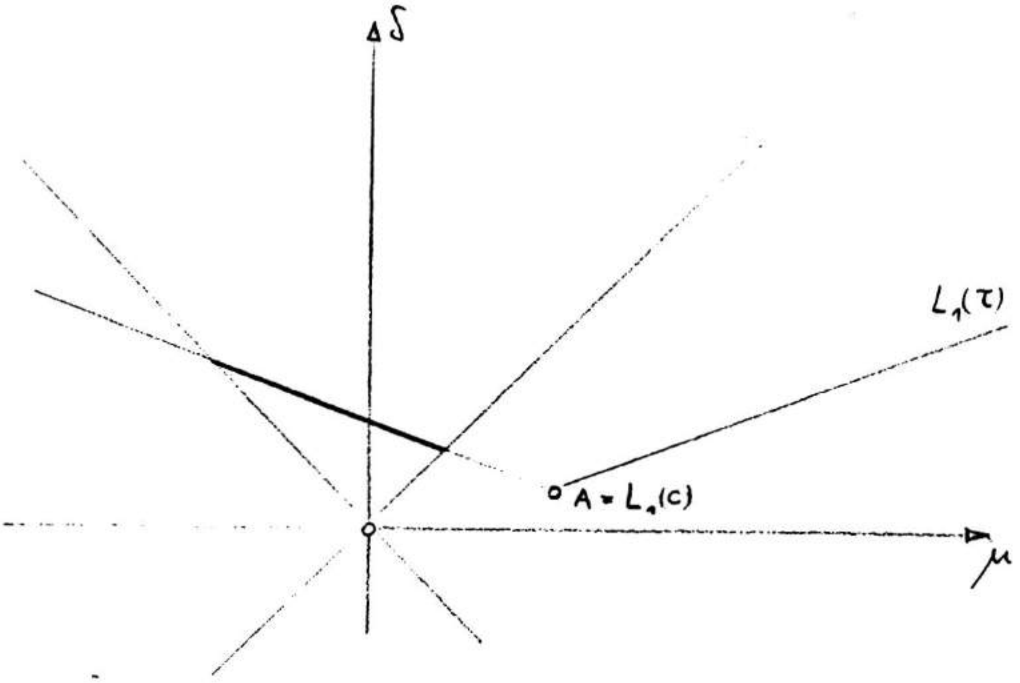


Abb. 5.2

<u>Fall 2:</u> Für $B \in T \setminus \{o\}$ ist mit

$$M_1 := ((c-A/B) \,\overline{\cap}\, X_o) \lhd X_o =$$

$$((c-A/B) \lhd X_o) \,\overline{\cap}\, X_o \qquad \text{und}$$

$$M_2 := ((c-A/B) \,\overline{\cap}\, X_o] \rhd X_o =$$

$$((c-A/B) \rhd X_o) \,\overline{\cap}\, X_o \qquad \text{und}$$

$$H_1 := \begin{cases} M_1 & \text{für} \quad M_1 \in \mathbb{IR} \\ \emptyset & \text{für} \quad M_1 \in \overline{\mathbb{IR}} \setminus \mathbb{R} \end{cases}$$

$$H_2 := \begin{cases} M_2 & \text{für} \quad M_2 \in \mathbb{IR} \\ \emptyset & \text{für} \quad M_2 \in \overline{\mathbb{IR}} \setminus \mathbb{R} \end{cases}$$

$\nu(L_1, X_o) = H_1 \cup H_2$ (Abbildung 5.3)

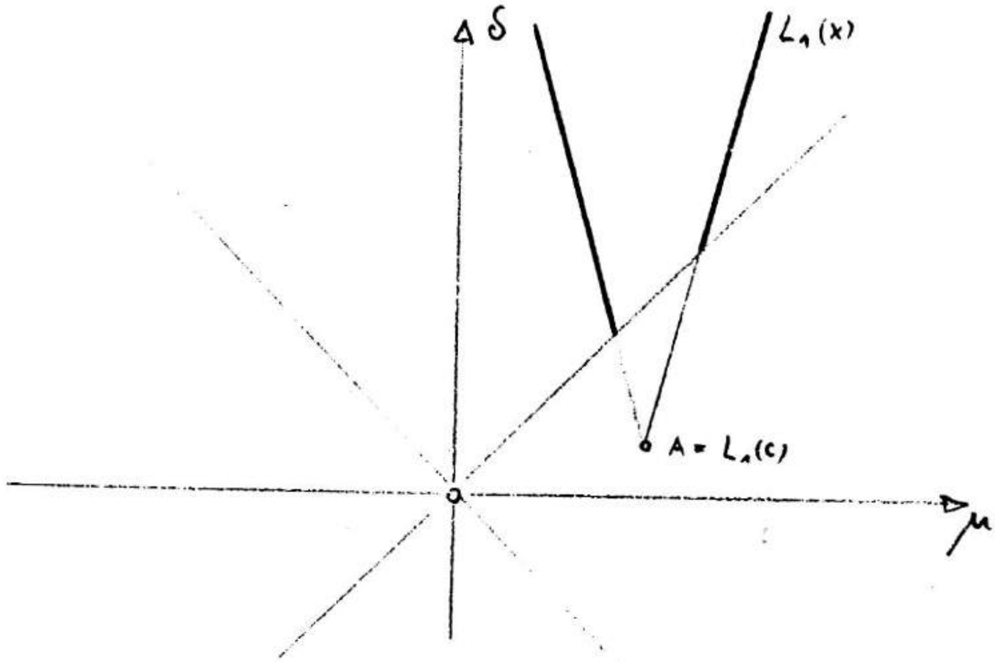


Abb. 5.3

<u>Bew.:</u>

<u>Fall 1:</u>

$$\nu(L_1, X_o) = \{x \in X_o \mid a + b(x-c) = o \quad \wedge \quad a \in A \quad \wedge \quad b \in B\} =$$

$$= \{x \in X_o \mid x = c - a/b \quad \wedge \quad a \in A \quad \wedge \quad b \in B\} =$$

$$= (c-A/B) \cap X_o$$ nach Satz 2.51 bzw. Definition I,2.5.

Nach Satz 2.40 und 2.42 ist

$(c-A/B) \cap X_o = \emptyset \iff (c-A/B) \sqcap X_o \in \overline{\mathbb{IR}} \setminus R$ und

$(c-A/B) \cap X_o \neq \emptyset \iff (c-A/B) \sqcap X_o \in \mathbb{IR}$, da

$A, B, X_o \in \mathbb{IR}$.

Somit ist

$$\nu(L_1, X_o) = \begin{cases} (c-A/B)\,\bar{\sqcap}\,X_o & \text{für } (c-A/B)\,\bar{\sqcap}\,X_o \in \mathbb{IR} \\ \\ \emptyset & \text{sonst.} \end{cases}$$

<u>Fall 2:</u>

$$\begin{aligned} \nu(L_1, X_o) &= \{x \in X_o \mid a+b(x-c) = o \;\wedge\; a \in A \;\wedge\; b \in B\} = \\ &= \{x \in X_o \mid x = c - a/b \;\wedge\; a \in A \;\wedge\; b \in B\} = \\ &= \{x \in \mathbb{R}^* \mid x = c - a/b \;\wedge\; a \in A \;\wedge\; b \in B\} \cap X_o = \\ &= \{x \in \mathbb{R} \mid x = c - ay \;\wedge\; a \in A \;\wedge\; y \in Y\} \cap X_o \end{aligned}$$

mit $Y := \{1/b \mid b \in B\} = (1/B \lhd P) \cup (1/B \rhd P)$

nach Satz 2.56.

Somit ist

$$\begin{aligned} \nu(L_1, X_o) = (&\{x \in \mathbb{R}^* \mid x = c-ay \;\wedge\; a \in A \;\wedge\; y \in (1/B \lhd P)\} \cup \\ &\{x \in \mathbb{R}^* \mid x = c-ay \;\wedge\; a \in A \;\wedge\; y \in (1/B \rhd P)\}) \cap X_o \end{aligned}$$

Nun ist $1/B \subseteq P$ und somit nach Satz 2.55(1)b) sowohl $(1/B \lhd P) \in \mathbb{IR}$ als auch $(1/B \rhd P) \in \mathbb{IR}$. Damit folgt nach Satz 2.51 bzw. Definition I,2.5 in $\mathbb{IR}$:

$$\begin{aligned} \nu(L_1, X_o) = \;&(c-A \ast (1/B \lhd P) \cap X_o \quad \cup \\ &(c-A \ast (1/B \rhd P) \cap X_o \end{aligned}$$

Nach Satz 2.55(4)(5) und (8) folgt weiter:

$$\nu(L_1,X_o) = (c-(A/B \lhd A * P)) \cap X_o \quad \cup$$

$$(c-(A/B \rhd A * P)) \cap X_o \quad =$$

$$= ((c-A/B) \rhd (c-P)) \cap X_o$$

$$((c-A/B) \lhd (c-P)) \cap X_o \quad =$$

$$= ((c-A/B) \rhd P) \cap X_o \cup ((c-A/B) \lhd P) \cap X_o$$

$$= \quad G_2 \qquad\qquad \cup \qquad\qquad G_1$$

Nun treten die Fälle auf:

Ist $G_1 \neq \emptyset$ bzw. $G_2 \neq \emptyset$, so ist nach Satz 2.41 und 2.42 mit

$G_1 \in \mathbb{IR}$ stets $G_1 = ((c-A/B) \lhd P) \tilde{\sqcap} X_o$ bzw. mit

$G_2 \in \mathbb{IR}$ stets $G_2 = ((c-A/B) \rhd P) \tilde{\sqcap} X_o$

und nach Satz 2.55(2)

$$G_1 = ((c-A/B) \tilde{\sqcap} X_o) \lhd (P \sqcap X_o) = ((c-A/B) \tilde{\sqcap} X_o) \lhd X_o$$

und nach Satz 2.55(3)

$$= ((c-A/B) \lhd X_o) \tilde{\sqcap} X_o = M_1$$

bzw.

$$G_2 = ((c-A/B) \tilde{\sqcap} X_o) \rhd X_o = ((c-A/B \rhd X_o) \tilde{\sqcap} X_o = M_2 \text{ .}$$

Ist jedoch $G_1 = \emptyset$ bzw. $G_2 = \emptyset$, so ist dies nach Satz 2.41 genau dann der Fall, wenn $M_1 \in \overline{\mathbb{IR}} \setminus \mathbb{R}$ bzw. $M_2 \in \overline{\mathbb{IR}} \setminus R$.

