ÜBER DIE ABSOLUTEN BETRÄGE DER WURZELN ALGEBRAISCHER GLEICHUNGEN.

VON

EDUARD BATSCHELET

in BASEL.

1. In einer Arbeit unter dem Titel »*Recherches sur la méthode de Graeffe et les zéros des polynômes et des séries de Laurent*» hat Herr A. OSTROWSKI den folgenden Satz bewiesen[1]:

Sei

$$A_0 x^n + A_1 x^{n-1} + \cdots + A_{n-1} x + A_n, \qquad A_0 \neq 0,$$

ein Polynom n-ten Grades mit beliebigen reellen oder komplexen Koeffizienten, dessen Wurzeln $x_1, \ldots, x_n$ *nach steigenden absoluten Beträgen geordnet sein mögen. Es gelte also*

$$|x_1| \leqq |x_2| \leqq \cdots \leqq |x_n|.$$

Variiert man die Argumente der Koeffizienten $A_0, \ldots, A_n$, *hält aber ihre absoluten Beträge fest, so bleibt* $|x_k|$ *für festes k in einem Intervall, dessen relative Breite höchstens gleich einer nur von n abhängigen Schranke* C_n *ist, für die*

$$C_n = 0{,}73 \cdot (n+1)^2 \tag{1}$$

gesetzt werden darf.

Unter der *relativen Breite* eines Intervalls $\langle a, b \rangle$, $0 < a \leqq b$, ist dabei das Verhältnis $\frac{b}{a}$ zu verstehen.

Im Folgenden bezeichnen wir mit c_n die (offenbar vorhandene) kleinste Schranke, für welche die Bedingungen des Satzes zutreffen. Dafür gilt $1 \leqq c_n \leqq C_n$.

---

[1] Acta math., Bd. 72 (1940), p. 145.

Herr Ostrowski hat weiterhin gezeigt, dass c_n für wachsende n beliebig gross wird, indem er die Beziehung

$$\underline{\lim}\ \frac{c_n}{n} \geqq \frac{4}{\pi} = 1{,}273\ldots$$

herleitete[1]. In meiner Dissertation[2] habe ich die etwas schärfere Beziehung

$$\underline{\lim}\ \frac{c_n}{n} \geqq \frac{1}{\lg 2} = 1{,}442\ldots$$

bewiesen.

In der vorliegenden Mitteilung leiten wir die Ungleichung

$$c_n > \frac{1}{4}n^2 \tag{2}$$

her, woraus sich wegen (1) ergibt, dass c_n *die Grössenordnung von* n^2 *besitzt.* Dieses Ergebnis ist insofern bemerkenswert, als Herr Ostrowski bei der Herleitung der oberen Schranke (1) für c_n scheinbar sehr starke Vernachlässigungen gemacht hat[3].

2. Dem Beweis von (2) schicken wir zwei Hilfssätze voraus.

Hilfssatz I. *Es sei k eine positive ganze Zahl und $F(\xi_1, \ldots, \xi_k)$ ein multilineares Polynom der Variabeln $\xi_1, \ldots, \xi_k$, das also in Bezug auf jede Variable linear ist. Das konstante Glied sei gleich* 0, *die übrigen Koeffizienten aber irgendwelche reelle Zahlen, die nicht sämtlich verschwinden.*

Dann kann man jeder der Variabeln $\xi_1, \ldots, \xi_k$ die Werte $+1$ *oder* -1 *so beilegen, dass*

$$F(\xi_1, \ldots, \xi_k) < 0$$

gilt.

---

[1] l. c., p. 146, Ungleichung (26, 1).

[2] »*Untersuchungen über die absoluten Beträge der Wurzeln algebraischer, insbesondere kubischer Gleichungen*«, Verh. d. Naturf. Ges. in Basel, Bd. LV (1944).

[3] Es möge $c_n^{(k)}$ für festes k, $1 \leqq k \leqq n$, die kleinste, nur von n und k abhängige Schranke für die relative Breite des Intervalls bezeichnen, in dem $|x_k|$ variiert. Aus den Untersuchungen von Herrn Ostrowski, l. c. Formel (25, 3) auf p. 145, folgt $c_n^{(k)} \leqq 2{,}92 \cdot k(n-k+1)$.

