Über die Divergenzpunkte des Newtonschen Verfahrens zur Bestimmung von Wurzeln algebraischer Gleichungen IV.

Von BÉLA BARNA (Debrecen)

Wir beschäftigen uns in dieser Arbeit — wie in den vorigen Teilen I., II. und III. — mit dem Newtonschen Näherungsalgorithmus angewandt auf Polynome mit reellen Koeffizienten und Nullstellen. Durch die Anwendung der Ergebnisse des Teiles III. werden wir hier die Menge der Divergenzpunkte nach ihrem Maß studieren. Als Resultat wird sich ergeben, daß diese Menge eine Nullmenge ist.

Bilden wir die ersten invers-iterierten Intervalle des unmittelbaren Konvergenzintervalls $(-\infty, B_1)$ (siehe die Figur). Diese sind die Intervalle $\big((B_1)_{-1}, B_1\big)$ $(i=2, 3, \ldots, k-1)$, wo der Punkt $(B_1)_{-1}$ der (einzige) invers-iterierte Punkt von B_1 im Intervall (B_{i-1}, B_i) ist. Es sei

$$\big((B_1)_{-1}, B_i\big) \equiv \Omega_i'', \ (B_1)_{-1} \in (B_{i-1}, B_i), \qquad i=2, 3, \ldots, k-1.$$

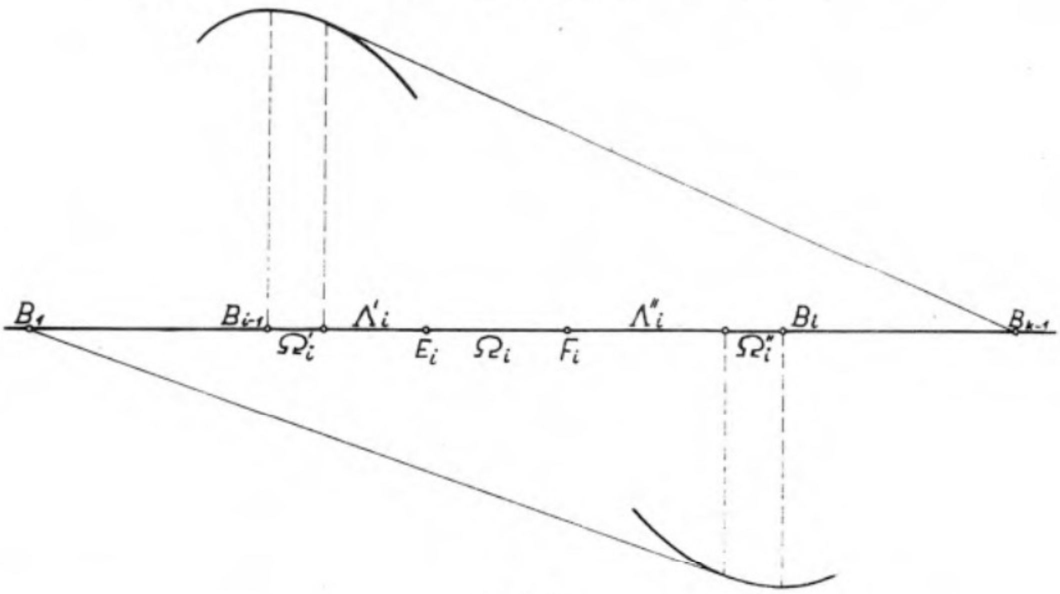

Fig. 1

Analogerweise bilden wir die ersten invers-iterierten Intervalle von $(B_{k-1}, +\infty)$. So bekommen wir die Intervalle

$$\big(B_{i-1}, (B_{k-1})_{-1}\big) \equiv \Omega_i', \ (B_{k-1})_{-1} \in (B_{i-1}, B_i), \qquad i=2, 3, \ldots, k-1.$$