Es ist bemerkenswert, daß schon die einfache Kurve, die quasilineare Kurve 1. Ordnung vom Grade 1 (für $B \in T$) unter geeigneten Umständen zwei Nullstellenintervalle besitzt. Entsprechend ist zu vermuten, daß eine quasilineare Kurve n-ter Ordnung vom Grade 1

$$L_n(\tau) = A + \sum_{i=o}^{n} B_i * (\tau - c_i)$$

n+1 Nullintervalle besitzen kann. Dies legt die Vermutung nahe, daß das in Satz 5.4 beschriebene Newton-Verfahren von sehr spezieller Art ist. Man könnte sich vorstellen, daß man $f(x)$ durch $L_n(x)$ einschließt und die Nullstellenmenge von L_n berechnet.

Für numerische Zwecke u.a. ist es wichtig zu wissen, ob die in Satz 5.2 aufgeführten Formeln inklusionsisoton sind. Offenbar gilt für $\tau \in X_o$ mit $L_1(\tau) \subseteq L_2(\tau)$, daß $\nu(L_1,X_o) \subseteq \nu(L_2,X_o)$ ist. Im Fall 1 des Satzes 5.2 ist dies trivial, da ja die Inklusionsisotonie in $\mathbb{IR} \setminus T$ nach Satz 1.2 auch für die Division gilt.

Anders liegt die Sache im Fall 2, da dort die Inversion in $T \setminus \{o\}$ auftritt und nach Satz 2.33 i.a. nicht mehr inklusionsisoton ist.

Dennoch gilt der

<u>Satz 5.3:</u>

Sei $L_1(\tau) := A + B * (\tau - c)$ und

$\tilde{L}_1(\tau) := \tilde{A} + \tilde{B} * (\tau - c)$ mit $A \subseteq \tilde{A} \;\wedge\; B \subseteq \tilde{B}$,

$A, \tilde{A}, B, \tilde{B} \in \mathbb{IR}$,

d.h.

$$\bigwedge_{\tau \in \mathbb{R}^*} L_1(\tau) \subseteq \tilde{L}_1(\tau) ,$$

so läßt sich mit den in Satz 5.2 angegebenen Formeln formal zeigen, daß

$$\nu(L_1, X_o) \subseteq \nu(\tilde{L}_1, X_o) \qquad \text{gilt .}$$

<u>Bew.:</u> Nach den Fallunterscheidungen von Satz 2.34 folgt im einzelnen:

(1) $B, \tilde{B} \in \mathbb{IR} \setminus T$, so gilt nach Inklusionsisotonie von $-$, $/$ und $\sqcap$:

$$\nu(L_1, X_o) = (c - A/B) \sqcap X_o \subseteq (c - \tilde{A}/\tilde{B}) \sqcap X_o = \nu(\tilde{L}_1, X_o)$$

(2) $B \in \mathbb{IR} \setminus T \;\wedge\; \tilde{B} \in T \setminus \{o\} \;\wedge\; A \in T$, so ist

$\nu(L_1, X_o) = (c - A/B) \sqcap X_o$ und

$\nu(\tilde{L}_1, X_o) = X_o$, da $\tilde{A}/\tilde{B} = P$ nach Satz 2.29 und es ist trivialerweise

$$\nu(L_1, X_o) \subseteq X_o = \nu(\tilde{L}_1, X_o)$$

(3) $B \in \mathbb{IR} \setminus T \quad \wedge \quad \widetilde{B} \in T \setminus \{o\} \quad \wedge \quad A \in \mathbb{IR} \setminus T$,

so ist

$$\nu(L_1, X_o) = (c - A/B) \,\overline{\cap}\, X_o \quad \wedge \quad \nu(\widetilde{L}_1, X_o) = \widetilde{H}_1 \cup \widetilde{H}_2$$

nach Fall 2 von Satz 5.2.

Nun ist nach Satz 2.34 c)(3) je nach Fall mit $B \in S_r$ bzw. $B \in S_l$

$A/B \succ \widetilde{A}/\widetilde{B}$ bzw. $A/B \prec \widetilde{A}/\widetilde{B}$, also ist nach Satz 2.55(11)

$A/B \,\overline{\cap}\, X_o \subseteq (\widetilde{A}/\widetilde{B} \,\overline{\cap}\, X_o) \vartriangleright X_o$ bzw.

$A/B \,\overline{\cap}\, X_o \subseteq (\widetilde{A}/\widetilde{B} \,\overline{\cap}\, X_o) \vartriangleleft X_o$, d.h. aber gerade, daß

$\nu(L_1, X_o) \subseteq \widetilde{H}_1$ bzw. $\nu(L_1, X_o) \subseteq \widetilde{H}_2$, d. h.

$\nu(L_1, X_o) \subseteq \widetilde{H}_1 \cup \widetilde{H}_2 = \nu(\widetilde{L}_1, X_o)$

(4) $B \in T \setminus \{o\} \quad \wedge \quad \widetilde{B} \in T \setminus \{o\}$, so ist nach Satz 2.33 die Division / inklusionsisoton und somit gilt wie in (1) $\nu(L_1, X_o) \subseteq \nu(\widetilde{L}_1, X_o)$.

Nach diesen Vorbetrachtungen kann nun endlich das modifizierte Newton-Verfahren dargestellt werden. (Im Hinblick auf die Anwendung ist das Verfahren rekursiv definiert.)

Definition und Satz 6.4:

Ist $f \in C^1_{X_o} : X_o \to \mathbb{R}$, f nicht konstant, und besitzen f und f' in $X_o \in \mathbb{IR}$ höchstens endlich viele Nullstellen, so berechnen sich die vorhandenen Nullstellen in X_o nach folgendem modifiziertem Newton-Verfahren:

1. Newton 1(f,X) bezeichne das modifizierte Newton-Verfahren für den Fall $f'(X) \notin T$ (d.h. f monoton) und berechnet die Nullstelle $\hat{x} \in X$, falls sie existiert, nach folgender Vorschrift:

(1.1) $\hat{x} = \lim\limits_{i \to \infty} X_i$, mit $X_{i+1} := \left(m(X_i) - \dfrac{f(m(X_i))}{f'(X_i)}\right) \sqcap X_i$

für $i \geq o$, $X_o = X \wedge X_i \in \mathbb{IR}$.

(1.2) Ist für ein $i > o$ $X_i \in \overline{\mathbb{IR}} \setminus \mathbb{IR}$, so liegt keine Nullstelle in X_o.

2. Newton 2(f,X) bezeichne das modifizierte Newton-Verfahren für den allgemeinen Fall $f'(X) \in \mathbb{IR}$. Es berechnet alle Nullstellen $\hat{x} \in X$, falls überhaupt eine existiert, nach folgender Vorschrift:

(2.1) Für $f'(X) \notin T$ ist nach Newton 1(f,X) zu verfahren.

(2.2) Für $f'(X) \in T \setminus \{o\}$ ist mit

$$H_1 := \left(\left(m(X) - \frac{f(m(X))}{f'(X)}\right) \ominus X\right) \sqcap X \quad \text{und}$$

$$H_2 := \{(m(X) - \frac{f(m(X))}{f'(X)}) \diamond X\} \sqcap X$$

Newton $2(f, H_1)$ falls $H_1 \in I\!I\!R$ und

Newton $2(f, H_2)$ falls $H_2 \in I\!I\!R$ zu

berechnen.

(2.3) Ist $f'(X)=o \wedge f(m)=o \wedge X=m(X)$, so ist

$X = m(X)$ Nullstelle.

(2.4) In allen anderen Fällen liegt keine Nullstelle in X .

3. Die Nullstellen (falls sie existieren) in X_o berechnen sich mit Newton $2(f, X_o)$. Die Berechnung bricht nach endlich vielen Schritten ab, wenn keine Nullstellen in X_o liegen.

