

1903, 3850

Ueber

Tchebychefsche Annäherungsmethoden.

Inaugural-Dissertation

zur

Erlangung der Doktorwürde

der

Hohen Philosophischen Fakultät der Georg-Augusts-Universität

zu Göttingen

vorgelegt von

Paul Kirchberger

aus Weilburg a./Lahn.

Göttingen 1902.

Druck der Dieterich'schen Univ.-Buchdruckerei.

(W. Fr. Kaestner).

Referent: Herr Professor Hilbert.

Tag der mündlichen Prüfung: 17. Juli 1902.

Meinen lieben Eltern.

Einleitung.

Mit dem Begriff der Funktion ist das Postulat der numerischen Berechnung der Funktionswerte für irgendwelche Werte der unabhängigen Variablen gegeben. Da aber die vier elementaren Spezies der Addition, Subtraktion, Multiplikation und Division, oder streng genommen nur die ersten drei derselben, die einzigen numerisch ausführbaren Rechnungsarten, alle andern aber nur insoweit durchführbar sind, als sie sich auf diese zurückführen lassen, so folgt hieraus, dass wir sämtliche Funktionen nur insoweit numerisch beherrschen, als sie sich durch rationale Funktionen ersetzen, d. h. angenähert darstellen lassen. Hieraus erhellt die grosse Bedeutung der Annäherungsprobleme für die gesamte Mathematik und die ausgezeichnete Stellung, die die Probleme der Annäherung durch rationale oder ganze rationale Funktionen einnehmen. In der That setzt, wenigstens für die numerische Berechnung, jede Annäherung durch andere, z. B. trigonometrische Funktionen, die annäherungsweise Ersetzbarkeit dieser Funktionen durch rationale voraus.

Die sämtlichen Probleme der Annäherung können wir von verschiedenem Gesichtspunkt aus einteilen. Die verschiedene Natur der anzunähernden Funktion, die entweder stetig sein, oder auch aus diskreten Wertepaaren bestehen kann, lässt die eigentlichen Annäherungsprobleme und die Interpolationsprobleme unterscheiden. Ferner können wir einteilen nach der Art der annähernden Funktion, die ein Polynom, eine rationale oder eine transcendente Funktion sein kann. Vor allem aber können wir nach der Natur der Annäherung selber, d. h. nach dem Gesichtspunkt, von dem aus die gesuchte Funktion als die am besten annähernde betrachtet werden soll, verschiedene Gruppen von Problemen unterscheiden. Es kann ein bestimmter Punkt bevorzugt und hier z. B. in der Uebereinstimmung möglichst vieler Ableitungen das

Kriterium der besten Annäherung gesehen werden, wie dies bei den abgebrochenen Potenzreihen und Kettenbrüchen der Fall ist. Es kann aber auch ein ganzes Intervall gleichberechtigt erscheinen. Dann können entweder alle gemachten Fehler gleichmässig berücksichtigt werden, indem verlangt wird, dass die Summe ihrer absoluten Beträge, die Summe ihrer Quadrate u. s. w. möglichst klein sei. Dies führt, wenn die anzunähernde Funktion stetig ist, auf ein Problem der Variationsrechnung. Oder aber es kann gefordert werden, dass der grösste Fehler, der bei der Ersetzung der anzunähernden Funktion durch die annähernde begangen wird, möglichst klein sei.

Dies letzte Kriterium wurde zuerst von Poncelet aufgestellt und von Tchebychef systematisch ausgearbeitet.

Vergleichen wir diese Methode mit der gewöhnlichen, nach der Potenzreihen nach einer endlichen Anzahl Glieder abgebrochen werden, so leuchtet sofort ein Vorzug der Tchebychef'schen Methode ein: Jede Annäherung hat nur Sinn in einem endlichen Intervall, und in einem solchen erfüllen die abgebrochenen Potenzreihen keine Minimalbedingung. Die Tchebychef'sche Methode hat dagegen den Nachteil, dass bei ihr der Grad der annähernden Funktion fest vorgeschrieben sein muss, es also nicht, wie bei den Potenzreihen, möglich ist, durch Erhöhung des Grades mit Benutzung des vorangegangenen Resultates eine Besserung der Annäherung zu erzielen.

Zweck der vorliegenden Arbeit ist es, zunächst die Tchebychef'schen Resultate darzustellen; hierbei werden einige Fragen zu erledigen sein, die Tchebychef bei der Abfassung seiner Arbeiten, die um etwa 50 Jahre zurückliegt, noch fernliegen mussten, wie namentlich die Frage nach der Existenz der annähernden Funktion und der analytischen Natur ihrer Koeffizienten. Sodann wird prinzipiell nicht nur mit Tchebychef die Zahl, sondern auch das Vorzeichen der Fehler zu berücksichtigen sein. Schliesslich soll die Theorie auch auf Funktionen zweier Variablen übertragen werden.

Kapitel I. Allgemeine Theorie bei einer reellen Veränderlichen.

§ 1. Definition.

Es sei $\varphi(x)$ die gegebene, angenähert darzustellende Funktion der reellen Variablen x . Ferner sei ein Intervall gegeben, als

das wir in der Regel $-1 \dots +1$ annehmen wollen, und n sei der gegebene Grad der gesuchten ganzen rationalen Funktion. Ist

$$f_n(x) = p_n x^n + p_{n-1} x^{n-1} + \dots + p_1 x + p_0$$

irgend eine ganze rationale Funktion n ten Grades, so hat die Differenz

$$f_n(x) - \varphi(x),$$

absolut genommen im gegebenen Intervall ein Maximum. Dies Maximum sei mit L bezeichnet und Abweichung genannt. Wir suchen diejenige Funktion $f_n(x)$, für die L ein Minimum wird. Sie heisse die „Annäherungsfunktion“ von $\varphi(x)$ im Intervall $-1 \dots +1$.

§ 2. Existenzbeweis.

L ist eine Funktion der Parameter $p_0, p_1 \dots p_n$.

$$L = L(p_0, p_1 \dots p_n),$$

und es handelt sich um das Minimum dieser Funktion. Nach dem Weierstrassschen Fundamentalsatz hat eine stetige Funktion im endlichen Bereich ein Minimum. Wir wollen daher nachweisen, dass wir uns auf ein endliches Gebiet beschränken können, und zu diesem Zweck beweisen wir folgenden

Algebraischen Hilfssatz: Es sei für eine ganze rationale Funktion der Grad n und ein endliches Intervall der reellen Variablen vorgeschrieben. Nach Annahme einer endlichen, aber beliebig grossen Zahl A lassen sich den $(n+1)$ Koeffizienten unserer Funktion Intervalle zuordnen, derart dass, wenn auch nur ein Koeffizient aus seinem Intervall heraustrit, ohne dass über das Verhalten der übrigen Koeffizienten etwas bekannt zu sein braucht, die Funktion an mindestens einer Stelle des Intervalls einen Wert annimmt, der dem absoluten Betrage nach grösser ist als A .

Denkt man sich jede ganze rationale Funktion n ten Grades als Punkt eines $(n+1)$ dimensionalen Raumes dargestellt, und sei M der grösste Wert, den eine Funktion im vorgeschriebenen Intervall annimmt, so lautet unser Satz:

Ausserhalb eines $(n + 1)$ dimensionalen Prismas ist M dem absoluten Betrage nach überall grösser als A . Dies Prisma kann nur von dem Intervall, dem Grad und von A abhängen.

Die Funktion sei

$$p_n x^n + p_{n-1} x^{n-1} + \dots + p_1 x + p_0,$$

das Intervall

$$a - h \dots a + h.$$

Wir transformieren

$$\bar{x} = \frac{(x - a)(2n + 2)}{h},$$

wodurch

$$\bar{f}(\bar{x}) = \bar{p}_n \bar{x}^n + \bar{p}_{n-1} \bar{x}^{n-1} + \dots + \bar{p}_0 = f(x).$$

Die p sind einfache Funktionen der \bar{p} , und es lassen sich die den p zugehörigen Intervalle leicht berechnen, wenn die der \bar{p} bekannt sind. Das Intervall für \bar{x} ist $-(2n + 2) \dots + (2n + 2)$. Wir schreiben nun wieder x, p, f statt $\bar{x}, \bar{p}, \bar{f}$. Da wir zur Betrachtung der Wurzeln übergehen wollen, betrachten wir nicht $f(x)$, sondern $f(x) + x^{n+1}$; ich behaupte, für diese Funktion können als die gesuchten Intervalle genommen werden:

$$\begin{aligned} |p_n| &\leq f_1^{(n+1)}(A) \\ |p_{n-1}| &\leq f_2^{(n+1)}(A) \\ &\vdots \\ |p_1| &\leq f_n^{(n+1)}(A) \\ |p_0| &\leq f_{n+1}^{(n+1)}(A) \end{aligned}$$

wo unter $f^{(n+1)}$ die elementaren symmetrischen Funktionen von $(n + 1)$ Variablen verstanden sind, und die Werte aller dieser Variablen gleich A gesetzt werden sollen, also

$$f_1^{(n+1)}(A) = A + A + \dots + A = (n + 1) A$$

$$f_2^{(n+1)}(A) = AA + AA + \dots + AA = \frac{n(n + 1)}{2} A^2$$

\vdots

$$f_{n+1}^{(n+1)}(A) = AA \dots = A^{n+1}$$

d. h. die $f^{(n+1)}$ sind die Entwicklungskoeffizienten von

$$(x + A)^{n+1}.$$

Tritt auch nur ein p aus seinem Intervall heraus, so muss von den Wurzeln von $x^{n+1} + f(x) = 0$ eine dem absoluten Betrag nach grösser sein als A . Denn wären alle $\leq A$, so wären auch die Koeffizienten alle $\leq f^{(n+1)}$ (wo jedes p mit dem ihm entsprechenden $f^{(n+1)}$ zu vergleichen ist) wie aus der Beziehung zwischen Wurzeln und Koeffizienten einer Gleichung hervorgeht. Diese Wurzel sei x_1 ; $|x_1| > A$. Ist x_1 positiv, so betrachten wir in der Ebene der Variablen x diejenige Halbebene, in der der reelle Teil negativ ist, und umgekehrt; den Streifen $0 \dots (2n + 2)$ bzw. $0 \dots - (2n + 2)$ teilen wir durch Parallele zur imaginären Axe in $(n + 1)$ Streifen, von denen jeder 2 Einheiten breit ist. Es kann nun nicht in jedem dieser Streifen eine Wurzel von $x^{n+1} + f(x) = 0$ liegen, denn diese Gleichung hat ausser x_1 , das in der andern Halbebene liegt, nur noch n Wurzeln, während wir $(n + 1)$ Streifen haben. Es sei mit ξ der Mittelpunkt der reellen Axe in einem Streifen bezeichnet, in dem keine Wurzel liegt. Dann ist $|\xi^{n+1} + f(\xi)|$ grösser als A . Denn zerlegen wir $x^{n+1} + f(x)$ in Primfaktoren $(x - x_1)(x - x_2) \dots (x - x_{n+1})$, so ist für $x = \xi$ der erste Faktor absolut genommen grösser als A , alle andern sind grösser oder gleich 1. Nun handelt es sich nicht um $x^{n+1} + f(x)$, sondern um $f(x)$, das aber höchstens um x_{n+1} , d. h. im Maximum um $(2n + 2)^{n+1}$ kleiner sein kann. Wir müssen also A durch $A + (2n + 2)^{n+1}$ ersetzen. Wir kommen zu dem Resultat:

Der angekündigte Hilfssatz ist bewiesen; als die den Koeffizienten von

$$p_n x^n + p_{n-1} x^{n-1} + \dots + p_0$$

zuzuordnenden Intervalle können die Entwicklungskoeffizienten von

$$[x + A + (2n + 2)^{n+1}]^{n+1}$$

genommen werden, wenn das Intervall für $x - (2n + 2) \dots + (2n + 2)$ ist. Andernfalls ist noch eine Transformation erforderlich.

Kehren wir zu dem Existenztheorem des Minimums der Abweichung L zurück. Wenn M der absolut grösste Wert ist, den

die gegebene Funktion φ in dem Intervall annimmt, und L_1 ein irgendwie wirklich angenommener Werth der Abweichung (etwa $L_1 = M$, wenn $f(x) \equiv 0$), so setzen wir $A = |M| + |L_1|$; denn jetzt ist sicher ausserhalb unseres $(n + 1)$ dimensionalen Prismas $L > L_1$, weshalb wir uns bei der Aufsuchung des Minimums von L auf das $(n + 1)$ dimensionale Prisma selbst beschränken können. Hier im endlichen Bereich existiert aber das Minimum nach dem Weierstrassschen Fundamentalsatz, falls

$$L = L(p_0, p_1, \dots, p_n)$$

eine stetige Funktion ist. Das ist nun in der That der Fall; denn durch hinreichend kleine Aenderungen der p können wir erreichen, dass sich die ganze rationale Funktion $f(x)$ an jeder Stelle des Intervalls um weniger ändert als die beliebig klein vorgeschriebene Grösse ε ; dann ändert sich auch $f(x) - \varphi(x)$ an jeder Stelle weniger als ε , und also auch L , das Maximum dieser Differenz.

Diese Erwägung setzt die Stetigkeit von $\varphi(x)$ nicht voraus. Dagegen muss von $\varphi(x)$ Endlichkeit und Eindeutigkeit verlangt werden, da sonst die Definition von L ihren Sinn verliert. Wohl aber bleibt unser Existenzbeweis gültig für Funktionen $\varphi(x)$, die für einzelne Strecken des gegebenen Intervalles nicht definiert sind, und er gilt auch für den Fall des Interpolationsproblems, in dem $\varphi(x)$ nur an einzelnen diskreten Punkten definiert ist.

Der Beweis für die Existenz der Tchebycheffschen Annäherungsfunktion ist damit erbracht.

Verallgemeinerung für mehr Veränderliche. Dieser Existenzbeweis lässt sich ohne Schwierigkeit auch auf den Fall übertragen, dass eine Funktion mehrerer Variablen $\varphi(x, y)$ oder auch $\varphi(x_1, x_2, \dots, x_n)$ durch eine ganze rationale Funktion n ten Grades $f(x, y)$ bzw. $f(x_1, x_2, \dots, x_n)$ angenähert dargestellt werden soll, wo die Annäherungsfunktion wieder so definiert wird, dass das Maximum L von $|f - \varphi|$ im vorgeschriebenen Intervall ein Minimum wird. Wir können uns auf zwei Variablen beschränken, da die Verallgemeinerung durch den Schluss von n auf $n + 1$ nichts Neues bietet. Wir beweisen zunächst den algebraischen Hilfssatz. Es ist

$$f(x, y) = Cx^n + g_1(y) \cdot x^{n-1} + g_2(y) \cdot x^{n-2} + \dots + g_n(y).$$

Als Funktion von x aufgefasst können wir nach dem Vorigen

jedem Koeffizienten ein Intervall zuordnen, das nicht überschritten werden kann, ohne dass $f(x, y)$ an irgend einer Stelle des Intervalles A überschreitet. Sei etwa dem g_ν das Intervall $-a_\nu \dots + a_\nu$ zugeordnet. Wenden wir jetzt auf $g_\nu(y)$ unsern algebraischen Hilfssatz an, wobei a_ν die Rolle von A spielt, so lassen sich für die $\nu + 1$ Koeffizienten von g_ν Intervalle angeben, die nicht überschritten werden können, ohne dass $|g_\nu(y)|$ für irgend ein y des Intervalles grösser wird als a_ν . Dies sind die gesuchten Intervalle der Koeffizienten von $f(x, y)$, wie unmittelbar zu sehen. — Wir haben hierbei die Intervallgrenzen von x und von y als voneinander unabhängig angenommen, d. h. ein rechteckiges Intervall vorausgesetzt. Dies ist keine Beschränkung der Allgemeinheit, da wir zu jedem Intervall ein rechteckiges bestimmen können, das völlig innerhalb des gegebenen liegt. Dies rechteckige Intervall dient uns dann zur Bestimmung der Intervallgrenzen.

Es lassen sich also den Koeffizienten von $f(x, y)$ Intervalle zuordnen, derart, dass wenn auch nur ein Koeffizient aus seinem Intervall heraustritt, $|f(x, y)|$ für mindestens ein Wertepaar grösser wird als A . Der Existenzbeweis der Annäherungsfunktion gestaltet sich weiterhin nicht verschieden von dem Fall einer Variablen.

§ 3. Notwendige Bedingungen.

1. Ein allgemeiner Satz.

Wir verallgemeinern das Problem dadurch, dass wir die Differenz $f(x) - \varphi(x)$, wo $\varphi(x)$ die gegebene, $f(x)$ die gesuchte Funktion ist, als speziellen Fall einer Funktion $F(x, p_0, p_1 \dots p_n)$, die von einer Variablen x und $(n + 1)$ Parametern p abhängt, auffassen. Ueber $F(x)$ machen wir folgende Voraussetzungen:

1) $F(x)$ sei als Funktion von x in einem gegebenen Intervall, etwa $-1 \dots +1$ stetig.

2) Als Funktion der p sei $F(x, p_0, p_1 \dots p_n)$ im Reellen überall fortgesetzt differentierbar und in eine beliebig weit konvergierende Taylorsche Reihe entwickelbar.

3) Die Ableitungen nach p behalten den stetigen Charakter in x zwischen $-1 \dots +1$.