Es sei noch

$$(E_i, F_i) \equiv \Omega_i, \qquad i=2, 3, \dots, k-1.$$

Wir nennen die so erhaltenen vollständigen Konvergenzintervalle Ω_i', Ω_i Ω_i'' die *Ω-Intervalle*. Die von diesen nicht bedeckten Teilintervalle von (B_1, B_{k-1}) sind die Λ-Intervalle:

$$\Lambda_i' \equiv [(B_{k-1})_{-1}, E_i], (B_{k-1})_{-1} \in (B_{i-1}, E_i), \Lambda_i'' = [F_i, (B_1)_{-1}], (B_1)_{-1} \in (F_i, B_i).$$

In diesen (geschlossenen) Intervallen sind $N'(x)$ und $N''(x)$ stetige Funktionen, und so gibt es eine Zahl H, für welche $|N''(x)| \leqq H$ gilt, wenn nur der Punkt x in einem Λ-Intervall liegt.

Beachten wir jetzt, daß die Menge $\mathscr{K}$ aller Konvergenzpunkte $x \in (B_1, B_{k-1})$ aus den Punkten der invers-iterierten Intervalle der Ω-Intervalle besteht, d.h. aus den Intervallen $(\Omega)_{-n}$ $(n=0, 1, 2, \dots)$. Daraus folgt die Meßbarkeit der Punktmenge $\mathscr{K}$. Zur Bestimmung des Maßes von $\mathscr{K}$ müssen wir uns ausführlicher mit der Menge der Intervalle $(\Omega)_{-n}$ beschäftigen.

Wir betrachten vor allem die invers-iterierten Intervalle erster Ordnung der unmittelbaren Konvergenzintervalle Ω_i. Jedes Intervall Ω_i wird von der Funktion $N(x)$ auf sich selbst abgebildet, es ist also ein invers-iteriertes Intervall von sich selbst. Da aber zwischen den vollständigen Konvergenzintervallen das Intervall Ω_i vor kommt, verabreden wir diese *nicht* zu seinen eigen Invers-iterierten zu zählen, d.h. es ist

$$(\Omega_i)_{-1} \neq \Omega_i. \tag{1}$$

Dann können wir einsehen: sämtliche invers-iterierten Intervalle $(\Omega)_{-n}$ aller Ω-Intervalle sind paarweise fremde Intervalle (d.h. sie haben keinen gemeinsamen Punkt).

Enthalten nämlich irgend zwei $(\Omega)_{-n}$-intervalle einen gemeinsamen Teilintervall, so enthält einer von beiden wenigstens einenen Grenzpunkt des anderen, *oder* fallen sie zusammen.

Der erste Fall ist unmöglich, weil die Grenzpunkte als Punkte von der Form $(B_i)_{-n}$, $(E_i)_{-n}$ oder $(F_i)_{-n}$ Divergenzpunkte sind, sie können also nicht in einem Konvergenzintervall liegen. Es ist auch unmöglich, daß gewisse invers-iterierten Intervalle (gleicher oder ungleicher Ordnung) von verschiedenen Ω-Intervallen zusammenfallen. Dies bringt mit sich, daß die Punkte dieser gemeinsamen invers-iterierten Intervalls eine gegen zwei verschiedene Grenzwerte konvergierende Iteriationsfolge haben. Daraus schließlich, daß *ein* Ω-Intervall zwei zusammenfallende invers-iterierte Intervalle hat, z.B.

$$(\Omega)_{-n'} = (\Omega)_{-n''}, \; n'' \geqq n', \tag{2}$$

folgt im Fall $n'' > n'$ für die iterierten Intervalle $(n''-1)$-ter Ordnung

$$(\Omega)_{n''-n'-1} = (\Omega)_{-1}.$$

Dies ist nur dann möglich, wenn dieses Ω-Intervall ein Ω_i ist, und so folgt $\Omega_i = (\Omega_i)_{-1}$ im Gegensatz zu (1). Ist dagegen $n'' = n'$, so erhält man aus (2) (durch $(n'-1)$-malige Iteration) das Zusammenfallen zweier invers-iterierten Intervalle erster Ordnung, was der Tatsache widerspricht, daß die Invers-iterierten erster Ordnung eines Ω-Intervalls in verschiedenen (B_{i-1}, B_i) -Intervallen liegen.