<u>Bew.:</u> Zum Beweis dieses Satzes zunächst ein Lemma, das wichtige Eigenschaften zusammenstellt:

<u>Lemma 5.5:</u>

a) $$\hat{x} \in X \wedge f(\hat{x})=o \Longrightarrow \begin{cases} \hat{x} \in (m(X) - \frac{f(m(X))}{f'(X)}) \sqcap X & \text{falls } f'(X) \notin T \\ \\ \hat{x} \in H_1 \cup H_2 & \text{falls } f'(X) \in T \setminus \{o\} \end{cases}$$

b) Für $X \subseteq X_o$ gilt:

$$\left(m(X) - \frac{f(m(X))}{f'(X)}\right) \cap X \begin{cases} \subsetneqq X & \text{für } f'(X) \notin T \\ \\ = X & \text{für } X = \hat{x} = m(X) \in \mathbb{R} \\ & \text{und } f(\hat{x}) = o \;. \end{cases}$$

$$H_1 \subsetneqq X \quad \wedge \quad H_2 \subsetneqq X \quad \text{für } f'(X) \in T \setminus \{o\} \;.$$

<u>Beweis des Lemmas:</u>

Es sei $m = m(X)$

a) Nach Konstruktion ist ein Schritt im modifizierten Newton-Verfahren, nämlich die Berechnung von

$$X \longrightarrow \begin{cases} \left(m - \frac{f(m)}{f'(X)}\right) \cap X & \text{für} \quad f'(X) \notin T \\ \\ H_1 \cup H_2 & \text{für} \quad f'(X) \in T \setminus \{o\} \end{cases}$$

nach Satz 5.2 gerade die Berechnung der Nullstellenmenge $\nu(L_1, X)$ mit $L_1(\tau) = f(m) + f'(X) * (\tau - m)$. Nun gilt für $f \in C^1_{X_o}$ der Mittelwertsatz, d.h. es ist für $\tau, m \in X \subseteq X_o$:

$$f(\tau) = f(m) + f'(\xi)(\tau - m) \qquad \text{für geeignetes} \quad \xi \in X,$$

also ist $f'(\xi) \subseteq f'(X)$ und somit folgt nach dem Einschließungssatz

$$f(\tau) \subseteq f(m) + f'(X) * (\tau - m) = L_1(\tau) \quad \text{für} \quad \tau \in X_o .$$

Somit gilt für jede Nullstelle $\hat{x} \in X$:

$f(\hat{x}) = o \subseteq L_1(\hat{x})$ und nach Definiton 6.1

$\hat{x} \in \nu(L_1, X)$.

b) (b.1) Ist $f(m) = o$, d.h. $m = \hat{x}$ eine Nullstelle, so ist (insbesondere auch im Falle (2.3))

$$m - \frac{f(m)}{f'(X)} = m = \hat{x} \subseteq X$$

und es ist nach (2.3) $m = \hat{x} = X \in \mathbb{R}$

oder $m = \hat{x} \subset X \in \mathbb{IR} \setminus \mathbb{R}$.

(b.2) Ist $f(m) \neq o$ und ist $f'(X) \notin T$, so ist $\frac{f(m)}{f'(X)} \notin T$, d. h. es ist $\frac{f(m)}{f'(X)} \ll o$ bzw. $\frac{f(m)}{f'(X)} \gg o$. Somit ist nach Lemma 2.32 e):

$$m - \frac{f(m)}{f'(X)} \gg m \quad \text{bzw.} \quad m - \frac{f(m)}{f'(X)} \ll m .$$

o.B.d.A. gilt für

$$m - \frac{f(m)}{f'(X)} \ll m \underline{\underline{\leq}} \rho X:$$

$$\left(m - \frac{f(m)}{f'(X)}\right) \sqcap X = \left[\lambda\left(m - \frac{f(m)}{f'(X)}\right) \sqcup \lambda X,\ \rho\left(m - \frac{f(m)}{f'(X)}\right)\right]$$

$$\subset \left[\lambda\left(m - \frac{f(m)}{f'(X)}\right) \sqcup \lambda X,\ m\right] \subseteq$$

$$\subseteq \quad [\lambda X, \rho X] = X$$

(b.3) Ist $f'(X) \in T \setminus \{o\} \ \wedge\ f(m) \neq o$, so ist

$\frac{f(m)}{f'(X)} \in \overline{T} \setminus \{o\}$, d.h. $\frac{f(m)}{f'(X)} \subsetneqq 0$ und

aufgrund der Inklusionsisotonie gilt

$m - \frac{f(m)}{f'(X)} \subsetneqq m \subseteq X$.

Nach Satz 2.54(9) folgt schließlich

$\left(m - \frac{f(m)}{f'(X)}\right) \sqcap X \subsetneqq (X \sqcap X) = X$, d. h.

$H_1 \subsetneqq X$.

Auf analoge Weise erhält man $H_2 \subsetneqq X$.

Nun zum Beweis von Satz 5.4:

Zu (2.1) und (2.2) des Satzes:

Aus dem Lemma folgt, daß für eine Nullstelle $\hat{x} \in X_i$ stets $\hat{x} \in X_{i+1} \subset X_i$ gilt, solange $\hat{x} \neq X_i$ ist und nur im Falle $\hat{x} = X_i$ ist $X_{i+1} = X_i = \hat{x} \in \mathbb{R}$ nach Fall b.1 des Lemmas. Es geht also ausgehend von X_o keine Nullstelle $\hat{x} \in X_o$ verloren.

Zu (2.4) bzw. (3) des Satzes:
Diejenigen Intervalle X, in denen keine Nullstellen liegen, führen aufgrund der Aussage b) des Lemmas nach endlich vielen Schritten zu einem $X_i \subset X$ mit $X_i \in \overline{\mathbb{IR}}$, denn für kein $m \in X$ ist $f(m) = o$, so daß Fall (b.1) nie eintreten kann. Somit tritt der Fall $X_{i+1} = X_i$ in $\mathbb{IR}$ für kein $i \in \mathbb{N}_o$ ein.

Zu (2.3) des Satzes:
Der Fall $f'(X)=o$ für $X \in \mathbb{IR} \setminus \mathbb{R}$ und $X \subseteq X_o$ kann nur auftreten, falls $f(m) \neq o$ für $m \in X$. Es wäre sonst mit $f'(X)=o$ f = constant = o und die Nullstellen würden sich in X häufen.

Die unter der Modifikation Newton 2(f,X) stattfindende Aufspaltung in zwei neue Intervalle H_1 und H_2 kann nur endlich oft stattdinden, nämlich nur so lange Intervalle $X \subseteq X_o$ mit $f'(X) \in T$, d.h. $o \in f'(X)$ vorliegen.

Nach Voraussetzung aber hat f' in X_o nur endlich viele Nullstellen, so daß nach endlich vielen Schritten diese Nullstellen der Ableitung durch die fortgesetzte Aufspaltung und nach Aussage b) des Lemmas separiert sind. Ab diesem Zeitpunkt treten höchstens bei mehrfachen Nullstellen unter Newton 2(f,X) nur noch Intervalle mit höchstens einer Nullstelle $\hat{x}$ auf, die entweder in H_1 oder aber in H_2 liegt. Nach dem Beweis von (3) des Satzes kann jedoch nach endlich vielen Schritten entschieden werden, in welcher der beiden Intervalle $\hat{x}$ liegt.

Zur Veranschaulichung der verschiedenen Fälle siehe Abbildung 5.3, 5.4, 5.5 und 5.6.