Wir stellen uns die Aufgabe, die Parameter p so zu wählen, dass die Abweichung der Funktion $F(x)$ von 0 ein Minimum wird, und suchen hierfür zunächst notwendige Bedingungen. Es seien

$$x_1, x_2, \dots, x_n$$

diejenigen alle voneinander verschiedenen Punkte des Intervalles $-1 \dots +1$, an denen $F(x)$ die Abweichung L annimmt, was ausser an den Grenzen $x = \pm 1$ nur bei Extremwerten von $F(x)$ der Fall sein kann. Ich behaupte nun eine notwendige Bedingung dafür, dass $L = L(p_0, p_1, \dots, p_n)$ ein Minimum wird, ist die, dass die algebraischen Gleichungen:

$$(1) \quad \begin{aligned} \frac{\partial F(x_1)}{\partial p_0} \cdot \lambda_0 + \frac{\partial F(x_1)}{\partial p_1} \cdot \lambda_1 + \dots + \frac{\partial F(x_1)}{\partial p_n} \cdot \lambda_n &= s_1 \\ \frac{\partial F(x_2)}{\partial p_0} \cdot \lambda_0 + \frac{\partial F(x_2)}{\partial p_1} \cdot \lambda_1 + \dots + \frac{\partial F(x_2)}{\partial p_n} \cdot \lambda_n &= s_2 \\ \vdots & \\ \frac{\partial F(x_r)}{\partial p_0} \cdot \lambda_0 + \frac{\partial F(x_r)}{\partial p_1} \cdot \lambda_1 + \dots + \frac{\partial F(x_r)}{\partial p_n} \cdot \lambda_n &= s_r \end{aligned}$$

die so zu verstehen sind, dass nach der partiellen Differentiation das zu untersuchende Wertsystem p_0, p_1, \dots, p_n eingesetzt wird, keine endlichen Werte für die λ ergeben.

Die rechten Seiten bedeuten dabei Grössen, die nur dem Vorzeichen nach bestimmt sind; es habe

$$\begin{aligned} s_1 & \text{ dasselbe Vorzeichen wie } F(x_1) \text{ und sei } \neq 0 \\ s_2 & \text{ ,, ,, ,, } F(x_2) \text{ ,, ,, } \neq 0 \\ \vdots & \\ s_r & \text{ ,, ,, ,, } F(x_r) \text{ ,, ,, } \neq 0. \end{aligned}$$

Ich behaupte, die Gleichungen dürfen für kein Wertsystem s ein endliches Wertsystem der λ ergeben, wenn die p so gewählt sind, dass

$$L = L(p_0, p_1, \dots, p_n)$$

ein Minimum ist.

Nehmen wir also an, für irgend ein System der s und p ergäben die Gleichungen ein endliches System λ , so ist zu zeigen, dass der diesem Wertsystem p entsprechende Wert \bar{L} von L kein Minimum ist, dass also für ein dem angenommenen System p benachbartes System $L = L(p_0, p_1, \dots, p_n)$ kleiner wird, d. h. dass für dies Wertsystem p $F(x)$ auf der ganzen Strecke

$$-1 \leq x \leq 1$$

kleiner bleibt als \bar{L} . Als dies benachbarte Wertsystem wählen wir

$$\begin{array}{lll} p_0 - \lambda_0 \varepsilon & \text{statt} & p_0 \\ p_1 - \lambda_1 \varepsilon & \text{,,} & p_1 \\ \vdots & & \vdots \\ p_n - \lambda_n \varepsilon & \text{,,} & p_n \end{array}$$

wo ε eine kleine positive Grösse ist, über deren Kleinheit wir uns die Verfügung noch vorbehalten. $F_\varepsilon(x)$ bezeichne die Funktion mit den neuen, $F_0(x)$ die mit den alten Parametern. Wir entwickeln $F_\varepsilon(x)$ nach Potenzen von ε , was nach unsern Voraussetzungen erlaubt ist:

$$\begin{aligned} F_\varepsilon &= F_0 + \left(\frac{\partial F}{\partial p_0} \cdot \frac{\partial p_0}{\partial \varepsilon} + \frac{\partial F}{\partial p_1} \cdot \frac{\partial p_1}{\partial \varepsilon} + \dots + \frac{\partial F}{\partial p_n} \cdot \frac{\partial p_n}{\partial \varepsilon} \right) \varepsilon + \dots \\ &= F_0 - \left(\frac{\partial F}{\partial p_0} \cdot \lambda_0 + \frac{\partial F}{\partial p_1} \cdot \lambda_1 + \dots + \frac{\partial F}{\partial p_n} \cdot \lambda_n \right) \varepsilon + (\varepsilon^2) + \dots \end{aligned}$$

Nun sieht man, dass diese Grösse für $-1 \leq x \leq +1$ absolut genommen überall kleiner ist als \bar{L} . Denn

1) An den Stellen x_1, x_2, \dots, x_n können wir für

$$\left(\frac{\partial F}{\partial p_0} \cdot \lambda_0 + \frac{\partial F}{\partial p_1} \cdot \lambda_1 + \dots + \frac{\partial F}{\partial p_n} \cdot \lambda_n \right)$$

die Grössen s einsetzen, und haben so z. B. für x_1

$$F_\varepsilon(x_1) = F_0(x_1) - s_1 \varepsilon + (\varepsilon^2) + \dots,$$

welche Grösse, da s_1 vom selben Vorzeichen ist als $F_0(x_1)$, für genügend kleine ε dem absoluten Betrag nach kleiner ist als

$$|F_0(x_1)| = \bar{L}.$$

2) Für die Nachbarschaft von x_1, x_2, \dots, x_n gilt dasselbe. In der Nähe eines jeden dieser Punkte grenzen wir einen Bereich der Variablen x ab, in dem immer noch

$$\left(\frac{\partial F}{\partial p_0} \cdot \lambda_0 + \frac{\partial F}{\partial p_1} \cdot \lambda_1 + \dots + \frac{\partial F}{\partial p_n} \cdot \lambda_n \right)$$

dasselbe Vorzeichen hat, wie die entsprechende Grösse s . Nach unsern Voraussetzungen müssen wir, da $s \neq 0$, überall einen endlichen Bereich erhalten, dessen Grösse wohl von s , aber nicht von ε abhängt.

3) Ausserhalb dieser Bereiche ist $F_0(x)$ überall um ein endliches Stück kleiner als \bar{L} , wegen der Stetigkeit gilt für $F_\varepsilon(x)$ dasselbe, wenn ε genügend klein genommen ist.

$F_\varepsilon(x)$ bleibt also in dem ganzen Intervall $-1 \leq x \leq +1$ kleiner als \bar{L} . F_ε ist also eine Funktion kleinerer Abweichung als F_0 , d. h. $L = L(p_0 p_1 \dots p_n)$ ist kein Minimum, wenn das System der linearen Gleichungen endliche Lösungen zulässt. — (Wir müssen jedoch beachten, dass unsere Herleitung zunächst nur dann einen Sinn hat, wenn die Punkte x_1, x_2, \dots, x_n , an denen $F(x)$ seine Abweichung annimmt, diskrete Punkte sind. Es darf also nicht $|F(x)|$ in einem endlichen Intervall gleich L sein.) Unser Resultat ist: damit die Abweichung der Funktion $F(x, p_0 p_1 \dots p_n)$ von 0 ein Minimum werde, ist die Nichtauflösbarkeit der Gleichungen (1) notwendige Bedingung.

2. Anwendung des Vorigen.

Statt der allgemeinen Funktion $F(x, p_0 p_1 \dots p_n)$ des vorigen Abschnittes setzen wir nun die Differenz

$$f_n(x) - \varphi(x),$$

wo $\varphi(x)$ eine gegebene, zwischen -1 und $+1$ stetige Funktion und $f_n(x)$ die gesuchte Annäherungsfunktion darstellt, wodurch unsern Bedingungen über $F(x)$ sicher genügt ist. Damit die Gleichung

$$|f_n(x) - \varphi(x)| = L$$

für endliche Intervalle ausgeschlossen ist, braucht nur vorausgesetzt zu werden, dass $\varphi(x)$ für kein endliches Intervall mit einer ganzen rationalen Funktion n ten oder niedrigeren Grades zusammenfällt. Da also

$$F(x) = p_n x^n + p_{n-1} x^{n-1} + \dots + p_0 - \varphi(x)$$

so lauten unsere Gleichungen (1)

$$\begin{aligned}
 & \lambda_0 + \lambda_1 x_1 + \lambda_2 x_1^2 + \dots + \lambda_n x_1^n = s_1 \\
 & \lambda_0 + \lambda_1 x_2 + \lambda_2 x_2^2 + \dots + \lambda_n x_2^n = s_2 \\
 & \vdots \\
 & \lambda_0 + \lambda_1 x_v + \lambda_2 x_v^2 + \dots + \lambda_n x_v^n = s_v.
 \end{aligned}
 \tag{1^a}$$

Bezeichnen wir

$$\lambda_0 + \lambda_1 x + \lambda_2 x^2 + \dots + \lambda_n x^n \text{ mit } H(x),$$

so ist $H(x)$ eine ganze rationale Funktion n ten Grades, die durch unser Gleichungssystem nur insoweit bestimmt wird, als sie an vorgeschriebenen Stellen von vorgeschriebenem Vorzeichen sein muss. Für die Frage, ob unser Gleichungssystem durch passende Wahl der Koeffizienten von $H(x)$ befriedigt werden kann, kommt nun nur die Anzahl der Zeichenwechsel unter den s_1, s_2, \dots, s_v in Betracht. Bedeutet etwa der Uebergang von s_1 zu s_2 einen Zeichenwechsel, so können wir dem dadurch genügen, dass wir eine Wurzel von $H(x) = 0$ zwischen x_1 und x_2 festsetzen. Thun wir dies bei allen Zeichenwechseln, und sorgen, dass $H(x_i)$ das richtige Vorzeichen hat, so sind alle Bedingungen unserer Gleichungen (1^a) erfüllt. Sind weniger Zeichenwechsel als Wurzeln vorhanden, so können wir die überzähligen Wurzeln ausserhalb des Intervalles $-1 \dots +1$ annehmen. Da wir n Wurzeln zur Verfügung haben, so sind bei n und weniger Zeichenwechseln unsere Gleichungen stets lösbar.

Unsere notwendige Bedingung für $f_n(x)$ als Funktion kleinster Abweichung von $\varphi(x)$ können wir also auch so aussprechen:

Die notwendige Bedingung für die Funktion kleinster Abweichung ist die, dass unter s_1, s_2, \dots, s_v d. h. unter

$$f_n(x_1) - \varphi(x_1), \quad f_n(x_2) - \varphi(x_2), \quad \dots \quad f_n(x_v) - \varphi(x_v),$$

wo x_1, x_2, \dots, x_v die Annahmestellen der Abweichung bedeuten, mindestens $(n+1)$ Zeichenwechsel auftreten.

Alle Erwägungen dieses Paragraphen lassen sich ausdehnen auf den Fall, dass $\varphi(x)$ nur abteilungsweise oder nur in diskreten Punkten definiert ist. Wir können uns ferner von der Einschränkung befreien, dass die Abweichung nur an diskreten Punkten angenommen werden soll. Werde etwa L in dem endlichen Intervall

$$a \leq x \leq b$$

angenommen, so haben wir in die Gleichungen (1) die folgende aufzunehmen:

$$\frac{\partial F(x)}{\partial p_0} \cdot \lambda_0 + \frac{\partial F(x)}{\partial p_1} \cdot \lambda_1 + \dots + \frac{\partial F(x)}{\partial p_n} \cdot \lambda_n = s_k$$

für

$$a \leq x \leq b$$

wodurch sich an der Beweisführung nichts ändert. Die Funktion $H(x)$ muss dann im ganzen Intervall $a \dots b$ das Vorzeichen von s_k behalten, und wir sehen sofort, dass ein solches Intervall wie ein diskreter Punkt anzusehen ist.

Die Annahme der Abweichung muss $(n + 2)$ Mal in abwechselndem Sinn erfolgen, wobei ein endliches Intervall, in dem die Abweichung angenommen wird, als eine einzige Annahme zu betrachten ist.

§. 4. Diese Bedingungen sind hinreichend und eindeutig.

Aus der Eindeutigkeit würde nach dem Existenzbeweise folgen, dass die Bedingungen hinreichend sind, wir wollen aber beides zugleich durch folgenden nun zu beweisenden Satz darthun:

Ist $f_n(x)$ ein Polynom n ten Grades, das der notwendigen Bedingung genügt, so giebt es kein von ihm verschiedenes Polynom n ten Grades von ebenso kleiner oder noch kleinerer Abweichung.

Sei also $\bar{f}_n(x)$ ein Polynom, dessen Abweichung von $\varphi(x)$ nicht grösser ist als L .

Falls nun, was ja möglich ist $\nu > (n + 2)$, so wählen wir unter x_1, x_2, \dots, x_ν $(n + 2)$ Werte aus von der Art, dass je zwei aufeinander folgende einen Zeichenwechsel von $f_n(x) - \varphi(x)$ liefern, und es seien nun diese Werte mit x_1, x_2, \dots, x_{n+2} bezeichnet, (die möglicherweise mit x_1, x_2, \dots, x_ν zusammenfallen können).

Wir betrachten:

$$[f_n(x) - \varphi(x)] - [\bar{f}_n(x) - \varphi(x)],$$

was eine ganze rationale Funktion vom höchstens n ten Grade ist.

$$= f_n(x) - \bar{f}_n(x).$$

An den Stellen $x_1, x_2 \dots x_{n+2}$ ist

$$|f_n(x) - \varphi(x)| = L,$$

während

$$|\bar{f}_n(x) - \varphi(x)| \leq L;$$

nehmen wir an, es gelte das Ungleichheitszeichen, so ist für das Vorzeichen von

$$[f_n(x) - \varphi(x)] - [\bar{f}_n(x) - \varphi(x)]$$

das des Minuenden massgebend, und es müsste dann die Differenz, ebenso wie der Minuend $(n + 1)$ Zeichenwechsel, also auch $(n + 1)$ Nullstellen zwischen -1 und $+1$ haben, d. h. $f_n(x) - \bar{f}_n(x) \equiv 0$. Gilt aber in $\bar{f}_n(x) - \varphi(x) \leq L$ an einigen oder allen Stellen $x_1, x_2 \dots x_{n+2}$ das Gleichheitszeichen, so rücken die Wurzeln von den Intervallen $x_1 \dots x_2, x_2 \dots x_3, \dots \dots x_{n+1} \dots x_{n+2}$ in diese Punkte selbst, ihre Zahl wird dabei, wie man sich leicht überzeugt, nicht geändert.

Auch an dieser Betrachtung ändert sich nichts, wenn wir statt eines der Punkte ein endliches Intervall einführen, in dem die Differenz ein festgesetztes Vorzeichen haben soll.

Damit ist unsere notwendige Bedingung als zugleich hinreichend und eindeutig nachgewiesen. Es ist nicht nur das kleinste Minimum, sondern überhaupt das Minimum von L als Funktion von $p_0, p_1 \dots p_n$ eindeutig bestimmt. In geometrischer Interpretation können wir unsere Resultate in folgenden Satz zusammenfassen:

Zu jeder stetigen Kurve lässt sich eine und nur eine Parabel n ten Grades angeben, die in einem vorgegebenen Intervall die Kurve $(n + 1)$ mal durchsetzt und sich an den $(n + 2)$ Punkten (oder Intervallen) grösster Entfernung gleich weit von der gegebenen Kurve entfernt.

Aus der Eindeutigkeit der Annäherungsfunktion folgt sofort folgender Satz: Liegt das definierende Intervall symmetrisch zum Nullpunkt, so ist die Annäherungsfunktion gerade, wenn die anzunähernde Funktion gerade ist, und ungerade, wenn diese ungerade ist. Im ersten Fall haben wir in $f(-x)$ im zweiten in $-f(-x)$ eine zweite Funktion, die den Bedingungen der Annäherungsfunktion genügt, woraus denn unser Satz folgt.

§ 5. Stetigkeit.

Wir benutzen die vorigen Sätze, um zu beweisen, dass sich die Annäherungsfunktion bei festgehaltenem Grad und Intervall mit der anzunähernden Funktion stetig ändert. Es sei $\varphi(x)$ eine in dem betrachteten Intervall den allgemeinen Bedingungen der Stetigkeit und Eindeutigkeit genügende Funktion, $f(x)$ die zugehörige Annäherungsfunktion n ten Grades. Ferner sei $\varphi_\varepsilon(x)$ eine von $\varphi(x)$ nur wenig verschiedene stetige und eindeutige Funktion, die in dem betrachteten Intervall

$$|\varphi(x) - \varphi_\varepsilon(x)| \leq \varepsilon$$

genügt, wo wir uns die Verfügung über die kleine Grösse ε noch vorbehalten. $\varphi_\varepsilon(x)$ verläuft in dem betrachteten Intervall völlig innerhalb eines von zwei parallelen Kurven begrenzten Streifens. Wenn nun $f_\varepsilon(x)$ die zu $\varphi_\varepsilon(x)$ gehörende Annäherungsfunktion ist, so behaupte ich kann auch

$$|f(x) - f_\varepsilon(x)|$$

durch geeignete Wahl von ε während des ganzen Intervalles beliebig klein gemacht werden. Zum Beweis betrachten wir die Funktion

$$[f(x) - \varphi(x)] - [f_\varepsilon(x) - \varphi_\varepsilon(x)]$$

an denjenigen Stellen x_1, x_2, \dots, x_{n+2} , an denen $[f(x) - \varphi(x)]$ seine Abweichung L , und zwar in abwechselndem Sinn annimmt. Sei nun L_ε die Abweichung der Funktion $f_\varepsilon(x)$ von $\varphi_\varepsilon(x)$, so ist

$$L_\varepsilon \leq |L + \varepsilon|$$

Wäre nämlich $L_\varepsilon > L + \varepsilon$, so würde $f(x)$, dessen Differenz von $\varphi_\varepsilon(x)$ im ganzen Intervall nirgends grösser wird als $L + \varepsilon$, besser annähern als $f_\varepsilon(x)$;

$$|f_\varepsilon(x) - \varphi_\varepsilon(x)|$$

ist also im ganzen Intervall, und also auch an den Stellen x_1, x_2, \dots, x_{n+2} nicht grösser als $L + \varepsilon$. Nehmen wir, um die Vorstellung zu fixieren, an, es sei $f(x) - \varphi(x)$ an der Stelle x_1 positiv, also gleich $+L$, so ist

$$[f(x) - \varphi(x)] - [f_\varepsilon(x) - \varphi_\varepsilon(x)]$$

für $x = x_1$ algebraisch genommen nicht kleiner als $-\varepsilon$. An der Stelle x_2 ist

$$[f(x) - \varphi(x)] - [f_\varepsilon(x) - \varphi_\varepsilon(x)]$$

nicht grösser als $+\varepsilon$. In dieser Weise fahren wir bis x_{n+2} fort.