Zufolge der vorigen können wir sagen: *die unendliche Reihe der Maßzahlen sämtlicher Intervalle* $(\Omega)_{-n}$ $(n=0, 1, 2, \ldots)$ *ist (absolut) konvergent*. Diese summe ist das Maß der Menge $\mathscr{K}$.

Bei der Ausrechnung dieser Summe bezeichnen wir mit $|\sigma|$ die Länge des Intervalls σ. Es wird dies auch denn kein Mißverständnis hervorrufen, wenn wir das Symbol bei Zahlen in dem gewöhnlichem Sinn (als Zeichen des absoluten Wertes) verwenden. Die Summe der Längenmaße sämtlicher invers-iterierten Intervalle n-ter Ordnung des Intervalls σ schreiben wir in der Form $S_{-n}(\sigma)$, d.h. es ist

$$S_{-n}(\sigma) = \sum_{\sigma_{-n}} |\sigma_{-n}|.$$

Die Intervalle

$$\Omega_i', \Lambda_i', \Omega_i, \Lambda_i'', \Omega_i''$$

bedecken „vollständig und einfach" das Intervall (B_{i-1}, B_i); damit wollen wir ausdrücken, daß jeder Punkt von (B_{i-1}, B_i) zu einem, und nur zu einem der vorigen Intervalle gehört. Zufolge der Definition des Symbols S_{-n} können wir schreiben:

$$S_0(\Omega_i') + S_0(\Lambda_i') + S_0(\Omega_i) + S_0(\Lambda_i'') + S_0(\Omega_i'') = B_i - B_{i-1}.$$

Wir wollen noch mit $S_{-n}(\Omega)$ und mit $S_{-n}(\Lambda)$ die Summe der Längen aller $(\Omega)_{-n}$-, bzw. aller $(\Lambda)_{-n}$- Intervalle bezeichnen. Also gilt

$$S_0(\Omega) + S_0(\Lambda) = B_{k-1} - B_1.$$

Die Intervalle $(\Omega)_{-1}$ und $(\Lambda)_{-1}$ haben mit keinem Ω-Intervall Punkte gemeinsam; die Intervalle $(\Omega)_{-n}$ $(n=0, 1, 2, \ldots)$ sind nämlich (paarweise) fremde Intervalle, es liegen ferner in jedem Teilintervall eines jeden Λ-Intervalls $(\Omega)_{-n}$-Intervalle, also liegt in einem Ω-Intervall kein $(\Lambda)_{-n}$-Intervall (oder ein Teil davon). Daraus folgt, daß die Intervalle $(\Omega)_{-1}$ und $(\Lambda)_{-1}$ Teilintervalle von Λ-Intervallen sind. Anderseits ist $N(x)$ in Λ' und in Λ'' monoton abnehmend, und das Intervall

$$(\Lambda_i')_1 = [F_i, B_{k-1}]$$

wird von den Intervallen

(4) $$\Lambda_i'', \Omega_i'', \Omega_{i+1}', \ldots, \Omega_{k-1}''$$

und

$$(\Lambda_i'')_1 = [B_1, F_i]$$

von den Intervallen

(5) $$\Omega_2', \Lambda_2', \Omega_2, \ldots, \Omega_i, \Lambda_i'$$

vollständig und einfach bedeckt. Das bedeutet, daß die in (B_{i-1}, B_i) liegenden ersten invers-iterierten Intervalle der Intervalle (4) und (5) die Intervalle Λ_i' und Λ_i'' vollständig und einfach bedecken. Dies ist für die Werte $i=2, 3, \ldots, k-1$ gültig. So gewinnen wir, das sämtliche Intervalle $(\Omega)_{-1}$ und $(\Lambda)_{-1}$ zusammen sämtliche Λ-Intervalle vollständig und einfach bedecken, d.h.

