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

Von BÉLA BARNA (Debrecen)

Wir beschäftigen uns in dieser Arbeit — wie in den I. und II. Teilen — mit dem Newtonschen Näherungsalgorithmus angewandt auf Polynome mit reellen Koeffizienten und Nullstellen. Die auf reelle Veränderliche spezialisierten Ergebnisse der vorigen Arbeiten erreichen wir aber *mit elementaren Methoden* der reellen Analysis, ohne Bezugnahme auf die Iterationstheorie der rationalen komplexen Funktionen. Außerdem verallgemeinern wir die entsprechende Sätze von I. und II., indem wir *die Einfachheit der Nullstellen nicht fordern.*

§ 1.

Es sei $f(x)$ ein Polynom m-ten Grades mit reellen Koeffizienten, das nur reelle Nullstellen besitzt, und in dem der Koeffizient des höchsten Gliedes den Wert 1 hat. Wir bezeichnen mit A_i $(i=1,2,\dots,k)$ die verschiedenen Nullstellen des Polynoms, mit M_i $(\geqq 1)$ die entsprechende Multiplizität, und es sei

$$A_1 < A_2 < \cdots < A_k.$$

Dann gilt

$$f(x)=(x-A_1)^{M_1}\cdots(x-A_k)^{M_k}, \tag{1}$$

mit

$$M_1+M_2+\cdots+M_k=m, \tag{2}$$

und

$$f'(x)=(x-A_1)^{M_1-1}\cdots(x-A_k)^{M_k-1}\cdot\varphi(x),$$

wo das Polynom $\varphi(x)$ nur einfache Nullstellen hat, deren Anzahl nach (2)

$$m-1-(M_1-1+\cdots+M_k-1)=k-1$$

ist; diese (und nur diese) Punkte $B_i(i=1,2,\dots,k-1)$ sind die Extremalstellen von $f(x)$, welche keine Nullstellen sind. Dabei gelten die Ungleichungen:

$$A_1 < B_1 < A_2 < \cdots < B_{i-1} < A_i < B_i < \cdots < B_{k-1} < A_k.$$

Die Funktion $f(x)$ verhält sich links von B_1 und rechts von B_{k-1} streng monoton, und es liegt in jedem der Intervalle

$$(-\infty, B_1), (B_1, B_2), \ldots, (B_{i-1}, B_i), \ldots, (B_{k-1}, +\infty)$$

genau eine der verschiedenen Nullstellen A_i.

Wir bezeichnen mit C_j $(j = 1, 2, \ldots, l)$ diejenige Nullstellen von $f''(x)$ — etwa nach Größe geordnet — die von den Punkten A_i verschieden sind, d. h. die Punkte C_j (schlechthin: C-Punkte) entsprechen den Inflexionsabszissen der Kurve $y = f(x)$, für welche $f(C_j) \neq 0$ ausfällt. Auf die Lage dieser Punkte bezieht sich die folgende Überlegung.

Ist $M_1 \geqq 2$, so gilt $f'(A_1) = f'(B_1) = 0$, also liegt nach dem Rolleschen Satz zwischen A_1 und B_1 wenigstens ein C-Punkt; die Mindestzahl der Nullstellen von $f''(x)$ in dem Intervall $(-\infty, B_1)$ beträgt also, mit Multiplizität gerechnet, $(M_1 - 2) + 1 = M_1 - 1$, während im Fall $M_1 = 1$ diese Zahl gleich 0 ist. Es gibt also in diesem Intervall wenigstens $M_1 - 1$, und ähnlicherweise im Intervall $(B_{k-1}, +\infty)$ wenigstens $M_k - 1$ Nullstellen von $f''(x)$. Ist $M_i \geqq 2$ $(i \neq 1, k)$, so gibt es in beiden Intervallen

$$(B_{i-1}, A_i), (A_i, B_i)$$

wenigstens eine Nullstelle von $f''(x)$, also liegen in

$$(B_{i-1}, B_i) \tag{3}$$

wenigstens $(M_i - 2) + 2 = M_i$ Nullstellen. Im Fall $M_i = 1$ gibt es in (3) wenigstens eine Nullstelle. Es ist also M_i in jedem Fall die Mindestzahl der Nullstellen von $f''(x)$ in (3). Auf der ganzen Abszissenachse liegen also wenigstens

$$(M_1 - 1) + M_2 + \cdots + M_{k-1} + (M_k - 1)$$

Nullstellen von $f''(x)$. Da aber diese Summe nach (2) mit dem Grad des Polynoms $f''(x)$ übereinstimmt, kann $f''(x)$ *nicht mehrere* Nullstellen haben, und so folgt, daß es in dem Intervall $(-\infty, B_1)$ — und ähnlicherweise in $(B_{k-1}, +\infty)$ — entweder einen C-Punkt gibt, oder keinen (je nachdem, ob die dort liegende Nullstelle von $f(x)$ mehrfach, oder einfach ist); es folgt ferner, daß im Intervall (3) entweder zwei C-Punkte liegen (im Fall $M_i \geqq 2$), oder höchstens ein solcher Punkt (wenn nämlich $M_i = 1$ und damit $f''(A_i) \neq 0$ gilt).

§ 2.