Nun unterscheidet sich

$$f(x) - f_\varepsilon(x)$$

von

$$[f(x) - \varphi(x)] - [f_\varepsilon(x) - \varphi_\varepsilon(x)]$$

höchstens um ε ; also können wir von der Funktion

$$f(x) - f_\varepsilon(x),$$

die eine ganze rationale Funktion n ten Grades ist, folgendes sagen: Sie ist an der Stelle

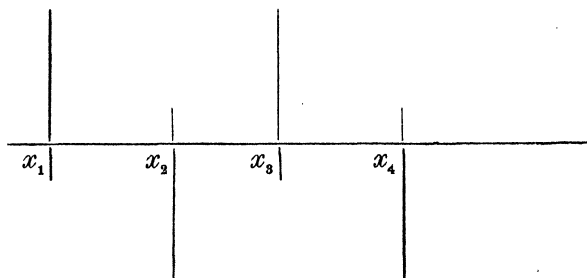
$$x_1 \text{ nicht kleiner als } -2\varepsilon$$

$$x_2 \text{ „ grösser „ } +2\varepsilon$$

$$x_3 \text{ „ kleiner „ } -2\varepsilon$$

.....

Die durch die Funktion dargestellte Kurve muss also durch die in unserer Figur an den Stellen x_1, x_2, \dots, x_{n+2} gezeichneten Geraden hindurchgehen.



Ich behaupte nun: In mindestens einem der Intervalle $x_1 \dots x_2, x_2 \dots x_3, \dots, x_{n+1} \dots x_{n+2}$ wird die Funktion $f(x) - f_\varepsilon(x)$ dem absoluten Betrage nach nicht grösser als 2ε .

Nehmen wir zunächst das Gegenteil an: In jedem der Intervalle läge eine Stelle, an der die Funktion $f(x) - f_\varepsilon(x)$ absolut genommen grösser wäre als 2ε ; ich behaupte: dann liegt in dem

Intervall $x_1 \dots x_3$ ein Minimum. Denn wenn 1) $f(x) - f_\varepsilon(x)$ an einer Stelle des Intervalls kleiner wird als -2ε , so muss diese Differenz, da sie an den Grenzen x_1 und x_3 nicht kleiner ist als -2ε , ein Minimum haben. Denn jede stetige Funktion hat im endlichen Intervall, wenn der kleinste Wert nicht an den Grenzen liegt, ein Minimum. Wenn aber 2) $f(x) - f_\varepsilon(x)$ im Intervall $x_1 \dots x_3$ nicht kleiner wird als -2ε , so muss es nach unserer Annahme in jedem der Teilintervalle $x_1 \dots x_2$ und $x_2 \dots x_3$ grösser werden als $+2\varepsilon$. Da es für x_2 nicht grösser sein kann als $+2\varepsilon$, so sieht man, dass auch in diesem Fall ein Minimum statthaben muss.

Auf diese Weise fahren wir fort: Ist unsere Annahme, dass in jedem Teilintervall die Funktion $f(x) - f_\varepsilon(x)$ aus dem Streifen $-2\varepsilon \dots +2\varepsilon$ heraustritt, richtig, so lässt sich zeigen, dass in jedem der Intervalle mit ungeraden Indices $x_1 \dots x_3, x_5 \dots x_7$ ein Minimum und in jedem der Intervalle mit geraden Indices $x_2 \dots x_4, x_4 \dots x_6 \dots$ ein Maximum liegen muss. Wir haben also, da n solcher Intervalle vorhanden sind: $x_1 \dots x_3, x_2 \dots x_4 \dots x_n \dots x_{n+2}$, den Widerspruch, dass $f(x) - f_\varepsilon(x)$ als ganze rationale Funktion n ten Grades n Extremwerte annehmen müsste.

Folglich war die Annahme unrichtig; die Funktion $f(x) - f_\varepsilon(x)$ wird in der That in mindestens einem der Intervalle

$$x_1 \dots x_3 \quad x_2 \dots x_4 \dots x_{n-1} \dots x_{n-2}$$

dem absoluten Betrag nach nicht grösser als 2ε .

Man sieht nun ferner leicht, dass, wenn eine ganze rationale Funktion n ten Grades in einem endlichen Intervall, dessen Grenzen von ε unabhängig sind, absolut genommen kleiner bleibt als ε (wie wir Kürze halber statt 2ε schreiben), die Koeffizienten dieser Funktion für hinreichend kleine ε beliebig klein gemacht werden können. Es genügt schon, dass die Funktion an $(n+1)$ festgehaltenen diskreten Punkten $\xi_1, \xi_2 \dots \xi_{n+1}$ die wir, wenn das ganze Intervall gegeben ist, etwa in gleichem Abstand von einander annehmen, kleiner bleibt als ε . Denn wenn wir die Koeffizienten bestimmen wollen, so geschieht dies aus den Gleichungen:

$$\begin{aligned} p_n \xi_1^n + p_{n-1} \xi_1^{n-1} + \dots + p_0 &= \varepsilon_1 \\ p_n \xi_2^n + p_{n-1} \xi_2^{n-1} + \dots + p_0 &= \varepsilon_2 \\ &\vdots \\ p_n \xi_{n+1}^n + p_{n-1} \xi_{n+1}^{n-1} + \dots + p_0 &= \varepsilon_{n+1} \end{aligned}$$

wo die

$$\varepsilon_1, \varepsilon_2 \dots \varepsilon_{p+1} \leq \varepsilon.$$

Die p erscheinen in der Form

$$\frac{d_1}{D} \varepsilon_1 + \frac{d_2}{D} \varepsilon_2 + \dots + \frac{d_{n+1}}{D} \cdot \varepsilon_{n+1},$$

wo D die Cauchysche Determinante und die d Unterdeterminanten der ξ , also von ε unabhängig sind; es werden also in der That alle p für hinreichend kleines ε beliebig klein.

Ist für einen der Koeffizienten von $f(x) - f_\varepsilon(x)$ eine obere Grenze des absoluten Betrages vorgeschrieben, so müssen wir für jedes der $(n + 1)$ Intervalle $x_1 \dots x_2, x_2 \dots x_3, \dots x_{n+1} \dots x_{n+2}$ auf die angegebene Weise ε bestimmen; der kleinste der so gefundenen Werte giebt die zulässige Grösse von

$$\varphi(x) - \varphi_\varepsilon(x)$$

an.

Ist $\varphi(x)$ nur durch einen „Funktionsstreifen“ gegeben, so liegt die zugehörige Annäherungsfunktion in einem Gebiet, das mit dem Funktionsstreifen beliebig schmal wird. Wir können auch sagen: Der topographische Charakter der Annäherungsfunktion ist nur von dem topographischen Charakter der anzunähernden Funktion abhängig (also z. B. nicht von ihrer analytischen Natur; dagegen wird sich zeigen, dass die analytische Natur der Koeffizienten der Annäherungsfunktion durch die analytische Natur der gegebenen Funktion bedingt ist.)

§ 6. Exkurs über die Darstellung durch rationale Funktionen.

Unser bisheriger Gedankengang lässt sich leicht übertragen für den Fall, dass die Annäherung nicht durch ganze, sondern durch gebrochene rationale Funktionen bewirkt werden soll.

Existenzbeweis.

Es sei eine in dem zu betrachtenden Intervall $-1 \dots +1$ endliche und eindeutige Funktion $\varphi(x)$ gegeben und n als Grad für den Nenner, m für den Zähler der Annäherungsfunktion vorgeschrieben. Wir wollen beweisen, dass die Abweichung der Funktionen der Form

$$\frac{q_m x^m + q_{m-1} x^{m-1} + \dots + q_0}{p_n x^n + p_{n-1} x^{n-1} + \dots + 1}$$

ein Minimum annimmt, und betrachten zu diesem Zweck die Funktionen der Form:

$$\frac{q_m x^m + q_{m-1} x^{m-1} + \dots + q_0}{p_n x^n + p_{n-1} x^{n-1} + \dots + p_0}$$

Aus dem $(m + n + 2)$ dimensionalen Raum, den diese Funktionen erfüllen, schneiden wir durch die Bedingungen

$$(2) \quad \begin{array}{l} |q_m| \leq 1 \quad |q_{m-1}| \leq 1 \dots |q_0| \leq 1 \\ |p_n| \leq 1 \quad |p_{n-1}| \leq 1 \dots |p_0| \leq 1 \end{array}$$

einen Würfel aus und betrachten die Wandungen dieses Würfels, an denen von (2) mindestens einmal das Gleichheitszeichen gilt; ich behaupte L als Funktion der q und p hat auf diesen Wandungen ein Minimum, für das $p_0 \neq 0$.

L kann in diesem Gebiet als stetige Funktion der p und q betrachtet werden, denn die einzigen Unstetigkeitsstellen sind die Unendlichkeitsstellen, die für unsere Minimumfrage nicht in Betracht kommen. L nimmt also sein Minimum nach dem Weierstrassschen Satz an. Ist für dieses $p_0 \neq 0$, so ist der Satz bewiesen, ist aber $p_0 = 0$, so muss notwendig auch $q_0 = 0$, weil sonst für $x = 0$ die Annäherungsfunktion, also auch L unendlich würde, ein Minimum also sicher nicht stattfände. Wir können alsdann Zähler und Nenner durch x teilen und so den Grad beider um 1 erniedrigen. Wir gehen deshalb zu einem um 2 Dimensionen niedrigeren Raum und betrachten in diesem den analogen Würfel. Auf diese Weise fahren wir fort, und wenn das Verfahren nicht früher abbricht, kommen wir entweder auf Funktionen der Form

$$\frac{q_\mu x^\mu + q_{\mu-1} x^{\mu-1} + \dots + q_0}{p_0}$$

d. h. auf ganze Funktionen, die wir bereits abgehandelt haben oder auf Funktionen der Form

$$\frac{q_0}{p_r x^r + p_{r-1} x^{r-1} + \dots + p_0}.$$

Bei dem Minimum dieser Funktionen kann nicht mehr $p_0 = 0$;

denn es müsste wegen des Punktes $x = 0$ alsdann auch $q_0 = 0$. Der Nenner kann aber nicht identisch 0 sein, weil ja mindestens ein $|p|$ gleich 1 sein muss. Wir würden also auf die identische Null geführt, die wir aber auch durch $q_0 = 0; p_\nu = p_{\nu-1} = \dots = p_1 = 0 \quad p_0 = 1$ erreichen können. Wir bekommen also stets ein Minimum, für das $p_0 \neq 0$. Dividieren wir alle Koeffizienten durch p_0 , so haben wir die gesuchte rationale Funktion. Es giebt keine besser annähernde rationale Funktion der Grade m und n ; denn gäbe es eine, so könnten wir alle Koeffizienten durch den absolut grössten unter ihnen dividieren, wodurch der Widerspruch zu Tage träte.

Zu jeder eindeutigen und endlichen Funktion giebt es in einem gegebenen Intervall eine am besten annähernde, rationale Funktion, deren Grad im Zähler und Nenner vorgeschrieben ist, und deren absolutes Glied im Nenner gleich 1 ist.

Notwendige Bedingungen.

Um notwendige Bedingungen zu finden, wenden wir wieder unsern allgemeinen Satz S. 14 und die Bedingungen (1) an, die hier, wenn

$$F(x) = \frac{f_m(x)}{f_n(x)} - \varphi(x).$$

$$f_m(x) = q_m x^m + q_{m-1} x^{m-1} + \dots + q_0$$

$$f_n(x) = p_n x^n + p_{n-1} x^{n-1} + \dots + p_1 x + 1$$

folgendermassen lauten: Es dürfen die Gleichungen:

$$\frac{1}{f_n(x_1)} (x_1^m \lambda_m + x_1^{m-1} \lambda_{m-1} + \dots + \lambda_0) - \frac{f_m(x_1)}{(f_n(x_1))^2} (x_1^n \lambda_{m+n} + x_1^{n-1} \lambda_{m+n-1} + \dots + x_1 \lambda_{m+1}) = s_1$$

$$(1^b) \frac{1}{f_n(x_2)} (x_2^m \lambda_m + x_2^{m-1} \lambda_{m-1} + \dots + \lambda_0) - \frac{f_m(x_2)}{(f_n(x_2))^2} (x_2^n \lambda_{m+n} + x_2^{n-1} \lambda_{m+n-1} + \dots + x_2 \lambda_{m+1}) = s_2$$

$$\dots \dots \dots \frac{1}{f_n(x_\nu)} (x_\nu^m \lambda_m + x_\nu^{m-1} \lambda_{m-1} + \dots + \lambda_0) - \frac{f_m(x_\nu)}{(f_n(x_\nu))^2} (x_\nu^n \lambda_{m+n} + x_\nu^{n-1} \lambda_{m+n-1} + \dots + x_\nu \lambda_{m+1}) = s_\nu$$

kein endliches System der λ zulassen, das sie befriedigt. Ich multipliziere nun die erste Gleichung mit $f_n(x_1)^2$, die zweite mit $(f_n(x_2))^2$ u. s. f. Der Sinn der rechten Seite wird dadurch nicht geändert, da ja s_1, s_2, \dots, s_v nur in Bezug auf das Vorzeichen bestimmt sind, das durch Multiplikation mit $(f_n(x_1))^2, \dots$ nicht geändert werden kann. Auch kann $f_n(x_1)$ nicht gleich 0 sein, weil $f_n(x)$ im Nenner der Annäherungsfunktion steht, die nicht unendlich werden kann; (die Möglichkeit gemeinsamer Wurzeln von $f_n(x)$ und $f_m(x)$ werden wir sogleich ausschliessen). $s_1 \cdot (f_n(x_1))^2$ bedeutet also dasselbe wie s_1 . Denken wir uns ferner, um das negative Vorzeichen zu entfernen, die $\lambda_{m+1}, \lambda_{m+2}, \dots, \lambda_{m+n}$ durch ihre negativen Werte ersetzt, so können wir unsere Gleichungen schreiben:

$$\begin{aligned} f_n(x_1) (x_1^m \lambda_m + x_1^{m-1} \lambda_{m-1} + \dots + \lambda_0) \\ + f_n(x_1) (x_1^n \lambda_{m+n} + x_1^{n-1} \lambda_{m+n-1} + \dots + x_1 \lambda_{m+1}) &= s_1 \\ f_n(x_2) (x_2^m \lambda_m + x_2^{m-1} \lambda_{m-1} + \dots + \lambda_0) \\ + f_n(x_2) (x_2^n \lambda_{m+n} + x_2^{n-1} \lambda_{m+n-1} + \dots + x_2 \lambda_{m+1}) &= s_2 \\ \vdots \\ f_n(x_v) (x_v^m \lambda_m + x_v^{m-1} \lambda_{m-1} + \dots + \lambda_0) \\ + f_n(x_v) (x_v^n \lambda_{m+n} + x_v^{n-1} \lambda_{m+n-1} + \dots + x_v \lambda_{m+1}) &= s_v. \end{aligned}$$

Diese Gleichungen sind vom $(m+n)$ ten Grade; wir können sie also schreiben:

$$\begin{aligned} b_{m+n} x_1^{m+n} + b_{m+n-1} x_1^{m+n-1} + \dots + b_1 x_1 + b_0 &= s_1 \\ b_{m+n} x_2^{m+n} + b_{m+n-1} x_2^{m+n-1} + \dots + b_1 x_2 + b_0 &= s_2 \\ \vdots \\ b_{m+n} x_v^{m+n} + b_{m+n-1} x_v^{m+n-1} + \dots + b_1 x_v + b_0 &= s_v \end{aligned}$$

wobei

$$\begin{aligned} b_0 &= \lambda_0 \\ b_1 &= \lambda_0 p_1 + \lambda_1 && + \lambda_{m+1} q_0 \\ b_2 &= \lambda_0 p_2 + \lambda_1 p_1 + \lambda_2 && + \lambda_{m+1} q_1 + \lambda_{m+2} q_2 \\ (4) \quad b_3 &= \lambda_0 p_3 + \lambda_1 p_2 + \lambda_2 p_1 + \lambda_3 && + \lambda_{m+1} q_2 + \lambda_{m+2} q_1 + \lambda_{m+3} q_0 \\ &\vdots \\ b_{m+n-1} &= \lambda_{m-1} p_n + \lambda_m p_{n-1} && + \lambda_{m+n-1} q_m + \lambda_{m+n} q_{m-1} \\ b_{m+n} &= \lambda_m p_n && + \lambda_{m+n} q_m. \end{aligned}$$

Die $(m+n+1)$ reihige Determinante dieses Systems lautet, wenn wir Zeilen und Spalten vertauschen:

$$D = \begin{vmatrix} 1 & p_1 & p_2 & p_3 & \dots & p_n & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & p_1 & p_2 & \dots & p_{n-1} & p_n & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & p_1 & \dots & p_{n-2} & p_{n-1} & p_n & 0 & \dots & 0 \\ 0 & 0 & 0 & 1 & \dots & p_{n-3} & p_{n-2} & p_{n-1} & p_n & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & q_0 & q_1 & q_2 & q_3 & \dots & q_m & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & q_0 & q_1 & q_2 & \dots & q_{m-1} & q_m & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & q_0 & q_1 & \dots & q_{m-2} & q_{m-1} & q_m & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & q_0 & \dots & q_{m-3} & q_{m-2} & q_{m-1} & q_m & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{vmatrix}$$

Denken wir uns vor dem q_0 noch ein \bar{q} stehen, das wir nachher gleich 0 setzen wollen, so sehen wir, dass D die Resultante der beiden Funktionen ist,

$$1 + p_1x + p_2x^2 + \dots + p_nx^n$$

und

$$\bar{q} + q_0x + q_1x^2 + \dots + q_mx^{m+1}$$

d. h. der beiden Funktionen

$$f_n(x) \quad \text{und} \quad x \cdot f_m(x).$$

Da nun das absolute Glied in $f_n(x)$ gleich 1 ist, $f_n(x)$ also für $x = 0$ nicht verschwindet, so haben diese Funktionen dann und nur dann einen Faktor gemein, wenn $f_n(x)$ und $f_m(x)$ einen Faktor gemein haben. Der zunächst zu betrachtende Fall $D \neq 0$ ist also identisch damit, dass

$$\frac{f_m(x)}{f_n(x)}$$

irreduzibel ist.