$$S_0(\Lambda) = S_{-1}(\Lambda) + S_{-1}(\Omega).$$

Dieses Formel können wir mit vollständiger Induktion verallgemeinern: Nehmen wir an, dass sämtliche Intervalle $(\Lambda)_{-n-1}$ von den Intervallen $(\Lambda)_{-1}$ und $(\Omega)_{-n}$ vollständig und einfach bedeckt werden. Dann liegt ein beliebiges $(\Omega)_{-n} \equiv \sigma$ in einem Intervall $(\Lambda)_{-(n-1)}\equiv\tau$. Jedes invers-iterierte Intervall τ_{-1} ist ein Teilintervall eines Λ-Intervalls, und hier ist $N(x)$ streng monoton abnehmend. Deshalb gilt $\sigma_{-1}\subset\tau_{-1}$ d.h. $(\Omega)_{-n+1}$ ist ein Teilintervall von einem $(\Lambda)_{-n}$. Für die Intervalle $(\Lambda)_{-(n+1)}$ sieht man das ebenso ein. Anderseits ist jedes $((\Lambda)_{-n})_1$ ein Intervall $(\Lambda)_{-(n-1)}$, welche nach der Annahme, von Intervallen $(\Omega)_{-n}$ und $(\Lambda)_{-n}$ vollständig und einfach bedeckt wird. Aus der Monotonität von $N(x)$ folgt jetzt, daß das betreffende Intervall $(\Lambda)_{-n}$ durch die Intervalle $(\Omega)_{-(n+1)}$ und $(\Lambda)_{-(n+1)}$ vollständig und einfach bedeckt wird.

Dies bedeutet:

$$S_{-n}(\Lambda) = S_{-(n+1)}(\Omega)+S_{-(n+1)}(\Lambda) \qquad (n=0, 1, 2, \dots),$$

und so folgt weiter sehr einfach:

$$S_0(\Lambda) = S_{-1}(\Omega)+S_{-2}(\Omega)+\dots+S_{-p}(\Omega)+S_{-p}(\Lambda),$$

oder nach (3)

$$S_0(\Omega)+S_{-1}(\Omega)+\dots+S_{-p}(\Omega)+S_{-p}(\Lambda) = B_{k-1}-B_1. \tag{6}$$

Wir beweisen: es ist

$$\lim_{n\to\infty} S_{-n}(\Lambda) = 0.$$

Lemma. *Ist σ ein Teilintervall eines Λ-Intervalls, so ist* $\lim S_{-n}(\sigma)=0$.

Beweis. Die Richtigkeit der Behauptung ist leicht einzusehen, wenn σ ein Konvergenzintervall ist. Dann liegt nämlich σ in einem vollständigen Konvergenzintervall ω, und $\sum_{n=0}^{\infty} S_{-n}(\omega)$ bildet eine Teilreihe der (absulut) konvergenten Reihe der Längen sämtlicher vollständiger Konvergenzintervalle, es ist also konvergent. Dann gilt aber $\lim S_{-n}(\omega)=0$, woraus zufolge der Ungleichungen $S_{-n}(\omega)\geqq S_{-n}(\sigma)>>0$ die Behauptung folgt.