Der aus dem beliebigen Punkt $x = x_0$ ausgehende Newtonsche Algorithmus entsteht durch die Iteration der *Grundfunktion*

$$N(x) = x - \frac{f(x)}{f'(x)}. \tag{4}$$

Die so erhaltene Folge

$$x_{n+1} = N(x_n) \quad (n = 0, 1, 2, \ldots)$$

ist die *Iterationsfolge*[1]) von x_0. Die *Iterierten* von $N(x)$ sind

$$N_0(x) = x,\ N_1(x) = N(x), \ldots, N_{n+1}(x) = N[N_n(x)].$$

Ist $N_\nu(\varkappa) = \varkappa$ und gilt $N_n(\varkappa) \neq \varkappa$ für $n = 1, 2, \ldots, \nu - 1$, so ist $\varkappa$ ein *Fixpunkt ν-ter Ordnung*. Die Fixpunkte erster Ordnung sind die Punkte A_i; wegen $f(A_i)/f'(A_i) = 0$ befriedigen nämlich die Zahlen A_i (und nur diese) die Gleichung

$$x - \frac{f(x)}{f'(x)} = x.$$

Im Falle eines Fixpuntes höherer Ordnung, d. h. wenn $\nu \geqq 2$ ist, sind die Punkte $\varkappa_0, \varkappa_1, \ldots, \varkappa_{\nu-1}$ verschiedene Fixpunte ν-ter Ordnung, und sie bilden einen ν-gliedrigen *Zyklus*. Die zu demselben Zyklus gehörigen Fixpunkte nennen wir *konjugierte* Fixpunkte. Die Iterationsfolge eines Fixpunktes enthält also nur endlich viele verschiedene Punkte.

Die Punkte x', deren Iterationsfolge den Punkt x_0 enthält, sind die *inversiterierten Punkte* des Punktes x_0; ist n die kleinste natürliche Zahl, für welche $N_n(x') = x_0$ gilt, so ist x' ein inversiterierter Punkt *n-ter Ordnung:* $x' = x_{-n}$.

Ist die Funktion $N(x)$ in dem Intervall $[a, b] \equiv \Delta x$ stetig, so bilden die iterierten Punkte $x_1 (x \in \Delta x)$ auch ein Intervall $(\Delta x)_1$, das *erste iterierte Intervall* von Δx. Dafür gilt

$$(\Delta x)_1 = [\min N(x),\ \max N(x)],\quad x \in \Delta x. \tag{5}$$

Wir definieren das *n-te iterierte Intervall* von Δx durch

$$(\Delta x)_n = ((\Delta x)_{n-1})_1.$$

Den Punkt x_0 nennen wir einen *Konvergenzpunkt*, bzw. *Divergenzpunkt* je nachdem ob die Iterationsfolge dieses Punktes konvergiert, oder divergiert.

---

[1] Die nach kleinen Buchstaben, nach Klammerausdrücken und nach dem Funktionszeichen N stehenden Indizes sind Iterationsindizes; nur die nach anderen großen Buchstaben stehenden Indizes und solche, die anderen Indizes angehängt sind, werden in dem gewöhnlichen, unterscheidenden Sinne verwendet.

Die Punkte x_n und $x_{-n}(n=1,2,\ldots)$ sind mit dem Punkt x_0 selbst Konvergenzpunkte, bzw. Divergenzpunkte.

Divergenzpunkte sind z. B. die Punkte B_i $(i=1,2,\ldots,k-1)$, und die Fixpunkte höherer Ordnung.

Gilt $\lim\limits_{n\to\infty} x_n=\xi$ und ist $N(x)$ im Punkt ξ stetig, so ist $N(\xi)=\xi$, d. h. ξ ist ein Fixpunkt erster Ordnung. Dies folgt aus der Gleichungen

$$\xi=\lim_{n\to\infty} x_{n+1}=\lim N(x_n)=N(\lim x_n)=N(\xi).$$

Es ist einfach zu beweisen, daß die Unstetigkeitspunkte von $N(x)$ d. h. die Punkte B_i $(i=1,2,\ldots,k-1)$ keine Grenzwerte von Iterationsfolgen darstellen können. Aus der Annahme $\lim x_n=B_i$ folgt nämlich

$$\lim (x_n-x_{n+1})=0,$$

also nach (4)

$$\lim \frac{f(x_n)}{f'(x_n)}=0;$$

dies ist aber unmöglich, weil $\lim f(x_n)=f(B_i)\neq 0$, und $\lim f'(x_n)=0$ ist. So folgt der

Satz 1. *Ist die Iterationsfolge* (x_n) $(n=0,1,2,\ldots)$ *konvergent, so ist ihr Grenzwert eine Nullstelle von* $f(x)$.

Ist x_0 ein Konvergenzpunkt und gilt $\lim x_n=\xi$, so sagen wir, daß x_0 zu dem Punkt ξ „gehört". Es gehört also jeder Konvergenzpunkt zu einem der Punkte A_i.