In diesem Fall sind die Gleichungen (1^b) vollständig äquivalent mit (3), und die Bedingung, dass (1^b) keine bestimmten Werte für λ ergeben sollen, ist identisch damit, dass (3) keine bestimmten Werte für b ergeben sollen. Dies führt, wie oben

bei den ganzen rationalen Funktionen auf die Bedingung, dass unter den $s_1, s_2 \dots s_v$ sich mindestens $m+n+1$ Zeichenwechsel finden.

Hinreichende und eindeutige Bedingungen.

Diese unter der Voraussetzung $D \neq 0$ notwendige Bedingung ist stets hinreichend und bestimmt das kleinste Minimum eindeutig, was wie oben durch Betrachtung der Funktion

$$\begin{aligned} & \left[\frac{f_m(x)}{f_n(x)} - \varphi(x) \right] - \left[\frac{\bar{f}_m(x)}{\bar{f}_n(x)} - \varphi(x) \right] \\ &= \frac{f_m(x)}{f_n(x)} - \frac{\bar{f}_m(x)}{\bar{f}_n(x)} = \frac{f_m \cdot \bar{f}_n - \bar{f}_m \cdot f_n}{f_n \cdot \bar{f}_n}, \end{aligned}$$

die nicht $(m+n+1)$ Zeichenwechsel haben kann, ohne dass

$$\begin{aligned} f_m \cdot \bar{f}_n - \bar{f}_m \cdot f_n &\equiv 0 \\ \frac{f_m}{f_n} &\equiv \frac{\bar{f}_m}{\bar{f}_n}, \end{aligned}$$

nachgewiesen wird. Der Nachweis der Eindeutigkeit erstreckt sich aber hier nur auf das kleinste Minimum, die Frage nach der Möglichkeit grösserer Minima, die $D = 0$ genügen, bleibt offen.

Die Untersuchung gestaltet sich also so, dass wir zuerst nach einer rationalen Funktion $\frac{f_m(x)}{f_n(x)}$ suchen, die $(m+n+1)$ Zeichenwechsel der Abweichung ergibt. Existiert diese, so ist sie ein Minimum, und es existiert dann keine rationale Funktion der Form $\frac{f_m(x)}{f_n(x)}$ mit ebenso kleiner oder noch kleinerer Abweichung. Existiert sie nicht, so bedeutet dies, dass die Zahlen m und n zu hoch gegriffen sind, dass wir mit Funktionen $\frac{f_{m-1}(x)}{f_{n-1}(x)}$ dieselbe Annäherung erreichen können. Wir suchen also unter diesen Funktionen solche, die $m+n-1$ Zeichenwechsel ergeben, u. s. f.

Zusammenfassend können wir sagen: Die Bedingung von $(m+n+1)$ Zeichenwechseln ist stets hinreichend, und wenn sie erfüllbar ist, auch notwendig.

§ 7. Interpolationsprobleme.

1. Geometrisches Beispiel für symmetrische Behandlung der beiden Variablen.

Wir wollen folgendes geometrische Problem behandeln, das

vielleicht geeignet ist, den Sinn unserer Methoden zu veranschaulichen. Es sei K ein Kegelschnitt und 1, 2 . . . 6 Punkte in der Ebene. Von jedem dieser Punkte betrachte ich das kürzeste der Lote, die sich von ihm an K ziehen lassen, und nenne es L_v . Die grösste der Strecken L_v sei mit L bezeichnet.

Es seien nun die Punkte 1, 2 . . . 6 gegeben und die Aufgabe gestellt, denjenigen Kegelschnitt zu finden, für den L ein Minimum wird.

Die Existenz der Lösung folgt, wenn wir bedenken, dass es zur Bestimmung des Kegelschnittes nur auf die Verhältnisse seiner Koeffizienten ankommt, aus Betrachtungen, die den S. 21—23 angestellten ähnlich sind.

Ich behaupte, eine notwendige Bedingung für das Minimum ist die Gleichheit aller L_v . Nehmen wir also an, es seien weniger als 6, etwa 2, der Grössen L_v gleich L . Auf den 2 Normalen, deren $L_v = L$, trage ich nun vom Kegelschnitt aus eine Strecke ε ab, die den gegebenen Punkten zugerichtet ist. Die Endpunkten dieser beiden Strecken zusammen mit 3 der übrigen 4 Normalen-Fusspunkte definieren einen neuen Kegelschnitt, dessen L , wenn ε genügend klein gewählt ist, kleiner ist, als das des erst angenommenen.

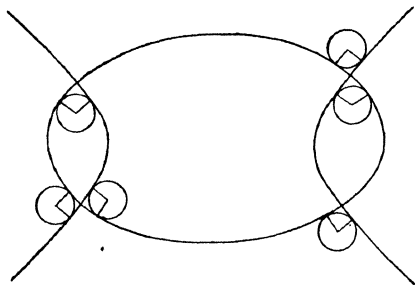
Denn von den zwei ausgezeichneten Punkten existiert eine Verbindungsstrecke zum Kegelschnitt von der Länge $L - \varepsilon$, die senkrechte Entfernung ist daher gleich oder kleiner $L - \varepsilon$. Von den drei hinzugenommenen Punkten existiert eine Verbindungsstrecke L_v , wo $L_v < L$. Die Entfernung des letzten Punktes aber, die kleiner war als L , bleibt dies auch bei dem neuen Kegelschnitt bei genügend kleinem ε wegen der Stetigkeit.

Fixieren wir, Einfachheit halber, unsern Kegelschnitt als Ellipse, so können wir weiter sagen, dass 1, 2 . . . 6 bei dem verlangten Kegelschnitt abwechselnd innerhalb und ausserhalb der Ellipse liegen müssen. Das „abwechselnd“ ist hierbei so zu verstehen: Beim Umlauf der Ellipse ergibt sich für die Fusspunkte der Lote eine cykliche Reihenfolge. Diese soll auch auf die gegebenen Punkte selbst übertragen werden. Nehmen wir etwa an, zwei aufeinanderfolgende Punkte, etwa 1 und 2 lägen ausserhalb der Ellipse. Ich betrachte nun die Schaar der Ellipsen, die durch die festgehaltenen Fusspunkte der Lote von den Punkten 3, 4, 5 und 6 gehen. Keine dieser Ellipsen kann auf dem Bogen 3 1 2 6 die ursprüngliche Ellipse schneiden, weil wir sonst zwei Ellipsen mit 5 gemeinsamen Punkten hätten. Jede dieser Ellipsen liegt daher

längs des Bogens 3 1 2 6 entweder ganz innerhalb oder ganz ausserhalb der ursprünglichen Ellipse. Diejenigen Nachbar-ellipsen, die ganz ausserhalb liegen, haben einen geringeren Abstand von den Punkten 1 und 2 als die ursprüngliche. Daher ist L entweder kleiner geworden, oder die L , sind für die neue Ellipse nicht mehr gleich, und für diesen Fall zeigten wir schon, dass L verkleinert werden kann.

Für Hyperbel und Parabel gilt derselbe Satz. Unser für die Ellipse gültiger Schluss, dass ein Schnittpunkt vorhanden sein müsste, falls der neue Kegelschnitt längs eines gewissen Bogens nicht entweder ganz innerhalb oder ganz ausserhalb des alten läge, lässt sich dort durch einfache projektive Erwägungen ersetzen.

Die Frage nach der Eindeutigkeit oder Vieldeutigkeit der Aufgabe ist hier nicht so einfach zu entscheiden, als bei dem Tchebycheffschen Problem, da Eindeutigkeit allgemein jedenfalls nicht existiert. Dies rührt daher, dass man von einer Reihenfolge der gegebenen Punkte, ehe der gesuchte Annäherungs-Kegelschnitt bekannt ist, nicht reden kann. In der That ist in unserer Figur, die einen Fall der Zweideutigkeit der Lösung zeigt, die Reihenfolge der Punkte in Bezug auf jeden der beiden Kegelschnitte eine verschiedene.



2. Das gewöhnliche Interpolationsproblem.

Es seien $(n + 2)$ Punkte

$$(x_1, y_1) \quad (x_2, y_2) \quad \dots \quad (x_{n+2}, y_{n+2})$$

gegeben, und es sei gesucht ein Polynom n ten Grades

$$y = x^n p_n + x^{n-1} p_{n-1} + \dots + p_0$$

von der Art, dass die grösste der Differenzen

$$y - y_\nu, \quad \nu = 1, 2, \dots, n + 2$$

möglichst klein sei.

Unsere Theorie liefert uns sofort die Lösung: Sind die Punkte in der Richtung der wachsenden Abscissen angeordnet,

also $x_1 < x_2 < x_3 \dots$ (den Fall zweier gleichen x , der einer Zweideutigkeit von $\varphi(x)$ entsprechen würde, schliessen wir aus), so müssen alle $(n+2)$ Werte $y-y_v$ absolut gleich, aber abwechselnd positiv und negativ sein. Dies führt uns auf folgendes System linearer Gleichungen:

$$p_n x_1^n + p_{n-1} x_1^{n-1} + \dots + p_0 - L = y_1$$

$$p_n x_2^n + p_{n-1} x_2^{n-1} + \dots + p_0 + L = y_2$$

⋮

$$p_n x_{n+2}^n + p_{n-1} x_{n+2}^{n-1} + \dots + p_0 + (-1)^{n+2} L = y_{n+1}$$

wo die x bekannt, die p und L zu bestimmen sind. Die Determinante dieses System ist die folgende:

$$\begin{vmatrix} x_1^n & x_1^{n-1} & \dots & 1 - 1 \\ x_2^n & x_2^{n-1} & \dots & 1 + 1 \\ x_3^n & x_3^{n-1} & \dots & 1 - 1 \\ \vdots & \vdots & \vdots & \vdots \\ x_{n+2}^n & x_{n+2}^{n-1} & \dots & 1 (-1)^{n+2} \end{vmatrix}$$

und man kann leicht sehen, dass diese Determinante, wie unsere allgemeine Theorie dies auch fordert, stets von 0 verschieden ist. Entwickeln wir nach den Elementen der letzten Spalte, so erhalten wir:

$$\begin{vmatrix} x_1^n & x_1^{n-1} & \dots & 1 \\ x_2^n & x_2^{n-1} & \dots & 1 \\ \vdots & \vdots & \vdots & \vdots \\ x_{n+1}^n & x_{n+1}^{n-1} & \dots & 1 \end{vmatrix} + \begin{vmatrix} x_1^n & x_1^{n-1} & \dots & 1 \\ \vdots & \vdots & \vdots & \vdots \\ x_n^n & x_n^{n-1} & \dots & 1 \\ x_{n+2}^n & x_{n+2}^{n-1} & \dots & 1 \end{vmatrix} + \dots + \begin{vmatrix} x_2^n & x_2^{n-1} & \dots & 1 \\ x_3^n & x_3^{n-1} & \dots & 1 \\ \vdots & \vdots & \vdots & \vdots \\ x_{n+2}^n & x_{n+2}^{n-1} & \dots & 1 \end{vmatrix}$$

wo die Summe das Vorzeichen $(-1)^{n-2}$ erhält. Diese Summe kann nicht verschwinden; denn alle Determinanten haben dieselben Vorzeichen, weil in jeder von ihnen die Basis der Elemente einer Zeile grösser ist als die aller vorhergehenden, und hierdurch das

Vorzeichen des Differenzenprodukts, dem diese Determinanten gleich sind, bestimmt ist, nämlich gleich dem von $(-1)^{\frac{n(n+1)}{2}}$

Die Bedingung $x_1 < x_2 < x_3 \dots$ ist dabei durchaus wesentlich, man findet z. B. für $n = 2$ durch Ausrechnen:

$$\begin{vmatrix} x_1^2 x_1 & 1 & -1 \\ x_2^2 x_2 & 1 & 1 \\ x_3^2 x_3 & 1 & -1 \\ x_4^2 x_4 & 1 & 1 \end{vmatrix} = 2(x_1 - x_2)(x_2 - x_4)(x_1 + x_3 - x_2 - x_4).$$

Die Determinante kann also für voneinander verschiedene x verschwinden, wenn x_2 und x_4 zwischen x_1 und x_3 liegen. Geometrisch bedeutet dies: Wenn 4 Punkte gegeben sind und eine Parabel gesucht wird, die in der Richtung der y -Axe von allen 4 gleich weit entfernt ist, so ist dies stets möglich, wenn die Punkte nach wachsenden Abscissen geordnet abwechselnd auf der einen und der andern Seite der Parabel liegen sollen, dagegen unter Umständen unmöglich, wenn zwei Punkte hintereinander auf derselben Seite der Parabel liegen sollen.

Sind mehr als $(n+2)$ Punkte gegeben, so ist folgendermassen zu verfahren: Man wähle auf alle möglichen Weisen aus den gegebenen Punkten $(n+2)$ aus und berechne die hierdurch eindeutig bestimmte Kurve nach dem oben angegebenen Verfahren. Unter allen so erhaltenen Kurven giebt es eine und nur eine, bei der der Abstand jedes der Punkte, die nicht zu ihrer Berechnung gedient haben kleiner ist als der Abstand der $n+2$. Unter Abstand ist dabei die Entfernung in der Richtung der Ordinate verstanden, bei „kleiner“ ist Gleichheit nicht ausgeschlossen.

Der Beweis ergibt sich aus der allgemeinen Theorie. Das Interpolationsproblem liefert also die wesentlichsten Resultate der Tchebycheffschen Annäherungsmethoden. Das Problem reduziert sich aber hier auf die Auflösung linearer Gleichungen.

Kapitel II. Aufstellung der Annäherungs-Funktion.

§ 1. Analytische Formulierung.

Wir kehren zu dem Problem zurück, zu der gegebenen stetigen Funktion $\varphi(x)$ für das Intervall $-1 \dots +1$ die Annähe-

rungs-Funktion zu bestimmen. Wenn wir nun zunächst unsere Bedingungen für die Annäherungsfunktion S. 15 ff. analytisch ausdrücken wollen, müssen wir bedenken, dass, weil $|f_n(x) - \varphi(x)|$ im Intervall $-1 \dots +1$ kleiner oder gleich L , die Werte von $|f_n(x) - \varphi(x)| = L$ Extremwerte darstellen, mit Ausnahme der Werte $x = \pm 1$, falls für sie die Gleichung statt hat.

$f_n(x)$ unterliegt dann den Bedingungsgleichungen:

$$\begin{aligned}
 & f_n(x_1) - \varphi(x_1) = L \\
 & (x_1 + 1) (f'_n(x_1) - \varphi'(x_1)) = 0 \\
 & f_n(x_2) - \varphi(x_2) = -L \\
 & f'_n(x_2) - \varphi'(x_2) = 0 \\
 (1) \quad & \cdot \cdot \cdot \cdot \cdot \cdot \cdot \\
 & \cdot \cdot \cdot \cdot \cdot \cdot \cdot \\
 & f_n(x_{n+1}) - \varphi(x_{n+1}) = (-1)^n L \\
 & f'(x_{n+1}) - \varphi'(x_{n+1}) = 0 \\
 & f_n(x_{n+2}) - \varphi(x_{n+2}) = (-1)^{n+1} L \\
 & (x_{n+2} - 1) (f'(x_{n+2}) - \varphi'(x_{n+2})) = 0.
 \end{aligned}$$

Praktisch ist jedoch häufig die folgende Form bequemer, obgleich sich in ihr der Grad der Gleichungen verdoppelt.

Es sollen die Gleichungen:

$$\begin{aligned}
 (2) \quad & (f_n(x) - \varphi(x))^2 - L^2 = 0 \\
 & (x^2 - 1) (f'_n(x) - \varphi'(x)) = 0
 \end{aligned}$$

zwischen $-1 \dots +1$ ($n+2$) voneinander verschiedene gemeinsame Wurzeln haben.

Indess sind diese analytisch formulierten Bedingungen nicht mehr hinreichend, sondern bloss noch notwendig. Denn sie enthalten nicht mehr unter allen Umständen die Bedingung, dass

$$|f_n(x) - \varphi(x)| \leq L \quad \text{für} \quad -1 \leq x \leq 1,$$

worauf wir später zurückzukommen haben.

Aus der Definition der Annäherungs-Funktion folgt ein Satz, der im Folgenden mehrfach angewandt werden wird: Die Annäherungs-Funktion lässt Multiplikation mit einer Konstanten zu. Ist $f_n(x)$ Annäherungsfunktion zu $\varphi(x)$, so ist im selben Intervall $C \cdot f_n(x)$ Annäherungsfunktion zu $C \cdot \varphi(x)$.

§ 2. Aufstellung für den einfachsten Fall.

Es sei

$$\varphi(x) = x^{n+1}$$

die Abweichung von $x^{n+1} - f_n(x)$ von 0 soll ein Minimum werden.

Die zweite der Gleichungen (2) ist dann nur vom Grade $(n+2)$; es müssen daher alle ihre Wurzeln auch Wurzeln der ersten Gleichung (2) sein. Nun sind die n Wurzeln, die nicht $(x^2-1) = 0$ sondern $f'_n(x) - \varphi'(x) = 0$ erfüllen, sämtlich Doppelwurzeln von $(f_n(x) - \varphi(x))^2 - L^2 = 0$, weil ja eben die Ableitung der linken Seite verschwindet, d. h.

$$(f_n(x) - \varphi(x))^2 - L^2 = 0$$

hat mit

$$(f'_n(x) - \varphi'(x))^2 = 0$$

$2n$ Wurzeln gemein. Die übrig bleibenden Wurzeln von $(f_n(x) - \varphi(x))^2 - L^2 = 0$ müssen nach dem Obigen $x = \pm 1$ sein. Also hat

$$(f_n(x) - \varphi(x))^2 - L^2 = 0$$

mit

$$(x^2 - 1) (f'_n(x) - \varphi'(x))^2 = 0$$

$2n + 2$ d. h. alle Wurzeln gemein, und wenn wir bedenken, dass in der ersten Gleichung der Koeffizient des höchsten Gliedes gleich 1, in der zweiten gleich $(n+1)^2$ ist, so haben wir, wenn wir kurz

$$\varphi(x) - f_n(x) = x^{n+1} - f_n(x) = y$$

setzen, die Differentialgleichung:

$$y^2 - L^2 = \frac{1}{(n+1)^2} (x^2 - 1) \left(\frac{dy}{dx} \right)^2$$

die durch Trennung der Variabeln ergibt:

$$\frac{dy}{\sqrt{L^2 - y}} = (n+1) \frac{dx}{\sqrt{1 - x^2}}$$

$$\arcsin \frac{y}{L} = (n+1) \arcsin x + C.$$

Hieraus ersieht man: Wenn x das Intervall $-1 \dots +1$ durch-

läuft, $\arcsin x$ also von $-\frac{\pi}{2} \dots + \frac{\pi}{2}$ und $(n+1) \arcsin x$ von $-\frac{n+1}{2} \pi$ bis $\frac{n+1}{2} \pi$ wandert, so durchwandert die rechte Seite ein Intervall von $(n+1)\pi$, daher schwankt $\frac{y}{L}$ $(n+1)$ mal von -1 zu $+1$ und y von $-L$ zu $+L$.