Ist σ kein Konvergenzintervall, so können wir so vorgehen: Es gibt immer (unendlich viele vollständige) Konvergenzintervalle in jedem Intervall. Es sei ω ein in σ liegendes Konvergenzintervall. Wegen $\omega\subset\sigma$, $\sigma_{-n}\subset\Lambda$ und der Monotonität von $N(x)$ liegt in jedem σ_{-n} ein invers-iteriertes Intervall n-ter Ordnung von ω: $\omega_{-n}\subset\sigma_{-n}$. Wir wollen für diese Intervallpaare beweisen, daß die Zahlen

$$\frac{|\sigma_{-n}|}{|\omega_{-n}|} \qquad (n = 0, 1, 2, \dots)$$

eine beschränkte Menge bilden.

Selbstverständlich ist $1\leqq\frac{|\sigma_{-n}|}{|\omega_{-n}|}$. Die Bestimmung einer oberen Schranke ist umständlicher. Ist

$$(\sigma_{-(n+1)})_1=\sigma_{-n},$$

so gilt nach dem Lagrangeschen Mittelwertsatz

$$\frac{|\sigma_{-n}|}{|\sigma_{-(n+1)}|} = |N'(\bar{x})|, \qquad \bar{x}\in\sigma_{-(n+1)}. \tag{7}$$

Es ist $\sigma_{-(n+1)}$ ein Teilintervall eines Λ-Intervalls, und in diesem ist $|N'(x)| \geqq Q > 1$. (Siehe III. S. 202.) So folgt also

$$|\sigma_{-n}| \geqq Q |\sigma_{-(n+1)}|, \quad Q > 1.$$

Unter Anwendung dieser Ungleichung für $n = 0, 1, 2, \ldots$ erhalten wir:

$$|\sigma| \geqq Q^{n+1} |\sigma_{-(n+1)}|. \tag{8}$$

Aus (7) folgt jetzt

$$\frac{|\omega_{-n}|}{|\omega_{-(n+1)}|} = |N'(x')|, \quad x' \in \omega_{-(n+1)},$$

also

$$\frac{|\sigma_{-(n+1)}|}{|\omega_{-(n+1)}|} = \frac{|\sigma_{-n}|}{|\omega_{-n}|} \cdot \frac{|N'(x')|}{|N'(\bar{x})|}. \tag{9}$$

Hier ist

$$\left| \frac{N'(x')}{N'(\bar{x})} \right| \leqq \frac{\max |N'(x)|}{|N'(\bar{x})|}, \quad x \in \omega_{-(n+1)},$$

und wegen $\omega_{-(n+1)} \subset \sigma_{-(n+1)}$, gilt

$$\max_{x \in \omega_{-(n+1)}} |N'(x)| \leqq \max_{x \in \sigma_{-(n+1}} |N'(x)|,$$

also ist

$$\left| \frac{N'(x')}{N'(\bar{x})} \right| \leqq \frac{\max |N'(x)|}{|N'(\bar{x})|}, \quad x \in \sigma_{-(n+1)}.$$

Es folgt noch, wegen $|N'(\bar{x})| \geqq \min |N'(x)|$ $(x \in \sigma_{-(n+1)})$,

$$\left| \frac{N'(x')}{N'(\bar{x})} \right| \leqq \frac{\max |N'(x)|}{\min |N'(x)|}, \quad x \in \sigma_{-(n+1)}.$$

Schreiben wir für $x \in \sigma_{-(n-1)}$

$$\max |N'(x)| = M', \quad \min |N'(x)| = m',$$

so erhalten wir aus der vorigen Ungleichung auf Grund von $m' \geqq Q > 1$:

$$\left| \frac{N'(x')}{N'(\bar{x})} \right| \leqq \frac{M'}{m'} = 1 + \frac{M' - m'}{m'} < 1 + M' - m'. \tag{10}$$

Aus der Monotonität der Funktion $|N'(x)|$ für $x \in \sigma_{-(n+1)}$ (siehe III. S. 202.) ergibt sich, daß diese Funktion den größten und kleinsten Wert in den Grenzpunkten des Intervalls annimmt, d.h. durch nochmalige Anwendung des Lagrangeschen Satzes folgt

$$\frac{M' - m'}{|\sigma_{-(n+1)}|} = |N''(x'')|, \quad x'' \in \sigma_{-(n+1)},$$

und so erhält man aus (8)

$$M' - m' \leqq \frac{|\sigma|}{Q^{n+1}} |N''(x'')|, \quad x'' \in \sigma_{-(n+1)}. \tag{11}$$