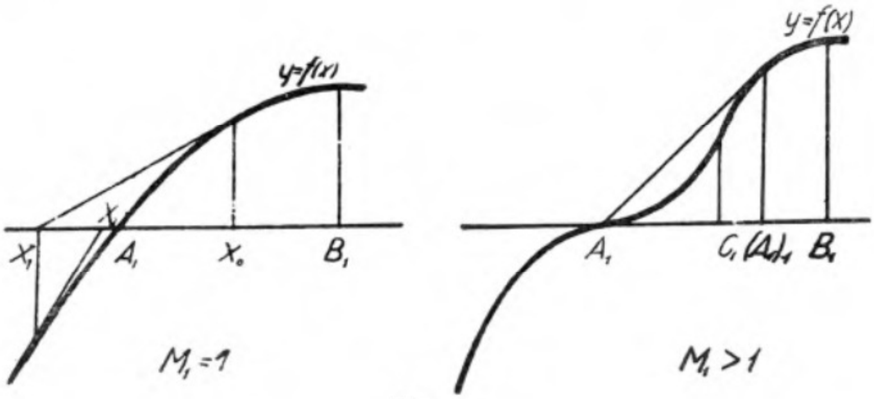

Fig. 1

Konvergenzpunkte sind z. B. die Punkte des Intervalls $(-\infty, B_1)$. Ist nämlich $M_1=1$, so ist die Iterationsfolge von x_0, mit $x_0\leqq A_1$, eine monoton wachsende Folge, und wegen der Konvexität der Kurve $y=f(x)$ für

$-\infty < x < B_1$, gilt $x_n \leqq A_1$. (S. Figur 1.) Es existiert also $\lim x_n$ und er hat — nach Satz 1. — den Wert A_1. Ist aber $A_1 < x_0 < B_1$, so gilt $x_1 < A_1$, und es folgt ebenfalls $\lim x_n = A_1$. Im Fall $M \geqq 2$ gibt es genau einen Punkt $(A_1)_{-1}$ mit

$$A_1 < (A_1)_{-1} < B_1,$$

und die Folge (x_n) strebt

für $x_0 \in (-\infty, A_1)$ monoton wachsend,

für $x_0 \in [A_1, (A_1)_{-1}]$ monoton abnehmend gegen A_1,

für $x_0 \in \big((A_1)_{-1}, B_1\big)$ ist $x_1 \in (-\infty, A_1)$ und so gilt ebenfalls $x_n \to A_1$.

Eine ganz analoge Erwägung zeigt, daß für jeden Punkt des Intervalls $(B_{k-1}, +\infty)$ $\lim x_n = A_k$ gilt. So erhalten wir: Die Iterationsfolgen des beliebigen Punkte x_0 der Intervalle $(-\infty, B_1)$ und $(B_{k-1}, +\infty)$ konvergieren nach A_1 bzw. nach A_k.[2])

Ist also $k=1$, so ist jeder Punkt Konvergenzpunkt, alle Punkte der Abszissenachse gehören zu A_1. *Es sei also im folgenden* $k > 1$.

Nun wollen wir die Funktion

$$N'(x) = \frac{f(x) f''(x)}{f'^2(x)} \tag{6}$$

in der Nähe der Punkte A_i $(i \neq 1, k)$ untersuchen.

Ist $M_i = 1$, dann gilt $N'(A_i) = 0$. Im Fall $M_i \geqq 2$ schreiben wir:

$$f(x) = (x - A_i)^{M_i} G(x),$$

dann ist — wie bekannt — $G(x)$ ein Polynom, mit $G(A_i) \neq 0$, und man kann $f'(x)$ in der Form

$$f'(x) = (x - A_i)^{M_i - 1} G_1(x)$$

darstellen, mit $G_1(A_i) = M_i \cdot G(A_i)$. Es folgt ferner:

$$f''(x) = (x - A_i)^{M_i - 2} G_2(x),$$

wo $G_2(A_i) = M_i \cdot (M_i - 1) \cdot G(A_i)$ gilt. Nach (6) ist also

$$N'(x) = \frac{G(x) G_2(x)}{G_1^2(x)},$$

d. h.

$$N'(A_i) = \frac{M_i - 1}{M_i}.$$

Jedenfalls gilt also

$$|N'(A_i)| < 1.$$

---

[2] Vgl. [3], S. 65—66.

Wegen der Stetigkeit von $N'(x)$ in A_i gibt es eine Umgebung

(7) $$(A_i - h, A_i + h) \qquad (h > 0)$$

dieses Punktes, in der $|N'(x)| \leqq q$ gilt, wo $q < 1$ ist. Wir bekommen für jeden Punkt dieser Umgebung — dem Lagrangeschen Mittelwertsatz zufolge —

$$|N(x_0) - N(A_i)| \leqq q\,|x_0 - A_i|,$$

d. h.

$$|x_1 - A_i| \leqq q\,|x_0 - A_i|,$$

also liegt auch x_1 im Intervall (7). Durch wiederholte Anwendung dieses Gedankenganges bekommen wir:

$$|x_n - A_i| \leqq q^n\,|x_0 - A_i|,$$

woraus für jeden Punkt x des Intervalles (7)

$$\lim x_n = A_i$$

folgt.

Das längste, den Punkt A_i einschließende Intervall, dessen jeder Punkt eine nach A_i konvergierende Iterationsfolge hat, nennen wir das *unmittelbare Konvergenzintervall* von A_i.