In einer Arbeit, die an anderer Stelle veröffentlicht wird, erhalten wir durch Abschätzung nach unten $c_n^{(k)} > k(n-k+1)$. Die Herleitung dieses Ergebnisses, von dem die obige Ungleichung (2) ein Spezialfall ist, geschieht auf einem ähnlichen Wege wie in der vorliegenden Arbeit, erfordert aber eine mehr in das Einzelne gehende Diskussion.

Beweis. Die Behauptung ist für $k = 1$ offenbar richtig, da sich dann F auf $R \cdot \xi_1$ reduziert, wo R irgend eine reelle Zahl $\neq 0$ ist. Für $k > 1$ führen wir den Beweis durch vollständige Induktion. Wir können $F(\xi_1, \ldots, \xi_k)$ in der Gestalt

$$\xi_k \cdot F_1(\xi_1, \ldots, \xi_{k-1}) + F_2(\xi_1, \ldots, \xi_{k-1})$$

schreiben, wo F_1 und F_2 multilineare Polynome der Variabeln $\xi_1, \ldots, \xi_{k-1}$ sind, die nicht beide identisch verschwinden. Bei F_2 fehlt ausserdem das konstante Glied.

Der Hilfssatz möge bereits für den Index $k-1$ bewiesen sein. Wenn $F_2 \not\equiv 0$ ist, so erfüllt F_2 die Bedingungen des Hilfssatzes, und wir können daher den Variablen $\xi_1, \ldots, \xi_{k-1}$ die Werte ± 1 so beilegen, dass $F_2 < 0$ wird. Setzen wir dann $\xi_k = +1$, wenn bei der getroffenen Wahl der Variabeln $F_1 \leqq 0$ ist, oder $\xi_k = -1$, wenn $F_1 > 0$ ist, so folgt unmittelbar $F(\xi_1, \ldots, \xi_k) < 0$.

Ist aber $F_2 \equiv 0$, so legen wir den Variabeln $\xi_1, \ldots, \xi_{k-1}$ die Werte ± 1 derart bei, dass $F_1 \neq 0$ wird, und geben ξ_k wiederum das zu demjenigen von F_1 entgegengesetzte Vorzeichen. Dann folgt die Behauptung auch für diesen Fall unmittelbar.

3. Hilfssatz II. *Es sei $n > 0$ eine ungerade Zahl und $m = \frac{n+1}{2}$ gesetzt. Dann gibt es stets eine algebraische Gleichung n-ten Grades*

$$A_0 x^n + A_1 x^{n-1} + \cdots + A_n = 0, \tag{3}$$

deren Koeffizienten reell sind und den Bedingungen

$$A_{n-\nu} = \pm A_\nu, \qquad \nu = 0, \ldots, m-1, \tag{4}$$

genügen, und die eine m-fache Wurzel $x = t$ besitzt, welche positiv ist und die Ungleichung

$$t > m \tag{5}$$

erfüllt.

Beweis. Eine Gleichung n-ten Grades mit einer m-fachen Wurzel t kann wegen $n = 2m - 1$ in der Gestalt

$$(x-t)^m \cdot (a_1 x^{m-1} + a_2 x^{m-2} + \cdots + a_m) = 0$$

angesetzt werden. Der Abkürzung halber setze man

$$T_\mu = (-1)^\mu \binom{m}{\mu} t^\mu, \qquad \mu = 0, 1, \ldots, m. \tag{6}$$

So erhält die Gleichung die Gestalt

$$(T_0 x^m + T_1 x^{m-1} + \cdots + T_m) \cdot (a_1 x^{m-1} + a_2 x^{m-2} + \cdots + a_m) \equiv$$
$$\equiv T_0 a_1 x^{2m-1} + (T_1 a_1 + T_0 a_2) x^{2m-2} + \cdots + T_m a_m = 0.$$

Der Koeffizient der Potenz $x^{n-\nu} = x^{(2m-1)-\nu}$ lautet

$$A_\nu = \sum_{\mu=1}^{m} T_{\nu-\mu+1} a_\mu,$$

wobei $T_{\nu-\mu+1} = 0$ zu setzen ist, wenn der Index negativ oder grösser als m ist.