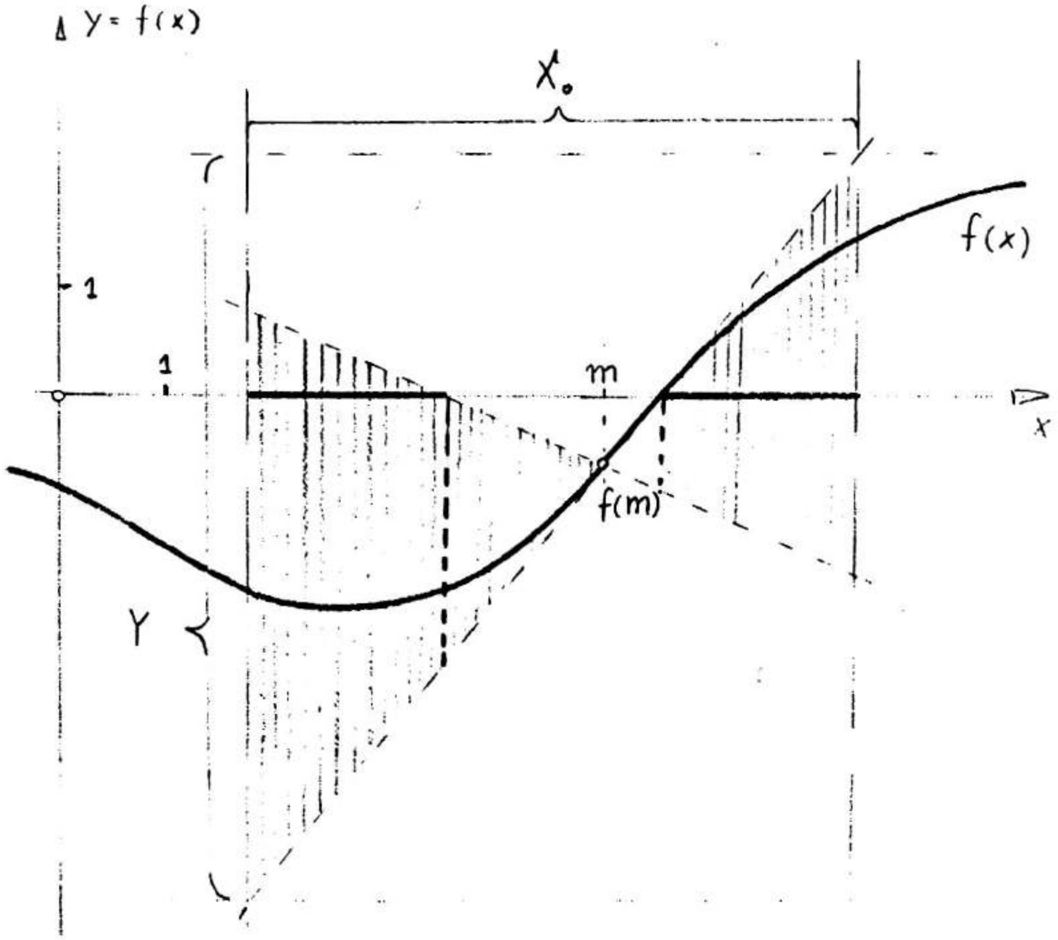


Abb. 5.3

Darstellung der Einschließung in der x-y-Ebene

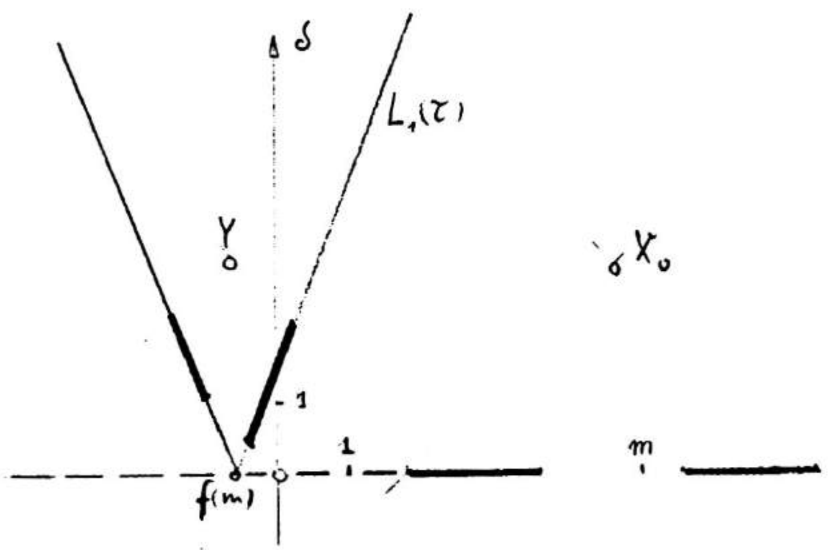


Abb. 5.4

Darstellung der Einschließung in der μ-δ-Ebene

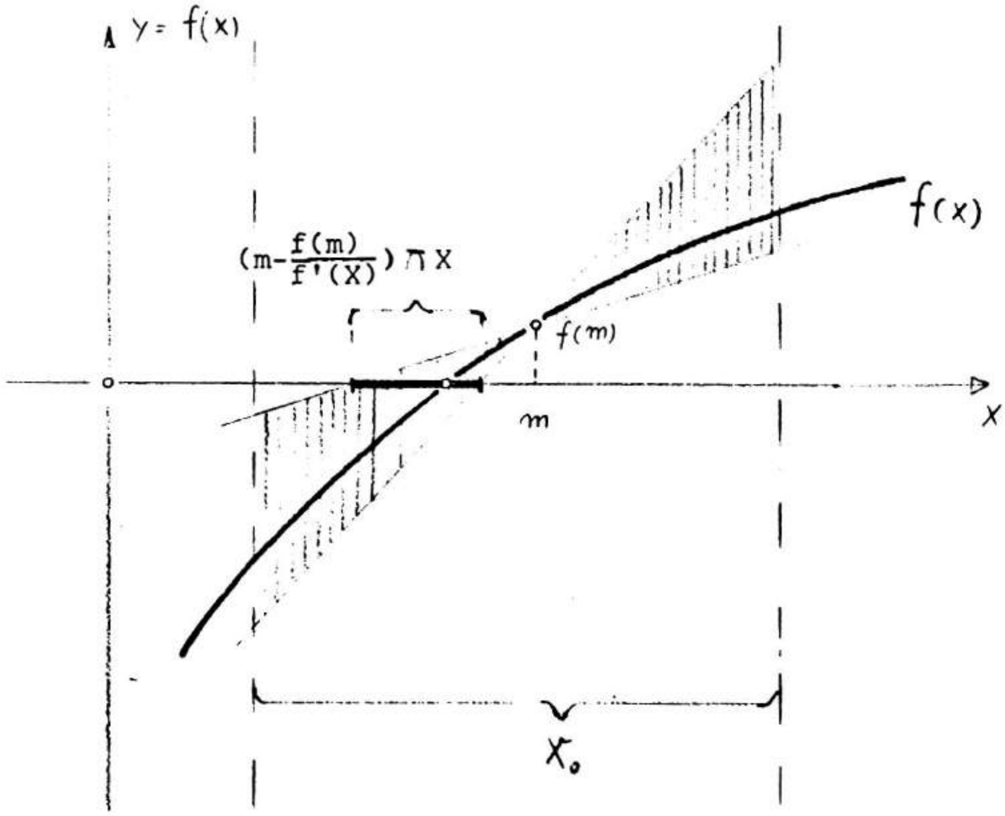


Abb. 5.5

Der Fall für Newton1(f,X)

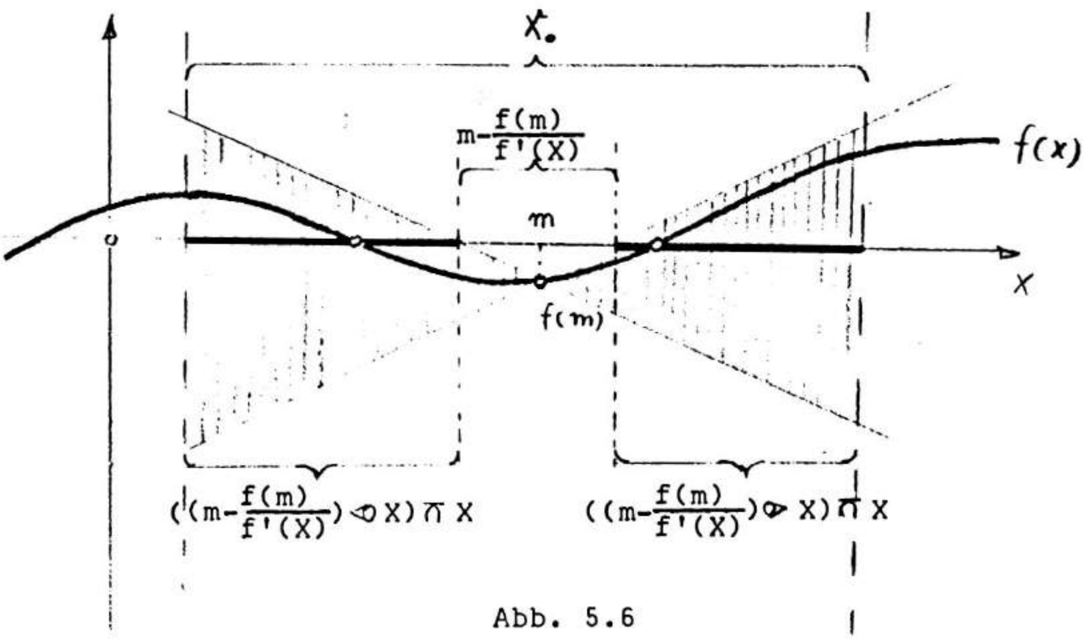


Abb. 5.6

Der Fall für Newton2(f,X)

Für die numerische Behandlung möge der Übersicht halber folgende Zusammenfassung gegeben werden:

Zusammenfassung:

Aus dem Satz 5.4 läßt sich folgender Algorithmus im Raster $(\mathcal{R}\mathbb{H}^*, \{0\}, \diamondplus, \diamondminus, \diamondtimes, \diamonddiv, \subseteq)$ von $\mathbb{H}^*$ zur Berechnung einer Einschließung aller Nullstellen in $X_o \in \mathcal{R}\mathbb{H}^*$ (sofern sie in $\mathcal{R}\mathbb{H}^*$ separierbar sind) angeben.

1) $A \mathbin{\circ\!\!\circ} B := [\lambda A \sqcap \lambda B, \rho A \sqcap \rho B|$

$A \mathbin{\bullet} B := |\lambda A \sqcup \lambda B, \rho A \sqcup \rho B|$

$A \mathbin{\pi} B := |\lambda A \sqcup \lambda B, \rho A \sqcap \rho B|$

$d(A) := \rho A - \lambda A$

$m_\alpha(A) := \alpha \cdot \lambda A + (1-\alpha) \cdot \rho A \qquad \alpha \in [0,1]$ fest .

$$\operatorname{sign}(A) := \begin{cases} 1 & A > 0 \\ 0 & A \in T \\ -1 & A < 0 \end{cases}$$

2) $A \in \mathbb{IR} \iff d(A) \geq 0$

3) $A \in T \iff \operatorname{sign}(A) = 0$

4) Seien αf bzw. $\alpha f'$ gerichtete Rasterrestriktionen von f bzw. f' in $\mathcal{R}\mathbb{H}^*$ mit

(A3) $\bigwedge_{X \in \mathcal{R}\mathbb{H}^*} f(X) \subseteq \alpha f(X)$, bzw. $f'(X) \subseteq \alpha f'(X)$

so läßt sich das modifizierte Newton-Verfahren auf folgende algorithmische Gestalt bringen:

Newton 1(f,X):

a) $X_{i+1}=(m(X_i) \ominus \alpha f(m(X_i)) \oslash \alpha f'(X_i)) \cap X_i$

für $i \geq o$, $X_o=X$, falls $d(X_i) \geq o$ und

$X_{i+1} \neq X_i$

b) Ist $X_{i+1} = X_i$, so ist X_i nullstellenverdächtig

c) sonst liegen keine Nullstellen in X,

Newton 2(f,X):

a) Newton 1(f,X) falls $\mathrm{sign}(\alpha f'(X)) \neq o$

b) Für $\mathrm{sign}(\alpha f'(X)) = o \wedge \alpha f'(X) \neq o$:

$Y:=m(X) \ominus \alpha f(m(X)) \oslash \alpha f'(X)$

$H_1:=(Y \lessdot X) \cap X$

$H_2:=(Y \gtrdot X) \cap X$

Newton $2(f,H_1)$ falls $d(H_1) \geq o$, $H_1 \neq X$

Newton $2(f,H_2)$ falls $d(H_2) \geq o$, $H_2 \neq X$

c) Ist $H_1 = X$ bzw. $H_2 = X$, so ist X nullstellenverdächtig.

d) $\alpha f'(X) = o \;\wedge\; m(X) = X \;\wedge\; \alpha f(m(X)) = o$, so ist $m(X) = X \in \mathbb{R}$ eine Nullstelle.

e) sonst liegen keine Nullstellen in X.