Um dies zu bestimmen nehmen wir die folgenden Eigenschaften der Funktion $N(x)$ in Betracht:

$$\lim_{x \to B_{i-1}+0} N(x) = +\infty, \quad \lim_{x \to B_i - 0} N(x) = -\infty;$$

ist $M_i = 1, f''(A_i) = 0$, so gibt es in dem Intervall

(8) $$(B_{i-1}, B_i)$$

keinen C-Punkt, und $N(x)$ nimmt monoton ab; ist $M_i = 1$, $f''(A_i) \neq 0$, so gibt es in (8) einen C-Punkt, und $N(x)$ nimmt in den Intervallen

$$(B_{i-1}, \min\{A_i, C_j\}) \; (\max\{A_i, C_j\}, B_i)$$

ab, während sie in

$$(\min\{A_i, C_j\}, \max\{A_i, C_j\})$$

monoton wächst; ist aber $M_i \geqq 2$, so gibt es zwei C-Punkte, und es gilt

$$B_{i-1} < C_j < A_i < C_{j+1} < B_i;$$

$N(x)$ ist in (B_{i-1}, C_j) und (C_{j+1}, B_i) monoton abnehmend, und in (C_j, C_{j+1}) monoton wachsend. (S. Figur 2.)

Aus diesen Eigenschaften folgt die Möglichkeit der folgenden Konstruktion: Wählen wir — etwa links von A_i — einen beliebigen Punkt x_0 in (7); dann liegt genau ein Punkt x_{-1} im Intervall (A_i, B_i). Bilden wir jetzt den ersten inversiterierten Punkt von x_{-1} in (B_{i-1}, A_i), er sei x_{-2}, und setzen wir diese Inversiteration fort. Dann liegen die so erhaltene Punkte

$$x_0, x_{-2}, x_{-4}, \dots \quad \text{bzw.} \quad x_{-1}, x_{-3}, x_{-5}, \dots$$

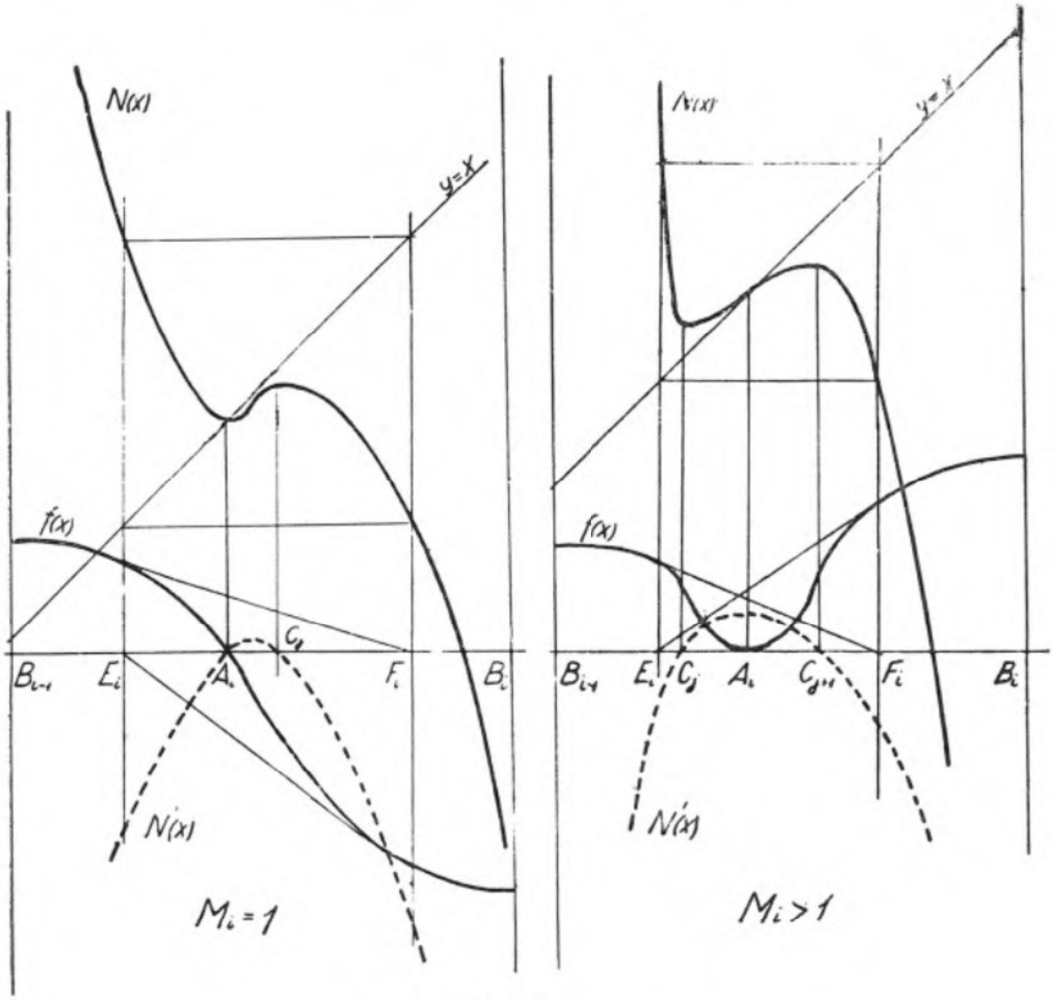

Fig. 2

auf entgegengesetzten Seiten von A_i, und bilden je eine, von diesem Punkt sich monoton entfernende Punktfolge zwischen B_{i-1} und B_i. Wir bezeichnen die Grenzwerte mit E_i, resp. F_i. Dann gilt

$$E_i < A_i < F_i$$

und die etwa vorhandenen C-Punkte liegen zwischen x_{-2} und x_{-1}, umsomehr zwischen E_i und F_i. Die Punkte E_i und F_i sind von x_0 unabhängig, für jeden Punkt x des Intervalles (E_i, F_i) gilt $x_n \in (E_i, F_i)$ und $x_n \to A_i$. Aus der