Um C zu bestimmen, beachten wir die obige Bemerkung, dass für $x = \mp 1$ y gleich $\pm L$ sein muss. Bei geradem n wird dem durch $C = 0$, bei ungeradem durch $C = \frac{\pi}{2}$ genügt, wobei, worauf es aber nicht ankommt, C nur bis auf Vielfache von π bestimmt ist.

1) n sei gerade; es sei $x = \sin z$

$$\arcsin \frac{y}{L} = (n+1) z$$

$$\frac{y}{L} = \sin (n+1) z$$

$$\frac{y}{L} = \text{dem imag. Teil von } \cos (n+1) z + i \sin (n+1) z$$

$$= \text{,, ,, ,, ,, } (\cos z + i \sin z)^{n+1}.$$

Wir ziehen von diesem Ausdruck den konjugierten ab und dividieren durch $2i$:

$$\begin{aligned} \frac{y}{L} &= \frac{(\cos z + i \sin z)^{n+1} - (\cos z - i \sin z)^{n+1}}{2i} \\ &= \pm \frac{1}{2} ((\sin z + i \cos z)^{n+1} + (\sin z - i \cos z)^{n+1}) \\ &= \pm \frac{1}{2} ((x + \sqrt{x^2 - 1})^{n+1} + (x - \sqrt{x^2 - 1})^{n+1}) \end{aligned}$$

der Koeffizient von x^{n+1} ist dabei gleich der Summe der an ungerader Stelle stehenden $(n+1)$ ten Binomialkoeffizienten. Diese finden wir, wenn wir $(1-1)^{n+1}$ formal entwickeln als $\frac{2^{n+1}}{2} = 2^n$.

Da nun der höchste Koeffizient in y gleich 1 sein muss, so folgt

$$L = \mp \frac{1}{2^n}$$

$$y = \left(\frac{x + \sqrt{x^2 - 1}}{2} \right)^{n+1} + \left(\frac{x - \sqrt{x^2 - 1}}{2} \right)^{n+1}$$

2) n sei ungerade;

$$\begin{aligned} \frac{y}{L} &= \sin \left(\frac{\pi}{2} + (n+1)z \right) \\ &= \cos (n+1)z \\ &= \text{dem reellen Teil von } \cos (n+1)z + i \sin (n+1)z \\ &= \frac{1}{2} ((\cos z + i \sin z)^{n+1} + (\cos z - i \sin z)^{n+1}) \\ &= \mp \frac{1}{2} ((x + \sqrt{x^2 - 1})^{n+1} + (x - \sqrt{x^2 - 1})^{n+1}) \end{aligned}$$

d. h. wir erhalten dasselbe Resultat wie oben.

$$(3) \quad y = \left(\frac{x + \sqrt{x^2 - 1}}{2} \right)^{n+1} + \left(\frac{x - \sqrt{x^2 - 1}}{2} \right)^{n+1}$$

ist diejenige ganze rationale Funktion $(n+1)$ ten Grades, die sich von -1 bis $+1$ am wenigsten von 0 entfernt und 1 zum höchsten Koeffizienten hat.

$x^{n+1} - y$ ist diejenige Funktion n ten Grades, die x^{n+1} am besten annähert. Man sieht, dass in (3) Glieder n ten Grades nicht vorkommen, wie dies auch aus unserer Bemerkung am Schluss von § 4 des vorigen Kapitels folgt. Daher ist $x^{n+1} - y$ nur vom $(n-1)$ ten Grade, und daher sicher diejenige Funktion $(n-1)$ ten Grades, die x^{n+1} am besten annähert.

Diejenige Funktion n ten Grades, die x^{n+2} am besten annähert, ist

$$x^{n+2} - \left[\left(\frac{x + \sqrt{x^2 - 1}}{2} \right)^{n+2} + \left(\frac{x - \sqrt{x^2 - 1}}{2} \right)^{n+2} \right].$$

§ 3. Die Annäherungsfunktion beliebiger analytischer Funktionen.

1. Analytischer Charakter der Koeffizienten.

Nachdem wir die Annäherungsfunktion für die einfachsten Fälle $\varphi(x) = x^{n+1}$ und $\varphi(x) = x^{n+2}$ aufgestellt haben, gehen

wir zu dem allgemeinen Fall der Darstellung analytischer Funktionen über. Wir werden dann freilich keine endliche geschlossene Form, sondern nur ein konvergentes Verfahren erwarten dürfen, das eine annäherungsweise Darstellung der eindeutig definierten Annäherungsfunktion mit jeder beliebigen Genauigkeit gestattet. Dabei ist der Grundgedanke: Das Intervall, durch das die Annäherungsfunktion definiert ist, als klein vorauszusetzen, die Intervallgrenzen als Parameter einzuführen, und die Annäherungsfunktion nach Potenzen dieses Parameters zu entwickeln.

Nehmen wir der Einfachheit halber an, die Intervallgrenzen seien $-h$ und $+h$, und die darzustellende Funktion $\varphi(x)$, von der Konvergenz in diesem Intervall vorausgesetzt wird, sei

$$\varphi(x) = k_0 + k_1x + \dots + k_nx^n + k_{n+1}x^{n+1} + \dots,$$

wo wir weiter noch voraussetzen wollen, dass k_{n+1} und k_{n+2} nicht beide zugleich verschwinden. Wir setzen

$$x = hz,$$

sodass das definierende Intervall nunmehr stets

$$z = -1 \dots +1.$$

Es wird

$$\varphi(x) = k_0 + k_1hz + k_2h^2z^2 + \dots + k_nh^n z^n + k_{n+1}h^{n+1}z^{n+1} + \dots.$$

Aus der Definition der Annäherungsfunktion folgt, dass die Annäherungsfunktion von $\varphi(x)$ gleich ist

$$k_0 + k_1hz + \dots + k_nh^n z^n$$

vermehrt um die Annäherungsfunktion von

$$k_{n+1}h^{n+1}z^{n+1} + k_{n+2}h^{n+2}z^{n+2} + \dots$$

und dass letztere wiederum gleich ist (nach dem Schluss des § 1):

$$h^{n+1}V,$$

wo V die Annäherungsfunktion von

$$k_{n+1}z^{n+1} + k_{n+2}hz^{n+2} + k_{n+3}h^2z^{n+3} + \dots$$

bedeutet. Setzen wir

$$k_0 + k_1 h z + \dots + k_n h^n z^n = \varphi_n(z),$$

so hat also die gesuchte Annäherungsfunktion die Form

$$\varphi_n(z) + h^{n+1} V.$$

Ich behaupte nun:

V lässt sich in eine in einem endlichen Intervall h konvergierende Reihe

$$V = V_0 + h V_1 + h^2 V_2 + \dots$$

entwickeln, deren Koeffizienten Polynome n ten Grades von z sind, deren Koeffizienten nur noch von den k abhängen.

Ist

$$V = p_n z^n + p_{n-1} z^{n-1} + \dots + p_0,$$

so lauten die Bedingungen für V nach (1):

$$\begin{aligned} & p_n z_1^n + p_{n-1} z_1^{n-1} + \dots + p_0 - (k_{n+1} z_1^{n+1} + h k_{n+2} z_1^{n+2} + \dots) - L = 0 \\ (4) \quad & p_n z_2^n + p_{n-1} z_2^{n-1} + \dots + p_0 - (k_{n+1} z_2^{n+1} + h k_{n+2} z_2^{n+2} + \dots) + L = 0 \\ & \vdots \\ & p_n z_{n+2}^n + p_{n-1} z_{n+2}^{n-1} + \dots + p_0 - (k_{n+1} z_{n+2}^{n+1} + h k_{n+2} z_{n+2}^{n+2} + \dots) + (-1)^{n+2} L = 0 \\ & (z_1 + 1) \frac{\partial}{\partial z_1} (p_n z_1^n + p_{n-1} z_1^{n-1} + \dots + p_0 - (k_{n+1} z_1^{n+1} + h k_{n+2} z_1^{n+2} + \dots)) = 0 \\ (5) \quad & \frac{\partial}{\partial z_2} (p_n z_2^n + p_{n-1} z_2^{n-1} + \dots + p_0 - (k_{n+1} z_2^{n+1} + h k_{n+2} z_2^{n+2} + \dots)) = 0 \\ & \vdots \\ & (z_{n+2} - 1) \frac{\partial}{\partial z_{n+2}} (p_n z_{n+2}^n + p_{n-1} z_{n+2}^{n-1} + \dots + p_0 - (k_{n+1} z_{n+2}^{n+1} + h k_{n+2} z_{n+2}^{n+2} + \dots)) = 0 \end{aligned}$$

die wir abgekürzt

$$\begin{aligned} F_1 - L &= 0 \\ F_2 + L &= 0 \dots \\ (z_1 + 1) \frac{\partial F_1}{\partial z_1} &= 0 \\ \frac{\partial F_1}{\partial z_2} &= 0 \dots \end{aligned}$$

schreiben wollen.

Für $h = 0$ haben wir die Gleichungen berechnet. Es ist dort $\frac{\partial F_1}{\partial z_1} \neq 0$ $\frac{\partial F_{n+2}}{\partial z_{n+2}} \neq 0$. Wegen der Stetigkeit muss dies auch noch für ein endliches Intervall h der Fall sein. Der Einfachheit halber lassen wir also die erste und letzte der Gleichungen (5) weg und ersetzen sie durch $z_1 = -1$, $z_{n+2} = +1$. Die Funktionaldeterminante der übrigen $2n + 2$ Gleichungen mit den $2n + 2$ Unbekannten

$$p_n, p_{n-1}, \dots, p_0, L, z_2, z_3, \dots, z_{n+1}$$

lautet jetzt:

$\frac{\partial F_1}{\partial p_n}$	$\frac{\partial F_1}{\partial p_{n-1}}$	\dots	$\frac{\partial F_1}{\partial p_0}$	$\frac{\partial F_1}{\partial L}$	$\frac{\partial F_1}{\partial z_2}$	$\frac{\partial F_1}{\partial z_3}$	\dots	$\frac{\partial F_1}{\partial z_{n+1}}$
$\frac{\partial F_2}{\partial p_n}$	$\frac{\partial F_2}{\partial p_{n-1}}$	\dots	$\frac{\partial F_2}{\partial p_0}$	$\frac{\partial F_2}{\partial L}$	$\frac{\partial F_2}{\partial z_2}$	$\frac{\partial F_2}{\partial z_3}$	\dots	$\frac{\partial F_2}{\partial z_{n+1}}$
\vdots								
$\frac{\partial F_{n+2}}{\partial p_n}$	$\frac{\partial F_{n+2}}{\partial p_{n-1}}$	\dots	$\frac{\partial F_{n+2}}{\partial p_0}$	$\frac{\partial F_{n+2}}{\partial L}$	$\frac{\partial F_{n+2}}{\partial z_2}$	$\frac{\partial F_{n+2}}{\partial z_3}$	\dots	$\frac{\partial F_{n+2}}{\partial z_{n+1}}$
$\frac{\partial^2 F_2}{\partial z_2 \partial p_n}$	$\frac{\partial^2 F_2}{\partial z_2 \partial p_{n-1}}$	\dots	$\frac{\partial^2 F_2}{\partial z_2 \partial p_0}$	$\frac{\partial^2 F_2}{\partial z_2 \partial L}$	$\frac{\partial^2 F_2}{\partial z_2^2}$	$\frac{\partial^2 F_2}{\partial z_2 \partial z_3}$	\dots	$\frac{\partial^2 F_2}{\partial z_2 \partial z_{n+1}}$
\vdots								
$\frac{\partial^2 F_{n+1}}{\partial z_{n+1} \partial p_n}$	$\frac{\partial^2 F_{n+1}}{\partial z_{n+1} \partial p_{n-1}}$	\dots	$\frac{\partial^2 F_{n+1}}{\partial z_{n+1} \partial p_0}$	$\frac{\partial^2 F_{n+1}}{\partial z_{n+1} \partial L}$	$\frac{\partial^2 F_{n+1}}{\partial z_{n+1} \partial z_2}$	$\frac{\partial^2 F_{n+1}}{\partial z_{n+1} \partial z_3}$	\dots	$\frac{\partial^2 F_{n+1}}{\partial z_{n+1}^2}$

Man sieht leicht, dass alle Elemente in dem Rechteck rechts oben verschwinden; denn es ist z. B. in der zweiten Zeile $\frac{\partial F_2}{\partial z_2} = 0$ nach (5) und $\frac{\partial F_2}{\partial z_3} = \frac{\partial F_2}{\partial z_4} = \dots = 0$, weil $z_3, z_4 \dots$ in E_2 nicht vorkommen. Aus letzterem Grunde verschwinden auch die Terme des unteren rechten Rechtecks mit Ausnahme der Diagonalglieder. Die Determinante ist demnach gleich:

$$\begin{vmatrix} z_1^n & z_1^{n-1} & \dots & z_1 & 1 & -1 \\ z_2^n & z_2^{n-1} & \dots & z_2 & 1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ z_{n+2}^n & z_{n+2}^{n-1} & \dots & z_{n+2} & 1 & (-1)^{n+2} \end{vmatrix} \cdot \frac{\partial^2 F_2}{\partial z_2^2} \cdot \frac{\partial^2 F_3}{\partial z_3^2} \cdot \dots \cdot \frac{\partial^2 F_{n+1}}{\partial z_{n+1}^2}.$$

Diese Determinante ist nach S. 29 von 0 verschieden, und die zweiten Ableitungen sind es jedenfalls für $h = 0$ auch.

Aus dem Nichtverschwinden der Funktionaldeterminante folgt, dass sich alle $2n + 2$ Grössen $p_n p_{n-1} \dots p_0 L z_1 \dots z_{n+1}$ nach Potenzen von h entwickeln lassen. Schreiten wir auf der reellen positiven Axe h weiter, so sind bis zu dem nächsten singulären Punkt alle diese Grössen eindeutig als Funktionen von h definiert. Es fragt sich aber, inwieweit durch sie unsere Annäherungsfunktion auch wirklich dargestellt wird. Denn die Gleichungen (4) und (5) sind ja nur notwendig, aber nicht hinreichend. Wir verfolgen den weiteren Verlauf in den z , weil die p sich unter allen Umständen eindeutig aus den z bestimmen.

Für $h = 0$ haben die beiden Gleichungen

$$F \pm L = 0 \quad \frac{\partial F}{\partial z} = 0$$

keine andern Wurzeln als die n : z_1, z_2, \dots, z_{n+1} gemeinsam. Dies sind zugleich alle Wurzeln von $\frac{\partial F}{\partial z} = 0$ für $h = 0$. Alle andern Wurzeln dieser Gleichung werden für $h = 0$ unendlich. Wir grenzen nun auf der reellen positiven Axe von h ein Gebiet so ab: Innerhalb dieses Gebietes soll $\frac{\partial F}{\partial z} = 0$ keine andern reellen Wurzeln zwischen -1 und $+1$ haben als diese n , oder es soll doch $|F|$ an den neu hinzukommenden Wurzeln kleiner sein als $|L|$. An der Grenze des Gebietes liegt dann eine Wurzel von $\frac{\partial F}{\partial z} = 0$, für die $F = \pm L$.

Innerhalb dieses Gebietes in h stellt die durch die Gleichungen (4) und (5) definierte analytische Funktion in der That unsere Annäherungsfunktion dar. Denn erstens wird L ($n + 2$) Mal im Intervall $z = -1 \dots +1$ angenommen. Denn da für $z = z_1 = -1$ $F = +L$ und für $z = z_2$ $F = -L$, und ebenso an dem positiven Ende des Intervalls, so ist klar, dass für den betrachteten Bereich in h unsere n Wurzeln im Intervall $-1 < z < 1$ bleiben müssen. Sodann kann auch für diesen Bereich $|F|$ im Intervall $-1 \leq z \leq 1$ nicht grösser werden als $|L|$. Denn dann müsste $|F|$ ein Maximum haben, für das es grösser wäre als L , was ja nach Definition ausgeschlossen ist. Hierdurch sind aber

auch die hinreichenden Bedingungen der Annäherungsfunktion erfüllt.

An der Grenze des Gebietes haben wir nun eine neue gemeinsame Wurzel von $F = \pm L$ und $\frac{\partial F}{\partial z} = 0$; z. B. kann $\frac{\partial F}{\partial z}$ für $z = \pm 1$ verschwinden. Lassen wir h darüber hinaus wachsen, so kann nun für diese Wurzel von $\frac{\partial F}{\partial z} = 0$ $|F|$ grösser werden als L , und dann stellt die so durch (4) und (5) definierte Funktion die Annäherungsfunktion nicht mehr dar. Wir haben nun zwei Systeme von Gleichungen, denen die Annäherungsfunktion genügen kann; das alte, dem sie für kleinere h genügt, und das neue, das dadurch entsteht, dass wir die neue gemeinsame Wurzel von $F = \pm L$ und $\frac{\partial F}{\partial z} = 0$ an Stelle der ihr im alten System benachbarten unter die

$$z_1 z_2 \dots z_{n+2}$$

einsetzen, was z. B. auch so geschehen kann, dass wir $\frac{\partial F_1}{\partial z_1} = 0$ statt $z_1 = -1$ einsetzen. Nach unsern Existenz- und Eindeutigkeitssätzen muss eines und nur eines dieser Gleichungssysteme die Annäherungsfunktion liefern. Dieses verfolgen wir auf der reellen Axe der h weiter, bis wir wieder an eine derartige ausgezeichnete Stelle gelangen. Die p sind auch an solchen Stellen stetige Funktionen von h , von den z aber erleidet dabei eines in der Regel einen Sprung.