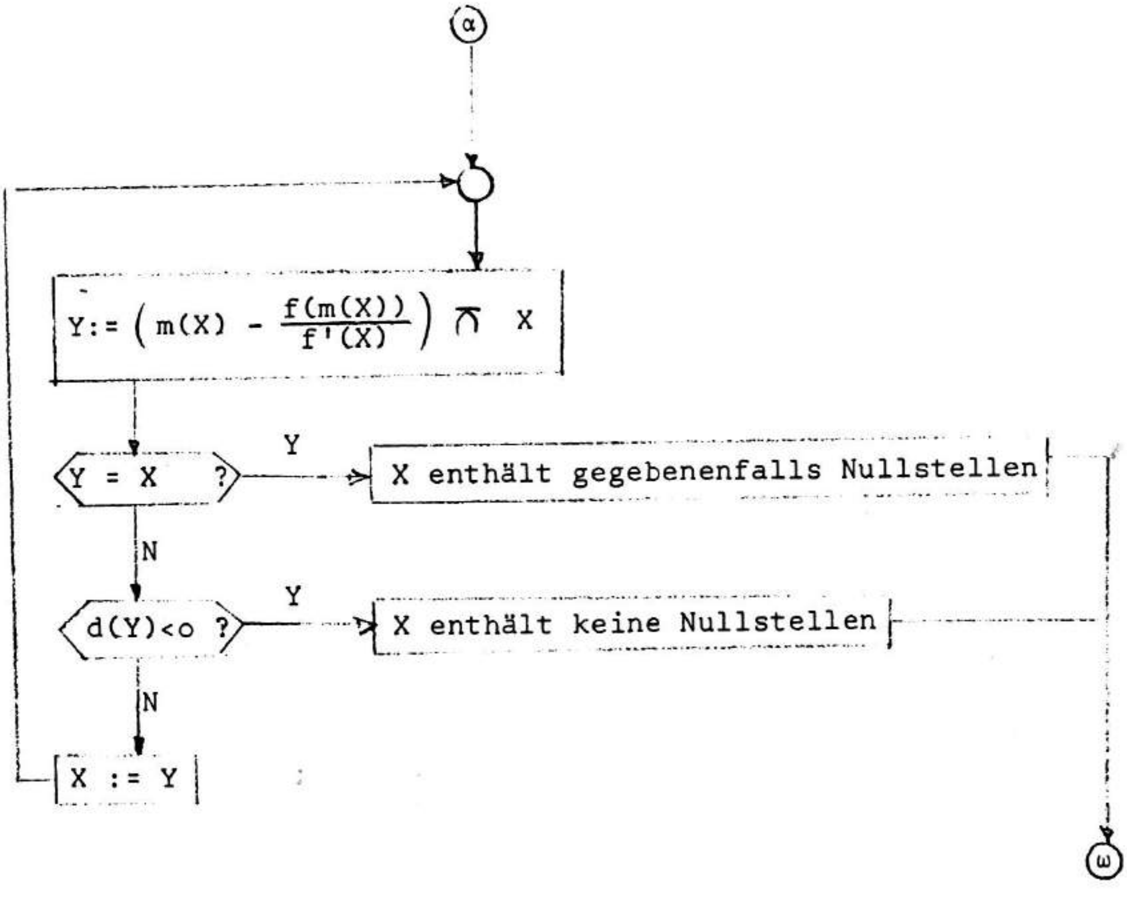
5) Der Algorithmus startet mit Newton $2(f,X_o)$ und berechnet nach Satz 5,3 Einschließungen aller Nullstellen in X_o.

<u>Bew.:</u>

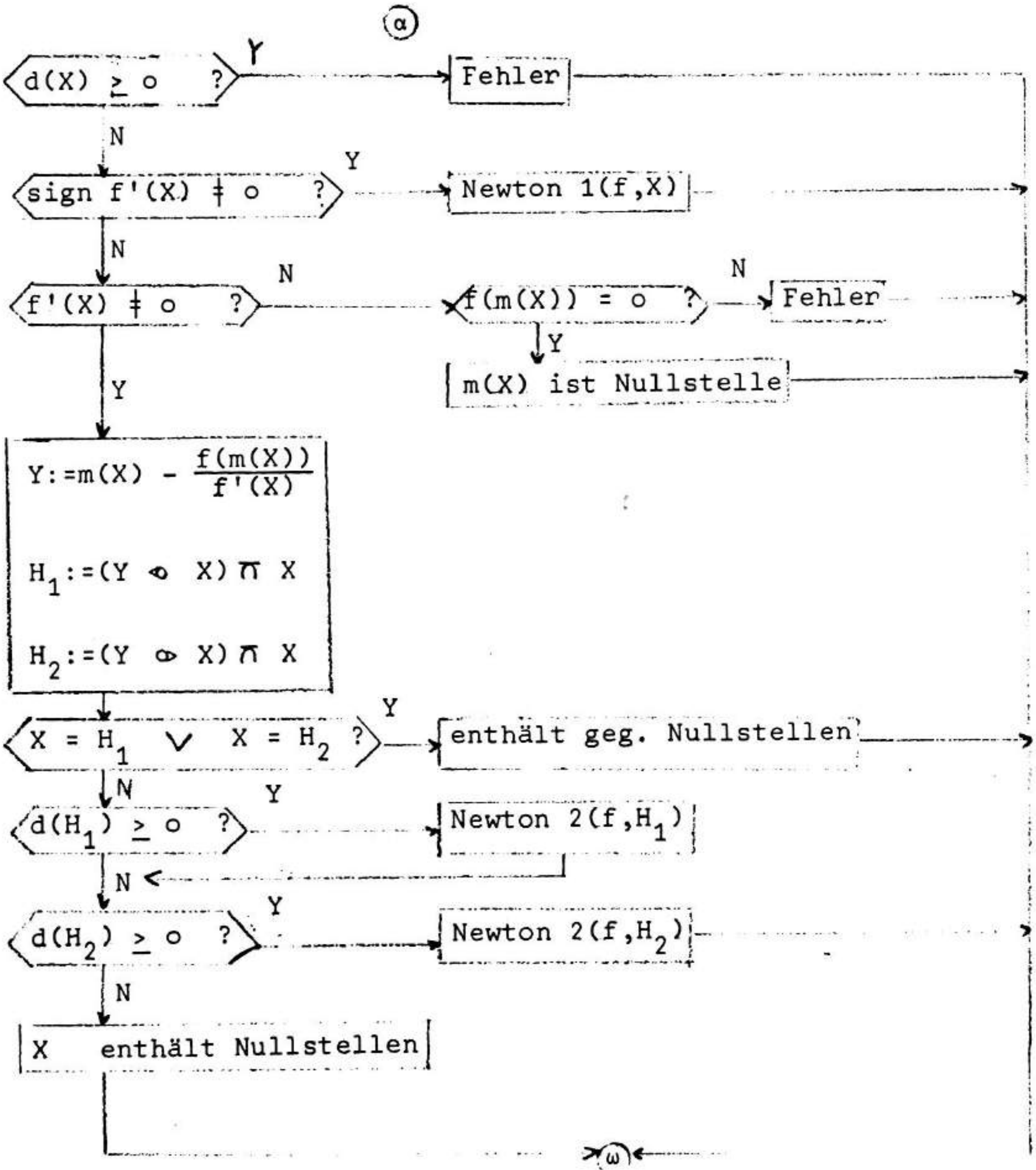
Aufgrund der nach außen gerichteten Verknüpfungen $\diamondminus$ und $\diamondslash$ muß die Aussage b) des Lemmas 5.5 so abgeschwächt werden, daß nur noch $\subseteq$ statt $\subsetneqq$ gilt. Somit brauchen in $\Re\mathbb{H}^*$ nicht mehr alle Nullstellen separierbar zu sein und es kann schon in $\mathbb{IR} \setminus \mathbb{R}$ der Fall $m(X) \diamondminus \alpha f(m(X)) \diamondslash \alpha f'(X) = X$ bzw. $H_1 = X$ bzw. $H_2 = X$ eintreten. Aus der durchweg Isotonie in $\Re\mathbb{H}^*$ bzw. Satz 5.3 für die Nullstellenbestimmung folgt unmittelbar, daß das in $\Re\mathbb{H}^*$ durchgeführte Verfahren in jedem Falle eine Einschließung liefert.

Zur Programmierung eignet sich am besten die rekursive Darstellung mittels der beiden Prozeduren Newton $1(f,X)$ und Newton $2(f,X)$. Die Berechnung startet mit Newton $2(f,X_o)$.