Stetigkeit der Funktion $N(x)$ folgert man noch durch die Gleichungen

$$\lim_{n\to\infty} x_{-2n} = \lim N(x_{-2n-1}) = N(\lim x_{-2n-1}),$$

daß $N(E_i) = F_i, N(F_i) = E_i$ gelten, d. h. E_i und F_i konjugierte Fixpunkte zweiter Ordnung, also Divergenzpunkte sind. Damit haben wir den Beweis dafür, daß *das unmittelbare Konvergenzintervall des Fixpunktes A_i das Intervall (E_i, F_i) ist.*

Da die Punkte B_i $(i = 1, 2, \ldots, k-1)$ Divergenzpunkte sind, so folgt aus den vorigen Überlegungen, daß *die folgenden Intervalle die sämtlichen unmittelbaren Konvergenzintervalle sind:*

$$(-\infty, B_1),\ (E_i, F_i)\ (i = 2, 3, \ldots, k-1),\ (B_{k-1}, +\infty).$$

Im Fall $k = 2$ ist B_1 der einzige Divergenzpunkt; alle anderen Punkte gehören zu A_1, bzw. zu A_2, je nachdem ob sie links, oder rechts von B_1 liegen.

Im Fall $k = 3$ sind außer den Punkten E_2, F_2 auch die Punkte $(B_1)_{-n}$ und $(B_2)_{-n} (n = 0, 1, 2, \ldots)$ Divergenzpunkte[3]), welche in den Intervallen $[B_1, E_2)$ und $(F_2, B_2]$ liegen. Es gibt also abzählbar unendlich viele verschiedene Divergenzpunkte. Jeder andere Punkt dieser Intervalle gehört zu A_1 oder zu A_3. Zu A_2 gehören nur die Punkte in (E_2, F_2).

Durch eine passende Anwendung der Inversiteration können wir — wie in II. S. 389. — beweisen, daß die Punkte E_i, F_i $(i = 2, 3, \ldots, k-1)$ Häufungspunkte der inversiterierten Punkte eines beliebigen Punktes sind.[4])

Auf die Menge sämtlicher Konvergenzpunkte bezieht sich der

Satz 2. *Die Konvergenzpunkte bilden eine offene Punktmenge.*

Beweis. Gehört der Punkt x_0 zur Wurzel A_i, so gibt es eine kleinste nichtnegative ganze Zahl p, für $x_p \in \Delta A_i$ gilt, wo ΔA_i das unmittelbare Konvergenzintervall des Fixpunktes A_i ist. Andererseits ist die Funktion $N_p(x)$ in dem Punkt x_0 stetig, dieser Punkt hat also eine Umgebung

$$(x_0 - \varepsilon, x_0 + \varepsilon) \qquad (\varepsilon > 0) \tag{9}$$

derart, daß für die Punkte x dieses Intervalls

$$N_p(x) = x_p \in \Delta A_i$$

gilt. Die Punkte von (9) gehören also zu A_i.

Nach einem bekannten Satz der Punktmengenlehre besteht also die Menge der Konvergenzpunkte aus den Punkten von gewissen offenen Intervallen. Eine ausführlichere Untersuchung dieser Intervalle lehrt, wie in I. § 3., daß

[3]) Siehe [4], S. 288.

[4]) Siehe [2], II. S. 389.

die Menge der Divergenzpunkte im Falle $k \geqq 4$ *die Mächtigkeit des Kontinuums hat.* Wir können die Divergenzpunkte ebenso klassifizieren, wie wir dies in I. § 4. getan haben.

§ 3.

Über die *Verteilung der Divergenzpunkte* gilt der

Satz 3. *In der Menge der Divergenzpunkte gibt es kein Intervall; diese Menge ist also punkthaft.*

Zum Beweis dieses Satzes wollen wir die Funktion $N'(x)$ näher betrachten.[5]) Man sieht sehr einfach, daß

$$\lim_{x \to B_i - 0} N'(x) = \lim_{x \to B_i + 0} N'(x) = -\infty \qquad (i = 1, 2, \ldots, k-1)$$

ist; so folgert man aus den Ungleichungen

$$N'(A_i) \geqq 0 \qquad (i = 1, 2, \ldots, k),$$

daß $N'(x)$ in jedem der Intervalle

(10) $$(A_1, B_1),\ (B_1, A_2),\ (A_2, B_2), \ldots, (B_{i-1}, A_i),\ (A_i, B_i), \ldots, (B_{k-1}, A_k)$$

jeden negativen Wert wenigstens einmal annimmt. Dies bedeutet, daß die Gleichung

(11) $$N'(x) = -a^2 \qquad (a \neq 0)$$

wenigstens $2k-2$ verschiedene Wurzeln haben muß. Sie kann aber auch *nicht mehr* haben. Um dies zu bestätigen, bezeichnen wir die verschiedenen mehrfachen Nullstellen mit