Ausser diesen Sprungstellen kommen noch irreguläre Punkte im funktionentheoretischen Sinn in Betracht, bei denen die Konvergenz der Reihen in h aufhört, was bei den eben besprochenen Punkten im Allgemeinen nicht der Fall ist. Betrachten wir die Funktionaldeterminante, so sieht man, dass diese nur dann gleich 0 wird, wenn der Faktor $\frac{\partial^2 F}{\partial z_2^2} \cdot \frac{\partial^2 F}{\partial z_3^2} \dots \frac{\partial^2 F}{\partial z_{n+1}^2}$, zu dem $\frac{\partial^2 F}{\partial z_1^2}$ und $\frac{\partial^2 F}{\partial z_{n+2}^2}$ noch hinzukommen kann, verschwindet. Die zweite Ableitung an diesen Stellen kann nur verschwinden, wenn zugleich auch die dritte verschwindet, weil diese Stellen ja wirkliche Maxima bzw. Minima sein müssen. Verschwinden aber diese Ableitungen, so hat an dieser Stelle $\frac{\partial F}{\partial z} = 0$ mit $F \pm L = 0$ eine Doppelwurzel

gemeinsam, die betreffende Grösse z also einen Verzweigungspunkt. Alsdann ist wie oben zu untersuchen, welcher der beiden Zweige im weitem Verlauf die Annäherungsfunktion darstellt. Ist die eine der beiden zusammenfallenden gemeinsamen Wurzeln von $F \pm L = 0$ und $\frac{\partial F}{\partial z} = 0$ im Allgemeinen komplex und nur gerade für den Verzweigungspunkt reell, so kann sie für die Annäherungsfunktion nicht in Betracht kommen.

Die Koeffizienten der Annäherungsfunktion sind abteilungsweise analytische Funktionen der Intervallgrenze h . Die einzigen hierbei vorkommenden Irregularitäten sind die Sprungstellen, an denen die Koeffizienten von einer analytischen Funktion zu einer andern, der ersten im funktionentheoretischen Sinne völlig fremden übergehen, und die Verzweigungspunkte, die dort liegen, wo $F = \pm L$ eine vierfache Wurzel hat.

2. Algorithmus der Ausrechnung.

Bestimmung von V_1 .

Wir setzen:

$$V = V_0 + hV_1 + h^2V_2 + h^3V_3 + \dots$$

$$L = L_0 + hL_1 + h^2L_2 + h^3L_3 + \dots$$

und ferner:

$$V_1 + hV_2 + h^2V_3 + \dots = \bar{V}_1$$

$$V_2 + hV_3 + h^2V_4 + \dots = \bar{V}_2$$

$$\dots \dots \dots$$

$$L_1 + hL_2 + h^2L_3 + \dots = \bar{L}_1$$

$$L_2 + hL_3 + h^2L_4 + \dots = \bar{L}_2,$$

so dass die V aus den \bar{V} und die L aus den \bar{L} durch Nullsetzung von h erhalten werden.

$$V = V_0 + h\bar{V}_1$$

ist die Annäherungsfunktion von

$$k_{n+1}z^{n+1} + hk_{n+2}z^{n+2} + h^2k_{n+3}z^{n+3} + \dots,$$

die aber für $h = 0$ in die Annäherungsfunktion von $k_{n+1}z^{n+1}$

übergeht, die uns aus dem vorigen Paragraphen bekannt ist. Es ist nämlich

$$V_0 = k_{n+1}(z^{n+1} - y),$$

wenn wir nach (3) unter y verstehen:

$$y = \left(\frac{z + \sqrt{z^2 - 1}}{2} \right)^{n+1} + \left(\frac{z - \sqrt{z^2 - 1}}{2} \right)^{n+1}.$$

Es ist also

$$V = k_{n+1}(z^{n+1} - y) + h\bar{V}_1.$$

Diese Funktion muss den einer Annäherungsfunktion charakteristischen Gleichungen (2) genügen:

$$[k_{n+1}(z^{n+1} - y) + h\bar{V}_1 - (k_{n+1}z^{n+1} + hk_{n+2}z^{n+2} + \dots)]^2 - L^2 = 0$$

$$(z^2 - 1) \frac{d}{dz} [k_{n+1}(z^{n+1} - y) + h\bar{V}_1 - (k_{n+1}z^{n+1} + hk_{n+2}z^{n+2} + \dots)] = 0$$

oder

$$(6) \quad (hk_{n+2}z^{n+2} + h^2k_{n+3}z^{n+3} + \dots + k_{n+1}y - h\bar{V}_1)^2 - L^2 = 0$$

$$(7) \quad (z^2 - 1) \frac{d}{dz} (hk_{n+2}z^{n+2} + h^2k_{n+3}z^{n+3} + \dots + k_{n+1}y - h\bar{V}_1) = 0$$

die $(n+2)$ Wurzeln miteinander gemein haben sollen. Statt der Gleichung (7) betrachte ich nun die einfachere:

$$(z^2 - 1) \frac{d}{dz} (k_{n+1}y) = 0$$

$$(8) \quad (z^2 - 1) \frac{dy}{dz} = 0$$

und behaupte: Ist \bar{z} eine Wurzel der Gleichung (8), so existiert eine Wurzel $\bar{z} + \mathfrak{B}(h)$ der Gleichung (7), wo \mathfrak{B} , wie stets im Folgenden, eine Potenzreihe ohne absolutes Glied bedeutet. In der That, denken wir uns irgend eine Wurzel von (7), so muss sie sich nach Potenzen von h entwickeln lassen, wenn nicht gerade $h = 0$ ein Verzweigungspunkt ist. Dies ist aber nicht der Fall, weil, wie wir wissen, die durch Nullsetzung von h aus (7) hervorgehende Gleichung (8) keine Doppelwurzeln hat. Das von h freie Glied ist natürlich der Wert der Wurzel für $h = 0$ d. h. \bar{z} . Dies Verfahren gilt bei den n Wurzeln von (7), die für $h = 0$

nicht unendlich werden, und ausserdem bei $z = \pm 1$, wo $\mathfrak{P}(h)$ identisch 0 ist, also bei $(n+2)$ Wurzeln. Es soll nun (6) durch $\bar{z} + \mathfrak{P}(h)$ befriedigt werden, und zwar identisch in h . Die Koeffizienten von h in (6) müssen also einzeln verschwinden, und ich muss folglich zu einer richtigen Gleichung kommen, wenn ich in (10), nachdem $\bar{z} + \mathfrak{P}(h)$ eingesetzt ist, die Koeffizienten von $h^2, h^3 \dots$ weglasse, und nur das absolute Glied und das mit h multiplicierte beibehalte. Ich kann dann, da alle Wurzeln $\bar{z} + \mathfrak{P}(h)$ nach (7) Doppelwurzeln von (6) sind, statt $\bar{z} + \mathfrak{P}(h)$ auch bloss \bar{z} einsetzen, und erhalte so: Es sollen:

$$(9) \quad (k_{n+1}^2 y^2 - L_0^2) + [2k_{n+1} y (k_{n+2} z^{n+2} - V_1) - 2L_1] h = 0$$

$$(8) \quad (z^2 - 1) \frac{dy}{dz} = 0$$

$(n+2)$ Wurzeln gemein haben. Nun wissen wir, dass für alle Wurzeln von (8) $k_{n+1}^2 y^2 - L_0^2 = 0$, und folglich haben auch folgende Gleichungen $n+2$ Wurzeln gemein:

$$(10) \quad -2hL_1 + 2k_{n+1} y (k_{n+2} z^{n+2} - V_1) h = 0$$

$$(8) \quad (z^2 - 1) \frac{dy}{dz} = 0.$$

Multiplicieren wir (10) mit y und setzen wieder $k_{n+1}^2 y^2 = L_0^2$ ein

$$(11) \quad -2hL_1 y + 2L_0^2 (k_{n+2} z^{n+2} - V_1) h = 0$$

$$(8) \quad (z^2 - 1) \frac{dy}{dz} = 0.$$

Diese Gleichungen sind beide vom Grade $(n+2)$; wenn sie $(n+2)$ Wurzeln gemein haben, so ist dies nur möglich wenn:

$$(12) \quad -2hL_1 y + 2L_0^2 (k_{n+2} z^{n+2} - V_1) h - C(z^2 - 1) \frac{dy}{dz} = 0$$

wo C von z , nicht aber von h unabhängig ist. In (12) kommt ein Glied mit z^{n+1} nur im ersten Term vor, und da (12) identisch in z gilt, folgt

$$L_1 = 0$$

$$(13) \quad V_1 = k_{n+2} z^{n+2} - \frac{C}{2hL_0^2} (z^2 - 1) \frac{dy}{dz}.$$

V_1 muss eine Funktion n ten Grades in z sein; der Koeffizient von z^{n+2} muss verschwinden; daher

$$(14) \quad C = \frac{2hL_0^2 k_{n+2}}{n+1}.$$

Wir erhalten somit als Schlussresultat:

$$(15) \quad V_1 = k_{n+2} \left(z^{n+2} - \frac{z^2 - 1}{n+1} \frac{dy}{dz} \right).$$

Diese Formel gilt, da oben bei Aufstellung von (8) durch k_{n+1} dividiert wurde, nur für $k_{n+1} \neq 0$. Die erhaltene Formel (15) ist nicht identisch mit der Annäherungsfunktion für $k_{n+2} z^{n+2}$. Die Annäherungsfunktion einer Summe ist nicht identisch mit der Summe der Annäherungsfunktionen.

Bestimmung der höheren Koeffizienten von V .

Wir erweitern die vorige Methode, sodass sie gestattet, die höheren Koeffizienten von

$$V = V_0 + hV_1 + h^2V_2 + \dots$$

zu berechnen, sobald die niedrigeren, die aber nicht 0 sein dürfen, bekannt sind. Diese Methode wird also auch bei $k_{n+1} = 0$ wenn $k_{n+2} \neq 0$ zum Ziel führen.

Es seien V_0, V_1, \dots, V_{i-1} gefunden, und wir wollen die folgenden Koeffizienten bis V_{2i-1} berechnen. Wir führen folgende Abkürzungen ein:

$$V_0 + hV_1 + \dots + h^{i-1} = V_i$$

$$k_{n+1} z^{n+1} + hk_{n+2} z^{n+2} + \dots + k_{n+i} h^{i-1} z^{n+i} - V_i = y_i$$

$$V_i + hV_{i+1} + h^2V_{i+2} + \dots + h^{i-1}V_{2i-1} = V^{(i)}$$

$$k_{n+i+1} z^{n+i+1} + hk_{n+i+2} z^{n+i+2} + \dots + k_{n+2i} h^{i-1} z^{n+2i} = \varphi^{(i)}$$

$$k_{n+2i+1} z^{n+2i+1} + hk_{n+2i+2} z^{n+2i+2} + \dots = \bar{\varphi}_{2i}.$$

Die für V charakteristischen Gleichungen lauten alsdann: Es sollen

$$[y_i + h^i \varphi^{(i)} + h^{2i} \bar{\varphi}_{2i} - h^i V^{(i)} - h^{2i} \bar{V}_{2i}]^2 - L^2 = 0$$

$$(z^2 - 1) \frac{d}{dz} [y_i + h^i \varphi^{(i)} + h^{2i} \bar{\varphi}_{2i} - h^i V^{(i)} - h^{2i} \bar{V}_{2i}] = 0$$

$n + 2$ Wurzeln zwischen $-1 \dots + 1$ gemeinsam haben. Wir sehen jetzt wieder, dass die mit h^2 multiplizierten Terme keinen Einfluss auf das von uns gesuchte $V^{(n)}$, das die mittleren Potenzen von V , von V_l bis V_{2l-1} umfasst, haben können. Denken wir uns nämlich eine Wurzel der zweiten Gleichung nach Potenzen von h entwickelt, was ja möglich ist, weil diese Gleichung für $h = 0$ in $(z^2 - 1) \frac{dy}{dz} = 0$ übergeht, ein Verzweigungspunkt also nicht vorliegen kann, so können die Koeffizienten von h^2 auch erst auf die $2l$ ten Potenzen dieser Wurzelentwicklung einwirken, während die niedrigeren Potenzen mit denen der Wurzeln von

$$(z^2 - 1) \frac{d}{dz} (y_s + h^l \varphi^{(n)} - h^l V^{(n)}) = 0$$

übereinstimmen. Die Wurzeln der zweiten Gleichung in die erste eingesetzt müssen diese identisch in h befriedigen; wir können also die Potenzen $h^2 \dots$ weglassen. Dies ergibt, dass auch die beiden folgenden Gleichungen $n + 2$ Wurzeln zwischen -1 und $+1$ gemein haben müssen:

$$(16) \quad y_s^2 + 2y_s(\varphi^{(n)} - V^{(n)})h^l - L_s^2 - 2h^l L_s L^{(n)} = 0$$

$$(z^2 - 1) \frac{d}{dz} (y_s + h^l \varphi^{(n)} - h^l V^{(n)}) = 0$$

wo aber, was wir durch unsere Zeichen nicht genau ausdrücken können, in der ersten Gleichung die noch darin enthaltenen Potenzen $h^2 \dots$ wegzulassen sind. Wir werden jedoch sehen, dass es hierauf nicht ankommt.

Wir gehen in der Vereinfachung noch weiter, indem wir die zweite Gleichung (16) durch eine Gleichung

$$(17) \quad (z^2 - 1) W = 0$$

ersetzen, die folgendermassen definiert ist: Die Wurzeln der zweiten Gleichung (16) sollen, soweit sie für $h = 0$ endlich bleiben, nach Potenzen von h entwickelt, in den ersten l Potenzen (h^l ausschliesslich) mit den Wurzeln von (17) übereinstimmen. (17) soll keine andern Wurzeln haben als diese.

Es ist zunächst zu zeigen, wie ganz allgemein eine solche Funktion W zu finden ist. Es sei etwa die Gleichung

$$u + hv = 0$$

gegeben, wo u und v ganze rationale Funktionen von x , u von h unabhängig und v eine ganze rationale Funktion von h ist. Der Grad in x von v sei höher als der von u , sodass für $h = 0$ mindestens eine Wurzel von $u + hv = 0$ unendlich wird. $u = 0$ habe keine Doppelwurzeln. Es ist die Aufgabe, eine Gleichung zu suchen, deren Wurzeln sich von denen von $u + hv = 0$ nur um Potenzen $h^n, h^{n+1} \dots$ unterscheiden, und deren Grad in x nicht höher ist als der von u .

Bei der vorausgesetzten Abwesenheit von Doppelwurzeln ist diese Bedingung mit der folgenden identisch: Die für $h = 0$ endlich bleibenden Wurzeln von $u + hv = 0$ sollen, in die linke Seite der gesuchten Gleichung eingesetzt, eine Potenzentwicklung erst mit dem m ten Glied beginnen lassen.

Die Wurzeln von $u + hv = 0$ lassen in

$$u \tag{I}$$

eingesetzt, die von h unabhängigen Glieder verschwinden. Für $m = 1$ ist $u = 0$ die gesuchte Gleichung; ist $m > 1$, so dividiere ich v durch u

$$v = qu + R.$$

R ist sicher nicht von höherem Grade als u . Die Wurzeln von $u + hv = 0$ lassen nun, in

$$u + hR \tag{II}$$

eingesetzt, das absolute und das h proportionale Glied verschwinden; denn $u + hR$ ist gleich

$$u + hv - hqu.$$

Da $u + hR$ von nicht höherem Grade ist als u , so ist es für $m = 2$ die verlangte Funktion. Für $m > 2$ dividiere ich v durch $u + hR$

$$v = q_1(u + hR) + R_1$$

und setze die Wurzeln von $u + hv = 0$ in

$$u + hR_1 \tag{III}$$

ein, worauf auch die quadratischen Glieder von h verschwinden, da $u + hR_1$ gleich ist

$$u + hv - hq_1(u + hR).$$

Der Rest der Division von v durch (III) liefert uns (IV), und es ist ersichtlich, dass wir nach $(m-1)$ Divisionen eine Funktion

$$u + hR_{m-2}$$

erhalten, die den verlangten Bedingungen genügt.