Mit den so dargestellten Koeffizienten erfülle man jetzt die Bedingungen (4) durch den Ansatz

$$(7) \qquad A_{n-\nu} + \xi_\nu A_\nu \equiv \sum_{\mu=1}^{m} T_{n-\nu-\mu+1} a_\mu + \xi_\nu \cdot \sum_{\mu=1}^{m} T_{\nu-\mu+1} a_\mu = 0, \qquad \nu = 0, \ldots, m-1,$$

wo $\xi_0, \xi_1, \ldots, \xi_{m-1}$ Grössen bezeichnen, die nur der Werte $+1$ oder -1 fähig sind. Der Hilfssatz II ist bewiesen, wenn es gelingt, für $a_1, \ldots, a_m$ reelle Zahlen zu finden derart, dass die Gleichungen (7) für geeignete $\xi_\mu = \pm 1$, $\mu = 0, \ldots, m-1$, und für einen geeigneten Wert $t > m$ erfüllt sind.

Nun lauten die Gleichungen (7), nach den Unbekannten $a_1, \ldots, a_m$ geordnet,

$$\sum_{\mu=1}^{m} (T_{n-\nu-\mu+1} + \xi_\nu T_{\nu-\mu+1}) a_\mu = 0, \qquad \nu = 0, \ldots, m-1,$$

oder, wenn man sie, mit $\nu = m-1$ beginnend, ausführlich schreibt:

$$\begin{array}{l}
(T_m + \xi_{m-1} T_{m-1}) a_1 + (T_{m-1} + \xi_{m-1} T_{m-2}) a_2 + \cdots \\
\qquad\qquad + (T_2 + \xi_{m-1} T_1) a_{m-1} + (T_1 + \xi_{m-1} T_0) a_m = 0 \\
\xi_{m-2} T_{m-2} a_1 + (T_m + \xi_{m-2} T_{m-3}) a_2 + \cdots + (T_3 + \xi_{m-2} T_0) a_{m-1} + T_2 \cdot a_m = 0 \\
\xi_{m-3} T_{m-3} a_1 + \xi_{m-3} T_{m-4} a_2 + \cdots + T_4 \cdot a_{m-1} + T_3 \cdot a_m = 0 \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
\xi_1 \cdot T_1 \cdot a_1 + \xi_1 \cdot T_0 \cdot a_2 + T_m \cdot a_{m-1} + T_{m-1} \cdot a_m = 0 \\
\xi_0 \cdot T_0 \cdot a_1 \qquad\qquad\qquad\qquad + T_m \cdot a_m = 0
\end{array}$$

Dies sind m homogene lineare Gleichungen für die m Unbekannten $a_1, \ldots, a_m$ Notwendig und hinreichend für die Existenz einer nicht trivialen Lösung ist das Verschwinden der Determinante des Systems. Wenn diese mit $D(t)$ bezeichnet wird, so hat man demnach

$$(8)\quad D(t) = \begin{vmatrix} T_m + \xi_{m-1} T_{m-1} & T_{m-1} + \xi_{m-1} T_{m-2} & \dots & T_2 + \xi_{m-1} T_1 & T_1 + \xi_{m-1} T_0 \\ \xi_{m-2} T_{m-2} & T_m + \xi_{m-2} T_{m-3} & \dots & T_3 + \xi_{m-2} T_0 & T_2 \\ \xi_{m-3} T_{m-3} & \xi_{m-3} T_{m-4} & \dots & T_4 & T_3 \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \xi_1 T_1 & \xi_1 T_0 & \dots & T_m & T_{m-1} \\ \xi_0 T_0 & 0 & \dots & 0 & T_m \end{vmatrix} = 0.$$