$$A_{i_1}, A_{i_2}, \ldots, A_{i_r};$$

dann ist die Anzahl der einfachen Nullstellen

$$s = m - (M_{i_1} + M_{i_2} + \cdots + M_{i_r})$$

und wir können das Polynom in der Form

$$f(x) = (x - A_{i_1})^{M_{i_1}} \cdots (x - A_{i_r})^{M_{i_r}} \cdot H(x)$$

schreiben, wo der Grad des Polynoms $H(x)$ gleich s ist; dann ist

$$f'(x) = (x - A_{i_1})^{M_{i_1}-1} \cdots (x - A_{i_r})^{M_{i_r}-1} \cdot \varphi(x)$$

wo der Grad des Polynoms $\varphi(x)$ gleich $s + r - 1$ ist, ferner ist

$$f''(x) = (x - A_{i_1})^{M_{i_1}-2} \cdots (x - A_{i_r})^{M_{i_r}-2} \cdot \psi(x)$$

---

[5]) Die hier folgende Methode ist lediglich die Verallgemeinerung des Beweisganges, den wir in [1] (§ 3, § 4) angewendet haben.

und der Grad von $\psi(x)$ beträgt $s+2r-2$, wo $H(x)$, $\varphi(x)$ und $\varphi(x)$, $\psi(x)$ relativ prim Polynompaare sind. Es ist also nach (6)

$$N'(x)=\frac{H(x)\cdot\psi(x)}{\varphi^2(x)}$$

und hier sind $H(x)\cdot\psi(x)$ und $\varphi^2(x)$ teilerfremde Polynome. Die Anzahl der Wurzeln von (11) kann also nicht größer sein, als der Grad der Gleichung

$$H(x)\cdot\psi(x)=-a^2\varphi^2(x),$$

d. h. als

$$s+(s+2r-2)=2(s+r-1);$$

es ist aber $s+r=k$, also $2(s+r-1)=2k-2$; genau so viele verschiedenen Wurzeln hat (11).

So folgt, daß die Funktion $N'(x)$ in jedem Teilintervall der Intervalle (10), wo sie negative Werte annimmt, eine *streng monotone Funktion ist*; insbesondere ist $N'(x)$ in $(B_{i-1}, E_i]$ negativ und monoton wachsend, in $[F_i, B_i)$ negativ und monoton abnehmend, woraus für $x\in(B_{i-1}, B_i), x\notin(E_i, F_i)$

$$|N'(x)|>\min_i\{|N'(E_i)|, |N'(F_i)|\}$$

folgt. Wir benötigen den folgenden

Hilfssatz: *Für jedes* $x\notin\Delta A_i$ $(i=2, 3, \ldots, k-1)$ *gilt* $|N'(x)|\geqq Q$ *mit* $Q>1$.

Es genügt — nach dem vorigen — zu beweisen, daß die Ungleichungen

$$N'(E_i)<-1 \qquad (i=2, 3, \ldots, k-1) \tag{12}$$

und

$$N'(F_i)<-1 \qquad (i=2, 3, \ldots, k-1) \tag{13}$$

gelten, wo die Relationen

$$E_i<A_i<F_i,$$

$$N(E_i)=F_i \tag{14}$$

$$N(F_i)=E_i \tag{15}$$

bestehen. Sodann kann $Q=\min\{|N'(E_i)|, N'(F_i)|\}$ gewählt werden.

Es sei

$$\Phi(x)=\frac{f'(x)}{f(x)},$$

dann ist

$$\Phi(x)=\frac{1}{x-N(x)}; \tag{16}$$

andererseits gilt die Gleichung

$$\Phi(x)=\sum_{j=1}^{k}\frac{M_j}{x-A_j}. \tag{17}$$

Aus (16) folgt nach (14) und (15)

$$\Phi(E_i)=\frac{1}{E_i-F_i}, \tag{18}$$

$$\Phi(F_i)=\frac{1}{F_i-E_i}, \tag{19}$$

und aus

$$N'(x)=\left(x-\frac{1}{\Phi(x)}\right)'=1+\frac{\Phi'(x)}{\Phi^2(x)}$$

folgert man, daß

$$1+\frac{\Phi'(E_i)}{(\Phi^2E_i)}<-1$$

mit der Ungleichung (12) äquivalent ist. Dies können wir nach (18) in der Form

$$1+\Phi'(E_i)(F_i-E_i)^2<-1,$$

d. h.

$$-\Phi'(E_i)>\frac{2}{(F_i-E_i)^2} \tag{20}$$

schreiben. Ähnlicherweise ist (13) mit

$$-\Phi'(F_i)>\frac{2}{(F_i-E_i)^2} \tag{21}$$

äquivalent. Nunmehr folgt aus (17)

$$-\Phi'(x)=\sum_{j=1}^{k}\frac{M_j}{(x-A_j)^2},$$

und so müssen wir nach (17), (18), (19) aus den Formeln

$$\frac{M_1}{A_1-E_i}+\cdots+\frac{M_i}{A_i-E_i}+\cdots+\frac{M_k}{A_k-E_i}=\frac{1}{F_i-E_i}, \tag{22}$$

$$\frac{M_1}{F_i-A_1}+\cdots+\frac{M_i}{F_i-A_i}+\cdots+\frac{M_k}{F_i-A_k}=\frac{1}{F_i-E_i} \tag{23}$$