Denken wir uns nun aus der zweiten der Gleichungen (16) W bestimmt und beachten wir, dass die zweite Gleichung ausdrückt, dass die in Frage kommenden Wurzeln Doppelwurzeln der ersten Gleichung sind, so kommen wir zu dem Satz: Die Wurzeln von

$$(17) \quad (z^2-1) W = 0$$

ergeben, in die erste der Gleichungen (16)

$$y_g^2 + 2y_g(\varphi^{(0)} - V^{(0)})h^l - L_g^2 - 2h^l L_g L_g^{(0)}$$

eingesetzt, eine Potenzreihe von h , die mit h^{2l} beginnt. Multiplizieren wir mit y_g :

$$(18) \quad (y_g^2 - L_g^2)y_g + 2y_g^2(\varphi^{(0)} - V^{(0)})h^l - 2y_g h^l L_g L_g^{(0)}.$$

Aus (16) ersehen wir, dass für die Wurzeln der zweiten Gleichung (16) und daher auch für die Wurzeln von (17) die beiden Potenzreihen

$$y_g^2 \quad \text{und} \quad L_g^2$$

sich nur um Glieder l ter Ordnung unterscheiden, wodurch also in (18) sich nur Glieder $2l$ ter Ordnung ändern, wenn wir $h^l y_g^2$ durch $h^l L_g^2$ ersetzen. Wir erhalten: Für die Wurzeln von

$$(z^2-1) W = 0$$

darf

$$(y_g^2 - L_g^2)y_g + 2L_g^2(\varphi^{(0)} - V^{(0)})h^l - 2y_g h^l L_g L_g^{(0)}$$

oder, da wir durch L_g^2 dividieren dürfen, da dies für $h = 0$ nicht verschwindet,

$$(19) \quad V^{(0)}h^l + \frac{L_g}{L_g^{(0)}} y_g \cdot h^l - \varphi^{(0)}h^l - \frac{(y_g^2 - L_g^2)}{2L_g^2} y_g$$

erst mit h^{2l} beginnen. Dividieren wir (19) durch $(z^2-1)W$

$$(19) = q(z^2-1)W + R,$$

so darf der Rest R gleichfalls erst mit dem Gliede h^{2l} beginnen;

denn für alle Wurzeln von $(z^2 - 1)W = 0$ müssen die niedrigeren Potenzen von h in R verschwinden, d. h., da der Grad in z von $(z^2 - 1)W$ höher ist als der von R , sie müssen identisch in z verschwinden. Führen wir die Division aus und bezeichnen den Rest der Division von

$$y, h^i \quad \text{durch} \quad (z^2 - 1)W \quad \text{mit} \quad R_0$$

$$\varphi^{(i)} h^i + \frac{y_j^2 - L_j^2}{2L_j^2} y_j \quad \text{„} \quad \text{„} \quad \text{„} \quad R_1,$$

so haben wir das Resultat, dass

$$V^{(i)} h^i + \frac{L^{(i)}}{L_j} R_0 - R_1$$

erst mit dem Gliede h^{2i} beginnen darf. Dividieren wir

$$R_1 \quad \text{durch} \quad R_0$$

$$R_1 = q_0 R_0 + r,$$

so wird nichts geändert, und wir haben:

$$(20) \quad V^{(i)} h^i - r - R_0 \left(q_0 - \frac{L^{(i)}}{L_j} \right) = \mathfrak{B}(h^{2i}).$$

Nun vergleichen wir die Grade in z : $(z^2 - 1)W$ ist vom Grade $n + 2$, folglich sind die Reste R_0 und R_1 höchstens vom Grade $n + 1$ und r höchstens vom Grade n .

Wir wollen zeigen, dass R_0 wirklich vom Grade $(n + 1)$ ist. Dividieren wir nämlich y_j durch $(z^2 - 1)W$, so ergibt dies für $h = 0$ die Division

$$y \quad \text{durch} \quad (z^2 - 1) \frac{dy}{dz}.$$

Da hierbei der Dividend vom Grade $n + 1$, der Divisor vom Grade $(n + 2)$ ist, so ist y der Rest der Division. R_0 hat also folgende Form:

$$R_0 = h^i y + \mathfrak{B}(h^{i+1})$$

d. h. R_0 hat ein mit h^i multipliciertes Glied z^{n+1} . Hieraus folgt, dass (20) nur richtig sein kann, wenn

$$(21) \quad q_0 - \frac{L^{(i)}}{L_j} = \mathfrak{B}(h^i),$$

und dann folgt aus (20)

$$(22) \quad V^{(n)}h^{(n)} = r + \mathfrak{B}(h^{2^n}).$$

Da nun

$$V^{(n)} = V_1 + hV_{1+1} + \dots + h^{l-1}V_{2^{l-1}}$$

keine höhern Potenzen von h als h^{l-1} , die linke Seite von (22) also keine höheren Potenzen von h als $h^{2^{l-1}}$ enthält, so ist durch (22) $V^{(n)}$ vollkommen bestimmt, indem wir in der Potenzreihe r bloss die höheren Potenzen von h^{2^i} an wegzulassen brauchen, um $V^{(n)}h^{(n)}$ zu finden. Man sieht auch leicht, dass, wie wir dies verlangen müssen, r mit dem Gliede h^l beginnt. Denn dies ist sowohl bei R_0 als bei R_1 , also auch bei dem Reste ihrer Division, bei r der Fall.

Die Gleichung (21), die wir auch, da L_1 ein von h freies Glied enthält

$$L^{(n)} = q_0L_1 + \mathfrak{B}(h^l)$$

schreiben können, liefert uns $L^{(n)}$.

Damit ist unsere Aufgabe, ein konvergentes Verfahren zur Aufstellung der Annäherungsfunktion einer beliebigen analytischen Funktion anzugeben, gelöst. Die Lösung geschieht durch eine Kette von Divisionen. Vom l ten zum $2l$ ten Grad der Annäherung gelangt man, indem man zunächst die Funktion W bildet, wozu nach unserm S. 45—46 angegebenen Verfahren $(l-1)$ Divisionen erforderlich sind. Zwei weitere Divisionen liefern R_0 und R_1 , und schliesslich liefert die Division von R_0 durch R_1 den Rest r , der die gesuchten Potenzen der Annäherungsfunktion, und den Quotienten q_0 , der die entsprechenden Potenzen der Abweichung bestimmt.

Nach diesem Verfahren kann auch die Annäherungsfunktion einer beliebigen nichtanalytischen Funktion $\varphi(x)$ bestimmt werden. Denn nach § 5 des vorigen Kapitels erhalten wir mit beliebiger Genauigkeit die Annäherungsfunktion von $\varphi(x)$, wenn wir die Annäherungsfunktion einer im ganzen Intervall hinreichend wenig von $\varphi(x)$ verschiedenen Funktion berechnen. Wir ersetzen also $\varphi(x)$ zunächst durch eine analytische Funktion, was mit jeder beliebigen Genauigkeit geschehen kann, und wenden auf diese unsern Algorithmus an.

Kapitel III. Rand- und Nebenbedingungen.

§ 1. Eine Randbedingung.

Es sei gefordert:

Eine ganze rationale Funktion

$$g_n(x) = p_n x^n + p_{n-1} x^{n-1} + \dots + p_1 x + p_0$$

zu finden, die an zwei gegebenen Intervallgrenzen gleich einer gegebenen eindeutigen Funktion $\varphi(x)$ wird, und deren Abweichung von $\varphi(x)$ ein Minimum ist.

Existenzbeweis: Es sei L_1 die Abweichung einer beliebigen ganzen rationalen Funktion n ten Grades, die der Randbedingung genügt. Es kommen für das Minimum also nur solche Funktionen in Frage, deren Abweichung kleiner oder gleich L_1 ist. Sehen wir von der Randbedingung ab, so erfüllen diese nach unserm algebraischen Hilfssatz S. 7 ein endliches $(n+1)$ -dimensionales Prisma. Aus diesem wird durch die Randbedingung ein $(n-1)$ -dimensionaler Bereich herausgeschnitten; denn wenn etwa $p_n p_{n-1} \dots p_2$ gegeben sind, so folgen durch die Randbedingung p_1 und p_0 stets eindeutig. In diesem endlichen $(n-1)$ -dimensionalen Bereich folgt aber die Existenz des Minimums nach dem Weierstrassschen Satz.

Wir nehmen $p_n p_{n-1} \dots p_2$ als frei veränderlich, p_1 und p_0 als abhängig an und haben, wenn $x_1 \dots x_2$ das gegebene Intervall, $\varphi(x_1) = a$ $\varphi(x_2) = b$

$$p_n x_1^n + p_{n-1} x_1^{n-1} + \dots + p_2 x_1^2 + p_1 x_1 + p_0 = a$$

$$p_n x_2^n + p_{n-1} x_2^{n-1} + \dots + p_2 x_2^2 + p_1 x_2 + p_0 = b$$

oder:

$$p_1 x_1 + p_0 = a - p_n x_1^n - p_{n-1} x_1^{n-1} - \dots - p_2 x_1^2$$

$$p_1 x_2 + p_0 = b - p_n x_2^n - p_{n-1} x_2^{n-1} - \dots - p_2 x_2^2$$

$$p_1 = c_n p_n + c_{n-1} p_{n-1} + \dots + c_2 p_2 + \gamma$$

$$p_2 = c'_n p_n + c'_{n-1} p_{n-1} + \dots + c'_2 p_2 + \gamma'$$

wobei

$$c_n = -\frac{x_1^n - x_2^n}{x_1 - x_2} \quad c'_n = -(x_1^n x_2 - x_2^n x_1)$$

$$\begin{array}{ll}
c_{n-1} = -\frac{x_1^{n-1} - x_2^{n-1}}{x_1 - x_2} & c'_{n-1} = -(x_1^{n-1}x_2 - x_2^{n-1}x) \\
\vdots & \vdots \\
c_2 = -\frac{x_1^2 - x_2^2}{x_1 - x_2} & c'_2 = -(x_1^2x_2 - x_2^2x_1) \\
\gamma = \frac{a-b}{x_1 - x_2} & \gamma' = ax_2 - bx_1.
\end{array}$$

Wir sehen also: Die Form einer ganzen rationalen Funktion n ten Grades, deren Willkürlichkeit nur dadurch beschränkt ist, dass $g_n(x_1) = a$, $g_n(x_2) = b$, ist diese:

$$\begin{aligned}
g_n(x) &= p_n x^n + p_{n-1} x^{n-1} + \dots + p_2 x^2 + (c_n p_n + c_{n-1} p_{n-1} + \dots + c_2 p_2 + \gamma)x \\
&\quad + c'_n p'_n + c'_{n-1} p'_{n-1} + \dots + \gamma' \\
&= (x^n + c_n x + c'_n) p_n + (x^{n-1} + c_{n-1} x + c'_{n-1}) p_{n-1} + \dots + (x^2 + c_2 x + c'_2) p_2 + \gamma x + \gamma'
\end{aligned}$$

wo $p_n p_{n-1} \dots p_2$ noch willkürlich sind.

Auf $g_n(x) - \varphi(x)$ wenden wir wieder unsern allgemeinen Satz S. 14 an. Die Bedingung (1) S. 12 besagt nun, wenn wir

$$(x^n + c_n x + c'_n) \lambda_n + (x^{n-1} + c_{n-1} x + c'_{n-1}) \lambda_{n-1} + \dots + (x^2 + c_2 x + c'_2) \lambda_2$$

mit $h(x)$ bezeichnen:

Es soll unmöglich sein, durch irgend ein System $\lambda_n \lambda_{n-1} \dots \lambda_2$ zu erreichen, dass $h(x)$ an ν zwischen x_1 und x_2 liegenden Punkten die Werte $s_1 s_2 \dots s_\nu$ annimmt.

$h(x)$ hat dieselbe Form wie $g_n(x)$, nur dass für $h(x)$ $\gamma = \gamma' = 0$ zu setzen ist, was dann und nur dann erreicht wird, wenn für $h(x)$ $a = b = 0$. $h(x)$ ist also eine Funktion n ten Grades, deren Willkürlichkeit nur dadurch beschränkt ist, dass x_1 und x_2 Wurzeln sind. Den Gleichungen (1) S. 12 kann stets dann genügt werden, wenn $s_1 s_2 \dots s_\nu$ ($n-2$) oder weniger Zeichenwechsel enthalten. Es sind mindestens ($n-1$) Zeichenwechsel dazu erforderlich, dass die Gleichungen unauflösbar sind.

Eine notwendige Bedingung für die Funktion kleinster Abweichung unter der gegebenen Randbedingung ist die, dass die Abweichung mindestens n mal in abwechselndem Sinn angenommen wird.

Der Beweis, dass diese Bedingung eindeutig und hinreichend ist, gestaltet sich dem oben S. 16—17 geführten ganz analog.

Die Abweichung der gefundenen Funktion $g_n(x)$ von $\varphi(x)$ sei L . Wenn wir nun nachweisen können, dass

$$|g_n(x) - \varphi(x)|,$$

nachdem es für die Intervallgrenzen 0 war, ausserhalb wieder den Werth L annehmen muss, und zwar in dem Sinn, dass wir hierdurch mit den im Intervall gelegenen Werten zusammen $(n+1)$ Zeichenwechsel erhalten, so ist in dem so erweiterten Intervall $g_n(x)$ mit der ohne Randbedingung aufgestellten Annäherungsfunktion identisch, und es ist also dann die Hinzunahme der Randbedingung mit einer Erweiterung des Intervalls gleichbedeutend. Dies ist z. B. sicher der Fall, wenn $\varphi(x) = x^{n+1}$.

§ 2. Eine Nebenbedingung.

Die Annäherungsfunktion soll monoton sein, d. h. in dem definierenden Intervall entweder nur wachsen, oder nur abnehmen. Wir betrachten nur den einfachsten Fall, in dem $\varphi(x)$ eine Potenz ist. Es soll diejenige von -1 bis $+1$ monotone ganze rationale Funktion der Form

$$F(x) = x^n + p_{n-1}x^{n-1} + p_{n-2}x^{n-2} + \dots + p_1x + p_0$$

gesucht werden, die zwischen -1 und $+1$ die Null möglichst gut annähert.

Existenzbeweis: Wir fassen wieder die Abweichung L als Funktion der Koeffizienten p_0, p_1, \dots, p_{n-1} auf, und es handelt sich darum nachzuweisen, dass diese Funktion ein Minimum hat. Wir grenzen in dem n -dimensionalen Raum der p ein Gebiet G so ab, dass ausserhalb G L so gross ist, dass wir uns für die Entscheidung der Frage ob L ein Minimum hat, auf G beschränken können. Ausserhalb G können auch noch Minima oder untere Grenzen liegen, wir beschränken uns aber auf das kleinste Minimum, bezw. die kleinste untere Grenze von L , die in G liegen muss.

Jeder Punkt von G repräsentiert eine Funktion der verlangten Form, aber diese Funktionen genügen nicht alle der Bedingung der Monotonität. Denken wir uns von den durch Punkte innerhalb G repräsentierten Funktionen die monotonen gesondert betrachtet, so werden die sie repräsentierenden Punkte ein Ge-

biet G_1 erfüllen, das einen Teil von G ausmacht. Es handelt sich um die Existenz eines Minimums von L in G_1 . Es lässt sich wieder der Weierstrasssche Satz von der Existenz eines Minimums einer stetigen Funktion im endlichen Gebiet anwenden, sobald wir wissen, dass die Grenzen des Gebietes G_1 zu diesem gerechnet werden dürfen.

Wir haben also zu entscheiden, ob die Grenzen von G_1 zu G_1 gerechnet werden dürfen oder nicht. Was bedeuten die Grenzen von G_1 ? Jeder auf einer solchen Grenze liegende Punkt scheidet Funktionen, die im Intervall $-1 \dots +1$ monoton sind, von nichtmonotonen Funktionen. Auf welche Weise können nun nichtmonotone Funktionen in monotone übergehen? Nichtmonotone Funktionen sind solche, deren Ableitung im betrachteten Intervall Wurzeln von ungerader Vielfachheit hat. Wenn eine Funktion diese Eigenschaft durch stetige Veränderung der Koeffizienten und also auch der Wurzeln der Ableitung verlieren soll, so ist dies offenbar nur entweder dadurch möglich, dass eine Wurzel aus dem betrachteten Intervall herausrückt im Grenzfall also in die Grenzen -1 und $+1$ hineinfällt, oder dadurch, dass zwei Wurzeln ungerader Vielfachheit zusammenfallen und eine Wurzel gerader Vielfachheit bilden. Die Grenzen von G_1 repräsentieren also Funktionen, deren Ableitungen Wurzeln von gerader Vielfachheit oder Wurzeln in den Grenzen -1 und $+1$ haben. Dies stört aber den monotonen Charakter der Funktion im Intervall $-1 \dots +1$ nicht, d. h. wir dürfen die Grenzen von G_1 zu G_1 rechnen. Die Existenz des Minimums ist damit bewiesen. Wir werden sehen, dass das Minimum in der That auf diesen Grenzen liegt, und zwar gerade da, wo sich möglichst viele Begrenzungsflächen von G_1 schneiden.

Reduktion der Aufgabe. Schreiben wir die gesuchte Funktion in der Form:

$$F(x) = \int_{-1}^x F'(x) dx + C,$$

so ergibt sich leicht die Konstante C ; es muss nämlich

$$F(-1) = -F(+1),$$

da sonst, wie man unmittelbar sieht, die Abweichung, die hier nur an den Intervallgrenzen angenommen werden kann, durch Addition einer Konstanten verkleinert werden könnte. Dies liefert:

$$\int_{-1}^{-1} F'(x) dx + C = -\int_{-1}^{+1} F'(x) dx - C$$

$$C = -\frac{1}{2} \int_{-1}^{+1} F'(x) dx$$

$$F(x) = \int_{-1}^x F'(x) dx - \frac{1}{2} \int_{-1}^{+1} F'(x) dx$$

die Abweichung L , absolut genommen, ist gleich $F(1)$

$$L = \frac{1}{2} \int_{-1}^{+1} F'(x) dx.$$

Wenn nun $F(x)$ die gesuchte monotone Funktion ist, so darf es keine andere monotone Funktion $\bar{F}(x)$ von der Form:

$$\bar{F}(x) = x^n + p_{n-1}x^{n-1} + \dots + p_0$$

geben, für die

$$\int_{-1}^{+1} \bar{F}'(x) dx$$

kleiner wäre als

$$\int_{-1}^{+1} F(x) dx.$$

Mit diesem Satz können wir die Behauptung beweisen:

Die Wurzeln von $F'(x) = 0$ liegen alle im Intervall $-1 \dots +1$.