Die Matrix der Determinante (8) ist die Summe der beiden Matrizen

$$\begin{pmatrix} T_m & T_{m-1} & T_{m-2} & \dots & T_2 & T_1 \\ 0 & T_m & T_{m-1} & \dots & T_3 & T_2 \\ 0 & 0 & T_m & \dots & T_4 & T_3 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & T_m & T_{m-1} \\ 0 & 0 & 0 & \dots & 0 & T_m \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} \xi_{m-1} T_{m-1} & \xi_{m-1} T_{m-2} & \dots & \xi_{m-1} T_1 & \xi_{m-1} T_0 \\ \xi_{m-2} T_{m-2} & \xi_{m-2} T_{m-3} & \dots & \xi_{m-2} T_0 & 0 \\ \xi_{m-3} T_{m-3} & \xi_{m-3} T_{m-4} & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \xi_1 T_1 & \xi_1 T_0 & \dots & 0 & 0 \\ \xi_0 T_0 & 0 & \dots & 0 & 0 \end{pmatrix}.$$

Von 0 verschiedene Elemente kommen nur auf einer Seite der Haupt-, resp. der Nebendiagonalen vor.

Wegen (6) ist $D(t)$ ein Polynom vom Grade m^2 in t. Die Entwicklung nach fallenden Potenzen von t beginnt mit zwei Gliedern, die sich aus den Elementen der Hauptdiagonalen ergeben:

$$D(t) \equiv T_m^m + \xi_{m-1} T_m^{m-1} T_{m-1} \pm \cdots = 0,$$

woraus sich

$$D^*(t) \equiv (-1)^{m^2} \cdot D(t) \equiv t^{m^2} - m\,\xi_{m-1}\, t^{m^2-1} \pm \cdots = 0$$

ergibt. Man wähle von vornherein $\xi_{m-1} = 1$, womit (8) in die Gleichung

$$(9)\qquad D^*(t) \equiv t^{m^2} - m\, t^{m^2-1} + R(t) = 0$$

übergeht. $R(t)$ fasst hier die Glieder vom Grade $\leqq m^2 - 2$ zusammen. Offenbar gilt

$$(10)\qquad D^*(m) = R(m).$$

Man beachte weiter, dass bei der Berechnung der Determinante durch Bildung von Produkten aus m Elementen, von denen keine zwei in der gleichen Zeile oder in der gleichen Kolonne vorkommen, nur ein einziges Produkt entsteht, das keinen der Faktoren ξ_μ enthält, nämlich T_m^m, und ebenso nur ein einziges, welches bloss den Faktor ξ_{m-1} enthält, nämlich $\xi_{m-1} T_m^{m-1} T_{m-1}$. Das sind gerade die beiden Anfangsglieder der Entwicklung, die zu (9) geführt hat. Alle übrigen Produkte und daher auch sämtliche Glieder von $R(t)$ sind mit einem oder mit mehreren der Faktoren $\xi_0, \xi_1, \ldots, \xi_{m-2}$ behaftet.

Da ξ_μ für jedes feste μ nur in *einer* Zeile der Determinante (8) auftritt, so ist $R(t)$ somit ein multilineares Polynom der Grössen $\xi_0, \ldots, \xi_{m-2}$ ohne konstantes Glied. Anderseits können nicht sämtliche Glieder verschwinden, da das Produkt $\xi_0 \xi_1 \ldots \xi_{m-1}$ offenbar den (von t unabhängigen) Koeffizienten $T_0^m = 1$ besitzt.

$R(t)$ erfüllt damit die Bedingungen des Hilfssatzes I. Es gelingt folglich, für jedes einzelne ξ_μ, $\mu = 0, \ldots, m-2$ die Werte $+1$ oder -1 so zu bestimmen, dass $R(m) < 0$, oder dass vermöge (10)

$$D^*(m) < 0 \tag{11}$$

wird. Bei dieser Wahl der ξ_μ folgt, dass *$D^*(t)$ eine positive Wurzel $t > m$ besitzt*, da in (9) das Glied höchsten Grades in t einen positiven Koeffizienten besitzt.

Die Existenz der Gleichung (3) mit den Bedingungen (4) und (5) ist damit gesichert, womit der Hilfssatz II bewiesen ist.