Newton 1(f,X)


α
$Y := \left(m(X) - \frac{f(m(X))}{f'(X)} \right) \bar{\cap} X$
Y = X ?
Y
X enthält gegebenenfalls Nullstellen
N
d(Y)<o ?
Y
X enthält keine Nullstellen
N
X := Y
ω

Newton 2(f,X)


α
d(X) ≥ o ?
Y
Fehler
N
sign f'(X) ≠ o ?
Y
Newton 1(f,X)
N
f'(X) ≠ o ?
N
f(m(X)) = o ?
N
Fehler
Y
m(X) ist Nullstelle
Y
Y:=m(X) - f(m(X)) / f'(X)
H1:=(Y ⊲ X) ⊓ X
H2:=(Y ⊳ X) ⊓ X
X = H1 ∨ X = H2 ?
Y
enthält geg. Nullstellen
N
d(H1) ≥ o ?
Y
Newton 2(f,H1)
N
d(H2) ≥ o ?
Y
Newton 2(f,H2)
N
X enthält Nullstellen
ω

6. ANWENDUNGSBEISPIELE ZU KAPITEL III

In diesem Abschnitt möge kurz skizziert werden, wie in $\mathbb{H}$ implementierbare Raster und Rundungen eingeführt und die Aussagen des Abschnittes I.3 ausgenutzt werden können.

6.1 Das Raster $\mathcal{R}\mathbb{H}^*$ und zwei optimale gerichtete Rundungen

Will man z. B. auf Rechenanlagen Intervallrechnung über $\mathbb{H}^*$ betreiben, so muß natürlich wie bei den reellen Zahlen $\mathbb{R}^*$ nach Abschnitt I.4 eine endliche abzählbare Menge von Intervallzahlen $\mathcal{R}\mathbb{H}^*$ aus $\mathbb{H}^*$ ausgewählt werden.

Da $\mathbb{H}^*$ isomorph zum $\mathbb{R}^{*2}$ ist, ist es üblich, die Maschinenintervalle $\mathcal{R}\mathbb{H}^*$ (isomorph zu $(\mathcal{R}\mathbb{R}^*)^2$) durch die Paarbildung $\mathcal{R}\mathbb{R}^* \times \mathcal{R}\mathbb{R}^*$ darzustellen.

<u>Satz 6.1:</u>

Ist $\mathcal{R}\mathbb{R}^*$ ein Raster von $\mathbb{R}^*$, so ist $\mathcal{R}\mathbb{H}^* := \mathbb{I}\mathcal{R}\mathbb{R}^* \cup \overline{\mathbb{I}\mathcal{R}\mathbb{R}^*}$ stets ein Raster von $\mathbb{H}^*$.

<u>Bew.:</u>

$\mathbb{H}^*$ ist nach Satz 2.28 ein Vollverband. Nach Voraussetzung ist $\mathcal{R}\mathbb{R}^*$ vollständiger Teilverband von $\mathbb{R}^*$. Auf Grund der in der Herleitung von Satz 2.40 beschriebenen Ordnungsisomorphie gilt für die Teilmenge $\mathcal{R}\mathbb{H}^*$:

$(\mathcal{R}\mathbb{H}^*, \subseteq)$ ist ordnungsisomorph zu dem vollständigen Teilverband $(\mathcal{R}\mathbb{R}^* \times \mathcal{R}\mathbb{R}^*, \leq')$, also ist $(\mathcal{R}\mathbb{H}^*, \subseteq)$ ein vollständiger Teilverband von $(\mathbb{H}^*, \subseteq)$. Damit ist (S2") aus Definition I,1.2 gezeigt. (S1") wird wie folgt erfüllt:

$o(\mathbb{H}^*) = \overline{P} = o(\mathcal{R}\mathbb{H}^*)$ und $i(\mathbb{H}^*) = P = i(\mathcal{R}\mathbb{H}^*)$
mit $P := [-p,p]$ und $p \in \mathbb{R}^*$ nach Abschnitt I,4.

<u>Satz 6.2:</u>

$\mathcal{R}\mathbb{H}^*$ besitzt folgende Symmetrieeigenschaften:

$$\bigwedge_{A \in \mathcal{R}\mathbb{H}^*} \overline{A} \in \mathcal{R}\mathbb{H}^*$$

Ist zusätzlich $\mathcal{R}\mathbb{R}^*$ symmetrisch, so gilt:

$$\bigwedge_{A \in \mathcal{R}\mathbb{H}^*} -A \in \mathcal{R}\mathbb{H}^* \wedge -\overline{A} \in \mathcal{R}\mathbb{H}^*$$

<u>Bew.:</u>

Mit $A \in \mathcal{R}\mathbb{H}^*$ ist $(\lambda A, \rho A) \in \mathcal{R}\mathbb{R}^* \times \mathcal{R}\mathbb{R}^*$ und nach Voraussetzung ist dann sowohl $\lambda A, \lambda\overline{A} = \rho A$ und $\lambda(-A) = -\rho(A)$ aus $\mathcal{R}\mathbb{R}^*$ als auch $\rho A, \rho(\overline{A}) = \lambda(A)$ und $\rho(-A) = -\lambda(A)$ aus $\mathcal{R}\mathbb{R}^*$.

<u>Satz 6.3:</u>

Sind $\Delta : \mathbb{IR}^* \to \mathcal{R}\mathbb{IR}^*$ und $\nabla : \mathbb{IR}^* \to \mathcal{R}\mathbb{IR}^*$ die nach Satz I,1.4 optimalen gerichteten Rundungen, so ist die Rundung

$\lozenge A := [\nabla,\Delta]A := [\nabla\lambda A, \Delta\rho A]$ ebenfalls eine optimale gerichtete Rundung $\lozenge : \mathbb{IH}^* \to \mathcal{R}\mathbb{IH}^*$.

<u>Bew.:</u>

Die Eigenschaften von Definition I,1.3 lassen sich stets analog auf die von Δ bzw. ∇ zurückführen.

Insbesondere gilt

(R4) $-\lozenge A = \lozenge(-A)$.

<u>Bemerkung 6.4:</u>

Analog wie in symmetrischen Rastern $\mathcal{R}\mathbb{IR}^*$ wegen $-\nabla a = \Delta(-a)$ die beiden Rundungen ∇ und Δ zusammenhängen, kann in $\mathcal{R}\mathbb{IH}^*$ eine zu $\lozenge$ entsprechend nach (innen) unten gerichtete Rundung $⧖$ in $\mathcal{R}\mathbb{IH}^*$ definiert werden.

Es gilt:

$\overline{\lozenge A} = ⧖\, \overline{A}$ mit $⧖ A := [\Delta\lambda A, \nabla\rho A]$.

Schreibt man formal symbolisch

$\lozenge = [\nabla,\Delta]$ und $⧖ = [\Delta,\nabla]$, so ist die formale Konjugation des Rundungsoperators:

$\overline{\Diamond} = \overline{\mathbb{X}}$ und es gilt die Regel:

$\overline{\Diamond A} = \overline{\Diamond}\ \overline{A} = \overline{\mathbb{X}}\ A$.

Nach Definition 4.6 ist $p : \mathbb{H}^* \times \mathbb{H}^* \to \mathbb{H}^*$ mit $p(X,Y) := X - \overline{Y}$ eine Prämetrik. Die Frage, ob nun bezüglich dieser Prämetrik $\Diamond$ eine quasilinear isoton beschränkte Rundung ist, läßt sich wegen $p(o,Y) = -\overline{Y}$ auf das Problem zurückführen, ob $\Diamond Y \subseteq \phi(Y)$ gilt mit einem $\phi \in \mathcal{L}_1$.

<u>Satz 6.5:</u>

Sind Δ und ∇ nach Definition II,2.10 quasilinear isoton beschränkte Rundungen, d.h. es gilt:

$$\text{(R6)} \quad \bigwedge_{x \in \mathbb{R}^*} (\nabla x \subseteq \psi_1(x) \quad \wedge \quad \Delta x \subseteq \psi_2(x))$$

so ist $\Diamond$ ebenfalls quasilinear isoton beschränkt mit

$$\text{(R6)} \quad \bigwedge_{X \in \mathbb{H}^*} \Diamond X \subseteq \phi(X)$$

<u>Bew.:</u>

Es ist $\Diamond X = [\nabla\lambda X, \Delta\rho X]$ und nach Satz III,2.49 gilt für $\Diamond X \in \mathbb{IR}^*$:

$$\Diamond X = [\nabla\lambda X, \Delta\rho X] = \nabla\lambda X \;\underline{\cup}\; \Delta\rho X$$

(nach Voraussetzung)

$$\subseteq \psi_1(\lambda X) \;\underline{\cup}\; \psi_2(\rho X)$$

Mit $\psi_1(x) = Ax+B$ und $\psi_2(Y) = Cy+D$ folgt weiter:

$$\subseteq (A\lambda X+B) \;\underline{\cup}\; (C\rho X+D)$$

$$\subseteq ((A \;\underline{\cup}\; C)\lambda X + (B \;\underline{\cup}\; D)) \;\underline{\cup}\; ((A \;\underline{\cup}\; C)\rho X + (B \;\underline{\cup}\; D))$$

und nach Satz 2.44:

$$\subseteq (A \;\underline{\cup}\; C)(\lambda X \;\underline{\cup}\; \rho X) + (B \;\underline{\cup}\; C)$$

$$= \tilde{A} * X + \tilde{B} = \phi(X)$$

Dieselbe Abschätzung folgt für $\Diamond X \in \overline{I\!I\!R^*}$ womit die Behauptung bewiesen ist.