die Ungleichungen

$$\frac{M_1}{(E_i-A_1)^2}+\cdots+\frac{M_i}{(E_i-A_i)^2}+\cdots+\frac{M_k}{(E_i-A_k)^2}>\frac{2}{(F_i-E_i)^2}, \tag{24}$$

$$\frac{M_1}{(F_i-A_1)^2}+\cdots+\frac{M_i}{(F_i-A_i)^2}+\cdots+\frac{M_k}{(F_i-A_k)^2}>\frac{2}{(F_i-E_i)^2} \tag{25}$$

herleiten, wo $2\leqq i\leqq k-1$ gilt. Um einfachere Formeln zu erhalten, führen wir die Zahlen (S. Figur 3)

$$D_j=\frac{E_i-A_j}{F_i-A_j}$$

ein, wodurch

$$A_j = \frac{E_i - F_i D_j}{1 - D_j}$$

wird. Dann entsprechen den $A_j, j \neq i$ positive Zahlen D_j, und es gilt $D_i < 0$; statt (22) und (23) erhalten wir

(26) $$\frac{M_1}{D_1} + \cdots + \frac{M_{i-1}}{D_{i-1}} + \frac{M_{i+1}}{D_{i+1}} + \cdots + \frac{M_k}{D_k} = m - 1 + \frac{M_i}{|D_i|},$$

bzw.

(27) $$M_1 D_1 + \cdots + M_{i-1} D_{i-1} + M_{i+1} D_{i+1} + \cdots + M_k D_k = m - 1 + M_i |D_i|,$$

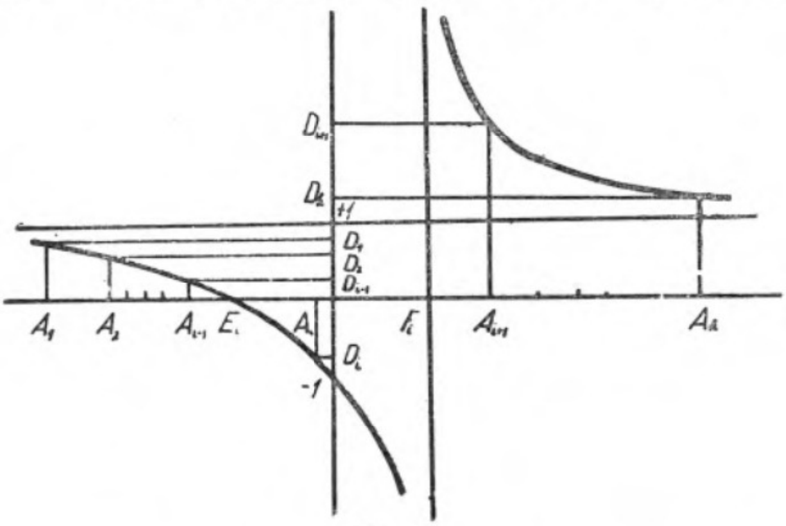

Fig. 3

während die Ungleichungen (24) und (25) — durch Anwendung von (26) und (27) — die Gestalt

(28) $$\sum_{j=1}^{k} \frac{M_i}{D_j^2} > m,$$

bzw.

(29) $$\sum_{j=1}^{k} M_j D_j^2 > m$$

annehmen. Wegen des paralellen Auftretens von D_j und $\frac{1}{D_j}$ genügt es aus (26) und (27) eine der Relationen (28) und (29) herzuleiten, z. B. die Ungleichung (29).

Dies geschieht sehr einfach mit Hilfe des folgenden Lemmas:

Lemma. *Sind m, p positive ganze Zahlen, $m-p \geqq 3$, S_q $(q = 1, 2, \ldots$ $\ldots, m-p)$ und T beliebige positive Zahlen, dann folgt aus den Gleichungen*

(30) $$S_1 + S_2 + \cdots + S_{m-p} = m - 1 + pT$$

und

$$\frac{1}{S_1}+\frac{1}{S_2}+\cdots+\frac{1}{S_{m-p}}=m-1+\frac{p}{T} \tag{31}$$

die Ungleichung

$$S_1^2+S_2^2+\cdots+S_{m-p}^2+pT^2>m. \tag{32}$$

Es ist nämlich[6]) im Fall

$$T<S_q \qquad (q=1,2,\ldots,m-p) \tag{33}$$

$$\sum_{q=1}^{m-p}(S_q-1)^2=\sum_{q=1}^{m-p}S_q^2-2\sum S_q+m-p=\sum_{q=1}^{m-p}S_q^2-2(m-1+pT)+m-p,$$

d. h.

$$\sum_{q=1}^{m-p}S_q^2=\sum_{q=1}^{m-p}(S_q-1)^2+2(m-1+pT)-m+p=$$

$$=\sum_{q=1}^{m-p}S_q\left(S_q-2+\frac{1}{S_q}\right)+m+2pT+p-2.$$

Hier wird jeder Summand — nach (33) und wegen $S_q-2+\frac{1}{S_q}\geqq 0$ — verkleinert, wenn man statt des Faktors S_q die Zahl T schreibt:

$$\sum_{q=1}^{m-p}S_q^2>\sum T\left(S_q-2+\frac{1}{S_q}\right)+m+2pT+p-2=$$

$$=T(m-1+pT)-2T(m-p)+T\left(m-1+\frac{p}{T}\right)+m+2pT+p-2=$$

$$=pT^2+(4p-2)T+m+2p-2>m-pT^2,$$

woraus (32) folgt.