4. Wir können nunmehr die Ungleichung (2) beweisen. Es möge $n > 0$ zunächst *ungerade* sein. Die Gleichung

$$A_0 x^n + A_1 x^{n-1} + \cdots + A_n = 0 \tag{12}$$

genüge den Bedingungen des Hilfssatzes II. Sie besitzt also eine m-fache Wurzel $t > m$, wo $m = \frac{n+1}{2}$ gesetzt ist.

Die Wurzeln $x_1, \ldots, x_n$ seien wie in Abschnitt 1. nach steigenden absoluten Beträgen geordnet. Wegen $m = \frac{n+1}{2}$ können wir x_m die mittlere Wurzel von (12) nennen. Die Gleichung kann höchstens $m-1$ Wurzeln besitzen, deren absoluter Betrag $> |x_m|$, und ebenso nur $m-1$ Wurzeln, deren absoluter Betrag $< |x_m|$ ist. Da nun t m-fache Wurzel ist, so muss

$$x_m = t$$

gelten.

Die aus (12) durch die Transformation $x = \frac{1}{y}$ hervorgehende Gleichung

$$A_n x^n + A_{n-1} x^{n-1} + \cdots + A_0 = 0 \tag{13}$$

besitzt offenbar die m-fache Wurzel $x_m^{-1} < m^{-1}$. Zugleich ist x_m^{-1} der Wert der mittleren Wurzel der Gleichung (13).

Wegen (4) gehört die Gleichung (13) zu der Schar der Gleichungen, die aus (12) entstehen, wenn man die Argumente der Koeffizienten beliebig variiert. Der absolute Betrag der mittleren Wurzel schwankt somit in einem Intervall, dessen relative Breite $> \frac{m}{m^{-1}} = m^2$ ist.

Nach Definition von c_n gewinnen wir daraus die Ungleichung

$$c_n > m^2 = \frac{(n+1)^2}{4}, \tag{14}$$

woraus (2) für ungerade n unmittelbar folgt.

Es sei endlich n *gerade* und gleich $n_1 + 1$ gesetzt. Bezeichnen wir mit $e_0, e_1, \ldots, e_{n_1}$ komplexe Variable vom absoluten Betrage 1, so existiert nach dem eben Bewiesenen eine Schar von Gleichungen

$$e_0 A_0 x^{n_1} + e_1 A_1 x^{n_1 - 1} + \cdots + e_{n_1} A_{n_1} = 0$$

mit festen A_ν, welche die folgende Eigenschaft besitzt: Unter den nach steigenden absoluten Beträgen geordneten Wurzeln gibt es wenigstens eine, deren absoluter Betrag in einem Intervall schwankt, das eine relative Breite $> \frac{(n_1+1)^2}{4} = \frac{1}{4} n^2$ hat. Dasselbe muss offenbar von der Schar der Gleichungen n-ten Grades gelten, die man aus der vorigen durch Multiplikation mit x erhält, nämlich von der Gleichungsschar $e_0 A_0 x^n + e_1 A_1 x^{n-1} + \cdots + e_{n_1} A_{n_1} x = 0$. Damit ist (2) auch für gerade n bewiesen.

5. Zum Schluss möge noch das Beispiel $n = 3$ betrachtet werden. Die Gleichung (9) lautet hier für die in Abschnitt 3. erklärte Wahl von ξ_0 und ξ_1

$$t^4 - 2t^3 - 2t + 1 = 0.$$

Sie ist reziprok und besitzt die positive Wurzel

$$t_0 = 2{,}296 \ldots.$$

Es existiert also eine Schar von kubischen Gleichungen, die sich nur in den Argumenten der Koeffizienten unterscheiden, für welche der absolute Betrag der mittleren Wurzel in einem Intervall variiert, dessen relative Breite $\geqq \frac{t_0}{t_0^{-1}} = t_0^2$ ist. Es folgt daraus $c_3 \geqq t_0^2$. An anderer Stelle[1] habe ich gezeigt, dass auch $c_3 \leqq t_0^2$ und somit

$$c_3 = t_0^2 = 5{,}274 \ldots$$

gilt.

[1] S. meine in Fussnote 2 auf Seite 254 zitierte Arbeit.