Die Voraussetzungen des Satzes 2.11 können jedoch nicht alle erfüllt werden. Es ist $p(X,Y) := X - \overline{Y}$ nach Satz 4.6 eine Prämetrik, aber es gilt:

$$p(A * X, A * Y) = A * X - \overline{A * Y} \nleq A * (X - \overline{Y}) = A * p(X,Y).$$

Mit Ausnahme dieser Eigenschaft (P7) gelten jedoch alle anderen:

$$\mathcal{V}^* := \{D_i \mid D_i(X,Y) := X - \overline{\delta\phi^{(i)}Y} \;\wedge\; i \in \mathbb{N}_0\}$$

ist dann ein prämetrisches Feld, das den Eigenschaften (PF1) (PF2) (PF3) (PF5) und (PF6) mit Ausnahme von (PF4) genügt.

(PF4) gilt jedoch, wenn man statt des Intervallproduktes das hyperbolische Produkt einführt.

$(\mathbb{H}^*, +, \;, p)$ ist dann ein prämetrisches Ringoid und $v^{\ell *}$ ist dann nach Satz II,2.11 ein mit dem Ringoid-Rasterringoid $\mathbb{H}^* \times \mathcal{R}\mathbb{H}^*$ verträgliches quasilineares prämetrisches Feld, wobei (V) gilt, d. h.

$$\text{(V)} \quad \bigwedge_{o \in \{+, \;\}} \; \bigwedge_{X,Y \in \mathcal{R}\mathbb{H}^*} \; X \; \diamondsuit\!\!\!\!\circ \; Y = \Diamond (X \circ Y) \; .$$

Will man dann ein Intervallprodukt abschätzen, etwa $D_i(A * X, A * Y)$, so ist die Verknüpfung $*$ nach Tabelle 2.5 durch das hyperbolische Produkt darzustellen.

Eine einfache Methode bietet Satz II,3.5 durch die Tatsache, daß aus (R6)(R6') folgt, $(\mathbb{H}^*, +, *, q)$ nach Definition 4.9 ein metrisches Ringoid ist und (V) gilt, d. h.

$$\text{(V)} \quad \bigwedge_{o \in \{+, *\}} \; \bigwedge_{X,Y \in \mathcal{R}\mathbb{H}^*} \; X \; \diamondsuit\!\!\!\!\circ \; Y = \Diamond (X \circ Y) \; .$$

Somit ist

$$U^* := \{u_i \mid u_i(X,Y) := q(X,Y) + ((\alpha+1)^i - 1)(q(o,Y) + \frac{\beta}{\alpha}) \wedge i \in \mathbb{N}\}$$

mit

$$\text{(R6')} \; q(X, \Diamond X) \le \|X - \overline{\phi(X)}\| \le \alpha \; \|X\| + \beta$$

ein lineares metrisches Feld, das mit dem Ringoid-Rasterringoid $\mathbb{H} \times \mathcal{R}\,\mathbb{H}$ verträglich ist.

6.2 Anwendung von Abschnitt I.3

Eine Iterationsfolge $\{X_{i+1} = F(X_i)\}_{i \in \mathbb{N}_0}$ mit $X_0 \in D \subseteq \mathbb{H}^*$, das bei exakter Rechnung genau einen Fixpunkt $\hat{X}$ besitzen möge, hat bei gerundeter Rechnung in $\mathcal{R}\mathbb{H}^*$ um so mehr Häufungspunkte in der Umgebung des Fixpunktes $\hat{X}$, je gröber das zugrundegelegte Raster, d. h. je gröber die Rundung ist.

Dieser Effekt wird bei schlecht konditionierten Iterationsfunktionen F, etwa bei Lipchitzkonstanten nahe bei 1,um so größer.
Die inklusionsisotonen Iterationsverfahren, wie sie in [Mayer 68], [Herzberger 69], [Alefeld 68], [Moore 69], [Krawczyk 68], [Nickel 67] und [Barth 71] und andere mehr beschrieben sind, gehen stets von einer Einschließung $X_0 \in \mathbb{IR}$ des exakten Fixpunktes $\hat{X}$ aus. Die numerische Einschließung X^* muß dann notwendig alle oben erwähnten eventuell auftretenden reellen Häufungspunkte einschließen.(vergl. Satz I,3.3 a)).

Die Einschließung wird also um so gröber sein, je gröber der zugrundegelegte Intervallraster $\mathcal{R}\mathbb{H}^*$ ist und je schlechter konditioniert die Iterationsfunktion F ist. In diesen Fällen jedoch kommt nach Satz I,3.2 (c) das antitone Iterationsverfahren um so mehr zur Geltung.

Folgende Beispiele mögen diese Tatsache vorführen, wobei zur Herstellung eines starken Kontrastes sehr grobe Raster mit etwa 10 Dualstellen und Lipschitzkonstanten nahe bei 1 gewählt wurden!

6.2.1 Das Iterationsverfahren $X_{i+1} = F(X_i)$

Sei das Iterationsverfahren

$$(\ast) \quad X_{i+1} = F(X_i),\ i \geq 0,\quad X_i \in D \subsetneq \mathbb{H}^*$$

vorgelegt mit der intervallwertigen Intervallfunktion $F : \mathbb{H}^* \to \mathbb{H}^*$, von der bekannt sei, daß

$$\| F(X) - \overline{F(Y)} \| \leq L \| X - \overline{Y} \| \quad \text{mit} \quad 0 \leq L < 1$$

für $X, Y \in D \subseteq \mathbb{H}^*$ gelte, d. h. das Iterationsverfahren (1) konvergiert gegen den Fixpunkt $\hat{X} \in D$.

Das gerundete isotone Iterationsverfahren

$$X_{i+1} = \alpha F(X_i) \sqcap X_i, \quad i \geq 0,\ X_i \in \tilde{D} \subseteq \mathcal{R}\mathbb{H}^*,\ X_0 \supseteq \hat{X}\ .$$

ist nach Satz I,3.3 a) für die isoton gerichtete Rasterrestriktion αF von F konvergent und liefert eine isotone Einschließungsfolge des exakten Fixpunktes $\hat{X}$, die gegen einen oberen Häufungspunkt X^* mit $\hat{X} \subseteq X^*$ konvergiert.
Das in Satz I,3.3 c) beschriebene Iterationsverfahren lautet

$$Y_{i+1} = (\alpha F(Y_i) \sqcup Y_i) \,\overline{\sqcap}\, X \quad \text{für } i \geq 0,\ Y_0 = \overline{X^*} \subseteq \hat{X}\ ,$$

dessen Konvergenz ebenfalls gesichert ist. Es liefert nach Satz I,3.3(c) eine weitere Einschließung Y^* für die jedoch gilt:

$$\hat{X} \subseteq Y^* \subseteq X^* .$$

Die Abbruchkriterien lauten in beiden Fällen:

Ist $X_{i+1} = X_i$ bzw. $Y_{j+1} = Y_j$, so sind $X_i = X^*$ bzw. $Y_j = Y^*$ die gesuchten Einschließungen.

Im letzteren Falle ist zu beachten, daß i.a. der Iterationsprozeß im Gegensatz zum ersten Verfahren nicht beliebig abgebrochen werden darf. Der Grund liegt darin, daß z. B. für reelles $F = f : \mathbb{R} \to \mathbb{R}$ bei antitoner Iteration vorübergehend der Fall eintreten kann, daß ab einem Index $k \in \mathbb{N}$ $Y_k \nsubseteq \hat{X} \wedge \hat{X} \nsubseteq Y_k$ gilt. Nach Aussage von Satz I.3.3 gilt dies nur vorübergehend, d. h. nach endlich vielen Schritten n tritt stets der Fall ein daß $\hat{X} \subseteq Y_l$ für $l \geq k + n \geq o$ gilt.

Beispiel:

Das Beispiel wurde auf der Rechenanlage ELX8 durch Simulation einer $\tau = 10$ dualstelligen Mantisse, also in einem Raster $\mathcal{R}_2^{10}\mathbb{H}^*$, in Gleitpunktarithmetik gerechnet.

Die Iterationsfunktion F lautet:

F(x) = (0.95 * x/(x+500) + 0.05/(x+500)) * (x+500)

Die Lipschitzkonstante ist $L_F = 0.95$, der Fixpunkt von $x = F(x)$

lautet $\hat{x} = 1$.

Die Rasterrestriktion αF ergibt sich durch die übliche nach außen gerichtete Rundung der einzelnen Verknüpfungen.

$$\alpha F(x) = (0.95 \diamondsuit\!\!\!\div\ X \diamondsuit\!\!\!/\ (x \diamondsuit\!\!\!+\ 500) \diamondsuit\!\!\!+\ 0.05 \diamondsuit\!\!\!/\ (x \diamondsuit\!\!\!+\ 500)) \diamondsuit\!\!\!\div\ (x \diamondsuit\!\!\!+ 500)$$

a) Die isotone Iteration liefert mit

$X_o = [0.1, 2.7] \quad X^* = [0.891, 1.18]$.

b) Das antitone Verfahren liefert mit

$Y_o = \overline{X_o} = [2.7, 0.1] \quad Y^* = [0.924, 1.13]$.

Die Güte, das Verhältnis der beiden Durchmesser, liefert:

$$\sigma = \frac{d(X^*)}{d(Y^*)} = 1.4\ldots$$

6.2.2 Iterationsverfahren für lineare (quasilineare) Gleichungssysteme

Es sei

(1) $X = \mathfrak{A} * X + B$ vorgelegt mit $\mathfrak{A} \in M_n \mathbb{IR}, X, B \in V_n \mathbb{IH}$ und Spektralradius $\rho(|\mathfrak{A}|) < 1$.