Ist dagegen statt (33)

$$T\geqq \min_q\{S_q\}=S_1,$$

dann folgt — nach einer bekannten Ungleichung —

$$\sqrt{\frac{S_2^2+S_3^2+\cdots+S_{m-p}^2}{m-p-1}}\geqq\frac{S_2+S_3+\cdots+S_{m-p}}{m-p-1}=\frac{m-1+pT-S_1}{m-p-1},$$

[6]) Den Beweis des Lemmas (mit der unwesentlichen Beschränkung $p=1$) habe ich als eine Aufgabe deren Lösung mir (für $m\geqq 6$) unbekannt war, in *Mat. Lapok* **8** (1957), 134 gestellt. Den folgenden einfachen Beweis verdanke ich einer diesbezüglichen Mitteilung von Herrn J. Balázs, wofür ich ihm auch an dieser Stelle meinen herzlichen Dank aussprechen möchte. Seine urprüngliche Lösung ist inzwischen in *Mat. Lapok* **10** (1959), 332 erschienen. Lösungen wurden noch von J. Czipszer, Á. Fáy, R. O. Davies, E. Makai eingesandt.

also

$$S_2^2+S_3^2+\cdots+S_{m-p}^2 \geqq \frac{(m-1+pT-S_1)^2}{m-p-1} = \frac{(m-p-1+p+pT-S_1)^2}{m-p-1} =$$

$$= m-p-1+2(p+pT-S_1)+\frac{(p+pT-S_1)^2}{m-p-1} >$$

$$> m+p-1+2(pT-S_1) \geqq m > m-pT^2,$$

womit wir den Nachweis des Lemmas erbracht haben.

Nun erhält man den Beweis des Hilfssatzes, indem man in den Formeln (30) und (31) $T=-D_i$, $p=M_i$ setzt, und statt der Zahlen S_q — den Multiplizitäten M_j entsprechend — die Zahlen D_j schreibt. (Also ist $S_1=S_2=\cdots$ $\cdots=S_{M_{i-1}}=D_{i-1}$, usw.) Dann entstehen (26) und (27), ferner geht (32) in (29) über.

Beweis des Satzes 3. Es sei — entgegen der Behauptung — $\Delta x=[a,b]$ ein Divergenzintervall (also ein Intervall, dessen jeder Punkt Divergenzpunkt ist). Dasselbe kann nur in einem der Intervalle

$$(34) \qquad [B_{i-1}, E_i] \quad \text{oder} \quad [F_i, B_i] \qquad (i=2,3,\ldots,k-1)$$

liegen, die außerliegenden Punkte sind nämlich Konvergenzpunkte, und in jeder Nähe der Punkte B_i liegen auch Konvergenzpunkte. (Die Punkte des Intervalles $(B_i-\varepsilon, B_i)$ gehören für genügend kleines $\varepsilon>0$ zu A_1, und die des Intervalles $(B_i, B_i+\varepsilon)$ zu A_k. Es ist nämlich für $x<B_i$ $x_1<B_1$, und für $x>B_i$ $x_1>B_{k-1}$, wenn nur $|B_i-x|$ genügend klein ist.) So folgt, daß alle iterierten Intervalle $(\Delta x)_n$ — die ja auch Divergenzintervalle sind — in den Intervallen (34) liegen, wo $N(x)$ monoton abnimmt. Also ist (nach (5))

$$(\Delta x)_1=[N(b),\ N(a)]$$

und für die länge dieses Intervalls besteht laut des Lagrangeschen Mittelwertsatzes

$$|(\Delta x)_1|=|N'(\tau)|\cdot|a-b|=|N'(\tau)|\cdot|\Delta x|,\ \ \tau\in\Delta x,$$

wo — nach dem Hilfssatz — $|N'(\tau)|\geqq Q$, $Q>1$ ist, woraus

$$|(\Delta x)_1|\geqq Q|\Delta x|$$

folgt. Durch Iterieren entsteht für beliebiges n

$$|(\Delta x)_n|\geqq Q^n|\Delta x|.$$

Dies widerspricht aber der Beschränktheit der Intervalle (34), also ist die Existenz eines Divergenzintervalles unmöglich.

Literatur

[1] B. BARNA, Über das Newtonsche Verfahren zur Annäherung von Wurzeln algebraischer Gleichungen, *Publ. Math. Debrecen* **2** (1951), 40—63.

[2] B. BARNA, Über die Divergenzpunkte des Newtonschen Verfahrens zur Bestimmung von Wurzeln algebraischer Gleichungen I, II, *Publ. Math. Debrecen* **3** (1953), 109—118; **4** (1956), 384—397.

[3] M. BAUER, Az algebrai egyenletek valós gyökének meghatározása iterációval, *Math. és Phys. Lapok* **26** (1917), 57—66.

[4] A. RÉNYI, A Newton-féle gyökközelítő eljárásról (On Newton's method of approximation), *Mat. Lapok* **1** (1950), 278—293.

(Eingegangen am 1. Dezember 1960.)