Es ist aber in jedem Intervall σ_{-n}

$$|N''(x)| \leqq H,$$

deshalb gilt nach (11)

$$M' - m' \leqq \frac{R}{Q^{n+1}}, \qquad R = |\sigma| \cdot H,$$

und aus (9) und (10) erhalten wir:

$$\frac{|\sigma_{-(n+1)}|}{|\omega_{-(n+1)}|} \leqq \left(1 + \frac{R}{Q^{n+1}}\right) \frac{|\sigma_{-n}|}{|\omega_{-n}|}.$$

Schreiben wir diese Ungleichung für die nacheinander folgenden Zahlen n auf:

$$\frac{|\sigma_{-1}|}{|\omega_{-1}|} \leqq \left(1 + \frac{R}{Q}\right) \cdot \frac{|\sigma|}{|\omega|},$$

$$\frac{|\sigma_{-2}|}{|\omega_{-2}|} \leqq \left(1 + \frac{R}{Q^2}\right) \cdot \frac{|\sigma_{-1}|}{|\omega_{-1}|},$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\frac{|\sigma_{-n}|}{|\omega_{-n}|} \leqq \left(1 + \frac{R}{Q^n}\right) \cdot \frac{|\sigma_{-(n+1)}|}{|\omega_{-(n+1)}|},$$

so ergibt sich:

$$\frac{|\sigma_{-n}|}{|\omega_{-n}|} \leqq \left(1 + \frac{R}{Q}\right)\left(1 + \frac{R}{Q^2}\right) \cdots \left(1 + \frac{R}{Q^n}\right) \frac{|\sigma|}{|\omega|}.$$

Es ist aber das unendliche Produkt

$$P = \prod_{n=1}^{\infty} \left(1 + \frac{R}{Q^n}\right)$$

konvergent, und so bekommen wir für jedes n

$$\frac{|\sigma_{-n}|}{|\omega_{-n}|} < P \frac{|\sigma|}{|\omega|}.$$

Schreibt man dies in Form

$$|\sigma_{-n}| < P \frac{|\sigma|}{|\omega|} \cdot |\omega_{-n}|,$$

so folgt

$$0 < S_{-n}(\sigma) < P \frac{|\sigma|}{|\omega|} \cdot S_{-n}(\omega),$$

und wir erhalten — wegen der Konvergenz von $\sum_{n=0}^{\infty} S_{-n}(\omega)$ —

$$\lim_{n \to \infty} S_{-n}(\sigma) = 0.$$

Durch Anwendung dieser Gleichung auf die Intervalle

$$\sigma = \Lambda_2', \Lambda_2'', \Lambda_3', \Lambda_3'', \ldots, \Lambda_{k-1}', \Lambda_{k-1}''$$

folgt in (6)

$$\lim_{p\to\infty} S_{-p}(\Lambda) = 0,$$

d. h.

$$\sum_{n=0}^{\infty} S_{-n}(\Omega) = B_{k-1} - B_1.$$

Dies ist der Maß der Menge $\mathscr{K}$, der Menge der Konvergenzpunkte im Intervall $[B_1, B_{k-1}]$. Da es aber außer diesem Intervall keine Divergenzpunkte gibt, folgt der

Satz. *Die Menge der Divergenzpunkte ist eine Nullmenge.*

Literatur

B. Barna, Über die Divergenzpunkte des Newtonschen Verfahrens zur Bestimmung von Wurzeln algebraischer Gleichungen III, *Publ. Math. Debrecen,* **8** (1961), 193—207.

(Eingegangen am 23. April 1966.)