Der Fixpunkt $\hat{X}$ des Systems (1) werde durch ein Iterationsverfahren

(2) $X_{i+1} = (\mathfrak{A} * X_i + B) \sqcap X_i$ mit $X_o \in V_n \mathbb{IH}$, $i \geq o$

als Einzelschrittverfahren durchgeführt.

Die entsprechenden Verfahren im Gleitpunktraster $\mathcal{R}_2^{10}\mathbb{H}$ seien:

a) <u>isoton:</u> $X_{i+1} = (\mathfrak{A} \diamond X_i \diamond B) \sqcap X_i$, $X_o \geq \hat{X}$

b) <u>antiton:</u> $Y_{i+1} = ((\mathfrak{A} \diamond Y_i \diamond B) \sqcup Y_i) \sqcap X^*$ mit

$Y_o = \overline{X^*} \leq \hat{X}$, wobei X^*

Fixpunkt von a) sei.

Liegt der Spektralradius $\rho(|\mathfrak{A}|)$ nahe bei 1 wie es in den folgenden Beispielen der Fall ist, so liefert das antitone Verfahren b) i.a. eine bessere Einschließung als a), so daß gilt:

$$\hat{X} \leq Y^* \subset X^* .$$

Als Gütezahl der Verbesserung wird die Zahl

$$\tau = \max_{i=1,\dots,n} \tau_i = \max_{i=1,\dots,n} \frac{d(X^*)_i}{d(Y^*)_i} \quad \text{berechnet}$$

(vergl. Definition 3.1(6))

<u>Beispiel 1:</u> Raster $\mathcal{R}\mathbb{H}^* = \mathcal{R}_2^{10}\mathbb{H}^*$.

Mit $n = 1$, $\rho(|\mathfrak{A}|) = 0.99$, $\mathfrak{A} = 0.99$, $B = 0.01$ und $X = 1$ erhält man ausgehend von $X_o = [0,2]$ bzw. $Y_o = [2,0]$
$X^* = [0.781, 1.57]$ bzw. $Y^* = [0.877, 1.37]$
und $\tau = 1.6$.

Beispiel 2:

$$n = 2, \quad \mathfrak{A} = \begin{pmatrix} 0.95 & 0.00001 \\ 0.9 & 0.005 \end{pmatrix}, \qquad \rho(|\mathfrak{A}|) \leq 0.96$$

$$B = \begin{pmatrix} [\ 0.0401, & 0.001] \\ [100 & , 100\ ] \end{pmatrix}$$

In diesem Beispiel ist offensichtlich $\mathfrak{A}$ schlecht konditioniert bezüglich der 1. Komponente, was sich auch in dem Unterschied von X^* und Y^* zeigt:

$$X_o = \begin{pmatrix} [\ 0.01, & 0.06] \\ [80 & , 120\ ] \end{pmatrix}, \qquad Y_o = \overline{X^*},$$

$$X^* = \begin{pmatrix} [\ 0.0378, & 0.0452] \\ [105 & , 106\ ] \end{pmatrix},$$

$$Y^* = \begin{pmatrix} [\ 0.039, & 0.0441] \\ [105 & , 106\ ] \end{pmatrix},$$

$$\tau = 1.45 .$$

Beispiel 3:

$$n = 3, \quad \mathcal{A} = \begin{pmatrix} 0.55 & -0.44 & 0 \\ 0 & 0.55 & -0.44 \\ -0.44 & 0 & 0.55 \end{pmatrix}, \quad \rho(|\mathcal{A}|) = 0.99 ,$$

$$B = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \hat{X} = \begin{pmatrix} 1.12 \\ 1.12 \\ 1.12 \end{pmatrix}, \quad X_o = \begin{pmatrix} [0, & 2.5] \\ [0, & 2.5] \\ [0, & 2.5] \end{pmatrix}, \quad Y_o = \overline{X^*} ,$$

$$X^* = \begin{pmatrix} [0.686, & 1.57] \\ [0.687, & 1.57] \\ [0.683, & 1.57] \end{pmatrix}, \quad Y^* = \begin{pmatrix} [0.864, & 1.39] \\ [0.865, & 1.39] \\ [0.862, & 1.39] \end{pmatrix} ,$$

$$\tau = 1.7 .$$

Beispiel 4:

$$n = 3 \quad \mathcal{A} = \begin{pmatrix} 0.3 & 0.3 & 0.3 \\ 0.3 & 0.4 & 0.29 \\ 0.29 & 0.3 & 0.4 \end{pmatrix}, \quad \rho(|\mathcal{A}|) = 0.99 ,$$

$$B = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad X_o = \begin{pmatrix} [15, & 30] \\ [15, & 33] \\ [15, & 33] \end{pmatrix}, \quad \hat{X} \approx \begin{pmatrix} 24.5 \\ 26.9 \\ 26.8 \end{pmatrix} ,$$

$$X^{*} = \begin{pmatrix} [22.3, & 28.3] \\ [24.5, & 31.1] \\ [24.5, & 31.1] \end{pmatrix}, \qquad Y^{*} = \begin{pmatrix} [22.9, & 27.3] \\ [25.2 & 30.0] \\ [25,2 & 30.0] \end{pmatrix} ,$$

$$\tau = 1.37 \quad .$$

L i t e r a t u r

[Alefeld 68] Intervallrechnung über den komplexen Zahlen und einige Anwendungen; Dissertation an der Universität Karlsruhe (1968)

[Barth 71] Nullstellenbestimmung mit der Intervallrechnung Computing 8, 320-328 (1971)

[Bourbaki] Éléments de mathématique; livre II, Algèbre, chapitre I, §2,4., 1964

[Christ 68] Realisierung einer Maschinen-Intervallarithmetik mit beliebigen ALGOL 60-Compilern; Elektronische Rechenanlagen 10 (1968), H. 5, 217-222

[Collatz 68] Funktionalanalysis und numerische Mathematik; Springer Verlag, 1968

[Greub 67] Linear Algebra; Grundlehren der mathematischen Wissenschaften, Bd. 97, Springer Verlag 1967

[Henrici 72] Elemente der numerischen Analysis; Band 2, B I, deutsche Übersetzung 1972

[Herzberger 69] Metrische Eigenschaften von Mengensystemen und einige Anwendungen; Dissertation an der Universität Karlsruhe (1969)

[Kahan 68] A More Complete Interval Arithmetic; University of Toronto (1968)

[Krawczyk 68] Einige Newton-Algorithmen für Intervallarithmetik; Bericht d. Lehrst. f. Num. Math., Universität Karlsruhe, 1968

[Kulisch 69]

1. An Axiomatic Approach to Rounded Computations; Mathematics Research Center, University of Wisconsin, Technical Summary Report 1020, 1-29, 1969

2. Grundzüge der Intervallrechnung; Überblicke Mathematik 2, 1969

[Kulisch 70]

1. On The Concept Of A Screen; Mathematics Research Center, University of Wisconsin, Technical Summary Report 1084, 1-12, 1970

2. Rounding Invariant Structures; Mathematics Research Center, University of Wisconsin, Technical Summary Report 1103, 1-47, 1970

3. Interval Arithmetic Over Completely Ordered Ringoids; Mathematics Research Center, University of Wisconsin, Technical Summary Report 1105, 1-56, 1970

[Kulisch 72] Vorlesung, gehalten im Wintersemester 1971/72 an der Universität Karlsruhe

[Mayer 68] Über die in der Intervallrechnung auftretenden Räume und einige Anwendungen;
Dissertation an der Universität Karlsruhe (1968).

[Moore 69] Intervall Analysis;
Übersetzung von: Interval Analysis, Prentice Hall, 1966

[Nickel 67] Die vollautomatische Berechnung einer einfachen Nullstelle von F(t) = o einschließlich einer Fehlerabschätzung;
Computing 2, 1967, 232-245

[Ortolf 69] Eine Verallgemeinerung der Intervallarithmetik;
Bonn: Gesellschaft für Mathematik und Datenverarbeitung 1969

[Schäfke 70] Zum Anwendungsbereich einiger Iterationsverfahren;
Num. Math. 15, 109-121 (1970)

[Ullrich 72] Rundungsinvariante Strukturen mit äußeren Verknüpfungen;
Dissertation an der Universität Karlsruhe (1972).

[Walter 64] Differential- und Integral-Ungleichung; Springer-Verlag 1964

[Wilkinson 69] Rundungsfehler; Springer-Verlag 1969

Lebenslauf

	Edgar Kaucher
geboren	am 8.1.1945 in Bauschlott, Baden-Württemberg
Eltern	Willi A. Kaucher, techn. Kaufmann Maria Kaucher, geb. Seitter
Religion	evangelisch
Schulbildung	von 1951 bis 1956 Besuch der Volksschule in Stein, von 1956 bis 1965 des Kepler-Gymnasiums in Pforzheim
Reifeprüfung	im Februar 1965 am Kepler-Gymnasium Pforzheim
Wehrdienst	von 1965 bis 1966
Studium	der Mathematik und Physik an der Universität Karlsruhe vom Wintersemester 1966/67 bis Sommersemester 1971
Diplomprüfung	für Mathematik an der Universität Karlsruhe im Juni 1971
Berufstätigkeit	vom 1. August 1971 bis zum 31. August 1972 wissenschaftlicher Angestellter an der Fakultät für Informatik der Universität Karlsruhe, seit 1. September 1972 wissenschaftlicher Angestellter am Institut für Angewandte Mathematik der Universität Karlsruhe
Eheschließung	am 8.10.1971 mit Marianne, geb. Wächter aus Weingarten, Baden-Württemberg