Die im Intervall liegenden Wurzeln von $F'(x) = 0$ liegen entweder an den Grenzen oder sie sind von gerader Vielfachheit. Die im Innern des Intervalls liegenden Wurzeln seien

$$\alpha_1, \alpha_2, \dots, \alpha_m, \quad \text{und} \quad 2\lambda_1, 2\lambda_2, \dots, 2\lambda_m$$

sei der Grad ihrer Vielfachheit. 1 und -1 seien eine λ - bzw. λ_0 -fache Wurzel von $F'(x)$, wo λ und λ_0 auch 0 sein können. Daher ist

$$\frac{F'(x)}{(x-1)^\lambda (x+1)^{\lambda_0} (x-\alpha_1)^{2\lambda_1} (x-\alpha_2)^{2\lambda_2} \dots (x-\alpha_m)^{2\lambda_m}}$$

eine ganze Funktion von x , die zwischen -1 und $+1$ nicht verschwindet. Der absolut kleinste Betrag dieser Funktion im Intervall $-1 \dots +1$ mit dem ihm zukommenden Vorzeichen sei mit L_0 bezeichnet; jetzt ist

$$\frac{F'(x)}{(x-1)^\lambda (x+1)^{\lambda_0} \dots (x-\alpha_m)^{2\lambda_m}} - L_0$$

eine Funktion, die ihr Zeichen von -1 bis $+1$ nicht ändert, und die, da Minuend und Subtrahend von gleichem Vorzeichen sind, absolut genommen stets kleiner ist als der Minuend allein. Das gleiche gilt auch von der Funktion, die ich erhalte, wenn ich mit dem Nenner multipliziere, also von der Funktion

$$F'(x) - L_0(x-1)^\lambda \dots (x-\alpha_m)^{2\lambda_m};$$

hieraus folgt, dass

$$\int_{-1}^{+1} F'(x) - L_0(x-1)^\lambda (x+1)^{\lambda_0} \dots (x-\alpha_m)^{2\lambda_m} dx$$

absolut genommen kleiner ist als

$$\int_{-1}^{+1} F'(x) dx,$$

wobei

$$\int_{-1}^x F'(x) - L_0(x-1) \dots (x-\alpha_m)^{2\lambda_m} dx$$

eine monotone Funktion ist. Ist der Grad von

$$L_0(x-1)^\lambda (x+1)^{\lambda_0} \dots$$

kleiner als der von $F'(x)$, so hat sich die Form von $F'(x)$ durch die Subtraktion nicht geändert, und $F(x)$ ist nicht die gesuchte Funktion, da wir

$$\overline{F}(x) = F'(x) - L_0(x-1)^\lambda (x+1)^{\lambda_0} \dots (x-\alpha_m)^{2\lambda_m}$$

setzen können. Der Grad von $F'(x)$ muss daher gleich dem von $(x-1)^\lambda (x+1)^{\lambda_0} \dots (x-\alpha_m)^{2\lambda_m}$ sein, und wir können setzen

$$\begin{aligned} F'(x) &= n(x-1)^\lambda (x+1)^{\lambda_0} (x-\alpha_1)^{2\lambda_1} \dots (x-\alpha_m)^{2\lambda_m} \\ &= n(x-1)^\varrho (x+1)^{\varrho_0} U^\varrho, \end{aligned}$$

wo ϱ und ϱ_0 nur 0 oder 1 sein können, da höhere Potenzen in U hineingezogen werden. $\varrho + \varrho_0$ ist gerade oder ungerade, je nach dem vorgeschriebenen Grade n von $F(x)$; $\varrho = \begin{Bmatrix} 0 \\ 1 \end{Bmatrix}$ bestimmt das positive oder negative Vorzeichen von $F'(x)$ d. h. den wachsenden oder abnehmenden Charakter von $F(x)$ im Intervall $-1 \dots +1$. Daher sind ϱ und ϱ_0 als gegeben zu betrachten. Der Grad u von U ist $u = \frac{n-1-\varrho-\varrho_0}{2}$. Unsere Aufgabe ist somit

reduziert auf die, das Polynom U zu finden, für das

$$\int_{-1}^{+1} (1+x)^{\ell_0} (1-x)^{\ell} U^2 dx$$

ein Minimum wird.

Lösung der Aufgabe: Die Auflösung geschieht durch Funktionen

$$T_0 T_1 \dots T_n$$

die mit den gewöhnlichen Kugelfunktionen viel Aehnlichkeit haben und in einem speziellen Fall in sie übergehen. Wir definieren:

$$\varphi(s,x) = \frac{(1+s+\sqrt{1-2sx+s^2})^{\ell_0} (1-s+\sqrt{1-2sx+s^2})^{\ell}}{\sqrt{1-2sx+s^2}} = T_0 + T_1 s + T_2 s^2 + \dots$$

Man sieht leicht, dass T_m eine ganze Funktion m ten Grades von x ist. Man kann nun zeigen, (vergl. Tchebychef: Ueber die den Legendreschen Funktionen ähnlichen Funktionen) dass

$$\int_{-1}^{+1} \varphi(s,x) \varphi(t,s) (1+x)^{\ell_0} (1-x)^{\ell} dx,$$

gleich ist

$$\int_0^1 \frac{2^{-\ell-\ell_0+1} (1-stz)^{-\ell}}{z^{-\ell_0} (1-z)^{-\ell} (1-stz^2)} dz,$$

wo s und t nur noch in der Verbindung st vorkommen. Setzen wir links unsere Funktionen T ein, so wird das Integral:

$$\int_{-1}^{+1} (T_0 + T_1 s + T_2 s^2 + \dots) (T_0 + T_1 t + T_2 t^2 + \dots) (1+x)^{\ell_0} (1-x)^{\ell} dx.$$

Wenn dies Integral nur von st , nicht von s und t einzeln abhängen soll, so folgt

$$\int_{-1}^{+1} T_n T_m (1+x)^{\ell_0} (1-x)^{\ell} dx = 0 \quad \text{für} \quad n \neq m.$$

Durch diese Integralgleichung ist das System der Funktionen T im Wesentlichen (d. h. jede Funktion bis auf einen konstanten Faktor) eindeutig bestimmt.

Aus der Integralgleichung folgt für $n > m$

$$\int_{-1}^{+1} (1+x)^{\ell_0} (1-x)^{\ell} T_n (A T_{n-1} + B T_{n-2} + \dots + H T_0) dx = 0$$

wo $A, B \dots H$ beliebige Konstanten sind. Da nun jede Funktion $(n-1)$ ten Grades durch $AT_{n-1} + BT_{n-2} \dots + HT_0$ dargestellt werden kann, so hat T_n die Eigenschaft, mit jeder Funktion f_{n-1}

$$\int_{-1}^1 (1+x)^{e_0} (1-x)^e f_{n-1} T_n dx = 0$$

zu liefern. Nehmen wir an, \overline{T}_n habe dieselbe Eigenschaft; der höchste Koeffizient von T_n und \overline{T}_n sei als gleich angenommen, sodass $T_n - \overline{T}_n$ vom Grade $n-1$. Die Integraleigenschaft kommt, wenn T_n und \overline{T}_n , auch $T_n - \overline{T}_n$ zu, und wählen wir $f_{n-1} = T_n - \overline{T}_n$, so haben wir

$$\int_{-n}^1 (1+x)^{e_0} (1-x)^e (T_n - \overline{T}_n)^2 dx = 0.$$

Ist die Funktion \overline{T}_n reell, so hat die Funktion unter dem Integralzeichen stets gleiches Vorzeichen, und es folgt $\overline{T}_n = T_n$. Wird von \overline{T}_n Reellität nicht vorausgesetzt, so kommt die Integraleigenschaft dem reellen und imaginären Teil einzeln zu, und wir können den analogen Schluss ziehen.

Wir kehren jetzt zu der Aufgabe zurück, die Funktion U vom Grade u zu suchen, für die

$$\int_{-1}^{+1} (1+x)^{e_0} (1-x)^e U^2 dx$$

ein Minimum wird. Es lässt sich U in der Form schreiben

$$U = \overline{A} T_0 + \overline{B} T_1 + \dots + H T_u;$$

dividieren wir durch H : ($H \neq 0$)

$U = H(AT_0 + BT_1 + \dots + GT_{u-1} + T_u)$. Es soll

$$\int_{-1}^{+1} (1+x)^{e_0} (1-x)^e (AT_0 + BT_1 + \dots + GT_{u-1} + T_u)^2 dx = \text{minim.},$$

wobei $A, B \dots G$ zu bestimmen sind. Setzen wir nach der gewöhnlichen Regel die Ableitungen nach diesen Parametern gleich Null:

$$\int_{-1}^{+1} (1+x)^{e_0} (1-x)^e \cdot 2 \cdot (AT_0 + BT_1 + \dots + GT_{u-1} + T_u) T_0 \cdot dx = 0$$

$$\int_{-1}^{+1} (1+x)^{e_0} (1-x)^e \cdot 2 \cdot (AT_0 + BT_1 + \dots + GT_{u-1} + T_u) T_1 \cdot dx = 0$$

$$\dots \dots \dots \int_{-1}^{+1} (1+x)^{e_0} (1-x)^e \cdot 2 \cdot (AT_0 + BT_1 + \dots + GT_{u-1} + T_u) T_{u-1} \cdot dx = 0.$$

Benutzen wir die Integralgleichung der T , so folgt:

$$\begin{aligned} \int_{-1}^1 (1+x)^{\epsilon_0} (1-x)^{\epsilon} \cdot 2 \cdot A \cdot T_0^2 dx &= 0 \\ \int_{-1}^{+1} (1+x)^{\epsilon_0} (1-x)^{\epsilon} \cdot 2 \cdot B \cdot T_1^2 dx &= 0 \\ \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots & \end{aligned}$$

Da wieder eine monotone Funktion unter dem Integralzeichen steht, folgt

$$A = 0 \quad B = 0 \dots G = 0.$$

Die Existenz des Minimums folgt aus unserm Existenzbeweis in Verbindung mit der Eindeutigkeit der ganzen Entwicklung. Wir sehen dies aber auch direkt, wenn wir das zweite Glied der Taylorschen Entwicklung von

$$\int_{-1}^{+1} (1+x)^{\epsilon_0} (1-x)^{\epsilon} (AT_0 + BT_1 + \dots + T_u)^2 dx$$

nach $A, B, C \dots$ untersuchen. Dies Glied ist eine quadratische Form der Inkremente von $A, B \dots$, die hier direkt in der Normalform als Summe von Quadraten auftritt, da die gemischten Terme, die mit $\frac{\partial^2}{\partial A \partial B} \dots$ multipliziert sind, verschwinden. Die Koeffizienten dieser quadratischen Form

$$\begin{aligned} \int_{-1}^{+1} (1+x)^{\epsilon_0} (1-x)^{\epsilon} T_0^2 dx \\ \int_{-1}^{+1} (1+x)^{\epsilon_0} (1-x)^{\epsilon} T_1^2 dx \\ \dots \end{aligned}$$

haben offenbar alle gleiches Vorzeichen. Ein Minimum-Maximum ist demnach ausgeschlossen.

Die gesuchte Funktion U ist bis auf einen konstanten Faktor gleich T_u . Da U als höchsten Koeffizienten die Einheit hat, so ist es hierdurch eindeutig bestimmt.

Berechnung der Abweichung. Bezeichnen wir den höchsten Koeffizienten von T_u mit K_u , so ist

$$F(x) = \int_{-1}^x \frac{n}{K_u^2} (1+x)^{\epsilon_0} (1-x)^{\epsilon} T_u^2 dx - \frac{1}{2} \int_{-1}^{+1} \frac{n}{K_u^2} (1+x)^{\epsilon_0} (1-x)^{\epsilon} T_u^2 dx$$

$$L = \frac{n}{2K_u^2} \int_{-1}^{+1} (1+x)^{e_0} (1-x)^e T_u^n dx.$$

Bei der numerischen Ausrechnung dieses Integrals wollen wir uns auf Angabe der Resultate beschränken. Aus dem oben Gesagten folgt, dass

$$\int_{-1}^{+1} (1-x)^e (1+x)^{e_0} T_u^n dx$$

gleich ist dem Koeffizienten von $(st)^u$ in der Entwicklung des Integrals

$$\int_0^1 \frac{2^{-e-e_0+1} (1-stz)^{-e}}{z^{-e_0} (1-z)^e (1-stz^2)} dz,$$

das für die hier in Frage kommenden Werte von e und e_0 sich auswerten lässt und für

$$e + e_0 = 0 \quad \frac{1}{\sqrt{st}} \log \frac{1 + \sqrt{st}}{1 - \sqrt{st}}$$

$$e + e_0 = 2 \quad \frac{1}{2} \left[\frac{1}{2(st)^{\frac{2}{3}}} \log \frac{1 - \sqrt{st}}{1 + \sqrt{st}} - \frac{\log(1-st)}{s^2 t^2} \right]$$

$$e + e_0 = 1 \quad \frac{1}{2st} \log(1-st) \quad \text{ergiebt.}$$

Die Reihenentwicklung liefert:

$$2 + \frac{1}{2} st + \frac{2}{3} (st)^2 + \dots + \frac{2}{2u+1} (st)^u$$

$$\frac{1}{2 \cdot 2 \cdot 3} + \frac{2}{2 \cdot 3 \cdot 5} st + \frac{3}{2 \cdot 4 \cdot 7} (st)^2 + \dots + \frac{u+1}{2(u+2) \cdot (2u+3)} (st)^u.$$

$$\frac{1}{2} + \frac{1}{4} st + \frac{1}{8} (st)^2 + \dots + \frac{1}{2^{u+1}} (st)^u.$$

Also ist

$$\int_{-1}^{+1} (1+x)^{e_0} (1-x)^e T_u^n dx$$

gleich

$$\frac{2}{2u+1} \quad \text{für} \quad e + e_0 = 0$$

$$\frac{u+1}{2(u+2)(2u+3)} \quad \text{„} \quad e + e_0 = 2$$

$$\frac{2}{2u+1} \quad \text{„} \quad e + e_0 = 1$$

K_u ist der höchste Koeffizient von T_u ; T_u ist der Koeffizient von s_u in

$$\frac{(1+s+\sqrt{1-2sx+s^2})^{e_0} (1-s+\sqrt{1-sx+s^2})^e}{\sqrt{1-2sx+s^2}}$$

K_u ist also der von s unabhängige Koeffizient von $(sx)^u$, wenn ich diesen Quotienten als Funktion von sx und s auffasse, also der Koeffizient von α^u in

$$\frac{(1+\sqrt{1-\alpha^{e+e_0}})}{\sqrt{1-\alpha}}$$

Das allgemeine Glied einer Reihenentwicklung wird:

$$\begin{aligned} \varrho + \varrho_0 = 0 \\ \varrho + \varrho_0 = 2 \\ \varrho + \varrho_0 = 1 \end{aligned} \quad K_u = \begin{cases} \frac{1.3.5 \dots (2u-1)}{1.3.2 \dots u} \\ \frac{1.3.5 \dots (2u+1)(u+1)}{2.1.2.3 \dots (u+1)(u+2)} \\ \frac{1.3.5 \dots 2u+1}{2.1.2.3 \dots u+1} \end{cases}$$

Nun ergibt sich der Wert von L

$$\begin{aligned} \varrho + \varrho_0 = 0 \\ \varrho + \varrho_0 = 2 \\ \varrho + \varrho_0 = 1 \end{aligned} \quad L = \begin{cases} \left(\frac{1.2.3 \dots \frac{n-1}{2}}{1.3.5 \dots n-2} \right)^2 \\ \left(\frac{1.2.3 \dots \frac{n-1}{2}}{1.3.5 \dots n-2} \right)^2 \frac{n+1}{n-1} \\ 2 \left(\frac{1.2.3 \dots \frac{n}{2}}{1.3.5 \dots n-1} \right)^2, \end{cases}$$

wenn wir $u = \frac{n-1-\varrho-\varrho_0}{2}$ benutzen.

Diese Werte gestalten sich besonders einfach bei Benutzung der Wallisschen Formel für π . In allen Fällen ergibt sich als untere Grenze der Abweichung, der sie mit wachsendem n asymptotisch zustrebt,

$$\frac{n-1}{2^n} \pi.$$

Hieraus können wir z. B. schliessen: Beträgt die Differenz

der Werte einer ganzen rationalen Funktion für -1 und $+1$ weniger als $2 \cdot \frac{n-1}{2^n} \pi$, so ist die Funktion nicht monoton, die Ableitung hat in dem betrachteten Intervall eine Wurzel von ungerader Vielfachheit. (Vergl. Tchebycheff: Sur les fonctions, qui se différent le moins possible de zéro).

Kapitel IV. Von den konvexen Polyedern.

Aufgabe dieses Kapitels ist es, einen Satz zu beweisen, der im folgenden gebraucht werden wird. Es seien

$$x^{(1)} x^{(2)} \dots x^{(n)}$$

unabhängige reelle Variable; ein System von speziellen Werten

$$x_1^{(1)} x_1^{(2)} \dots x_1^{(n)}$$

nenne ich einen Punkt. Es seien eine Anzahl Punkte $p_1, p_2, p_3 \dots$ und hiervon verschiedene Punkte $\bar{p}_1, \bar{p}_2, \bar{p}_3 \dots$ gegeben. Es handelt sich darum, ob es möglich sei, eine lineare Funktion

$$a_1 x^{(1)} + a_2 x^{(2)} + \dots + a_n x^{(n)} + a_{n+1}$$

anzugeben, die an allen Punkten p positive und an allen Punkten \bar{p} negative Werte annimmt. Unser Satz lautet nun:

Entweder ist es möglich eine solche lineare Funktion anzugeben,

Oder man kann aus den gegebenen Punkten $n+2$ Punkte von der Art auswählen, dass schon sie allein die Existenz einer solchen linearen Funktion unmöglich machen.

§ 1. Polyeder.

Aus $n+1$ Punkten, d. h. speziellen Wertsystemen der Variablen $x^{(1)} x^{(2)} \dots x^{(n)}$ können wir eine Determinante der Form bilden:

$$\begin{vmatrix} x_1^{(1)} & x_1^{(2)} & x_1^{(3)} & \dots & x_1^{(n)} & 1 \\ x_2^{(1)} & x_2^{(2)} & x_2^{(3)} & \dots & x_2^{(n)} & 1 \\ \cdot & & & & & \\ \cdot & & & & & \\ \cdot & & & & & \\ x_{n+1}^{(1)} & x_{n+1}^{(2)} & x_{n+1}^{(3)} & \dots & x_{n+1}^{(n)} & 1. \end{vmatrix}$$

Sie sei kurz mit dem Symbol

$$| 1, 2, 3 \dots n, n+1 |$$

bezeichnet. Ein System von n Punkten

$$(1, 2, 3 \dots n)$$

d. h. die Matrix von $n+1$ Spalten und n Zeilen sei eine „Wandung“, ein System von $(n-1)$ Punkten

$$\{ 1, 2, 3 \dots n-1 \}$$

ein „Rand“ genannt. Die Anordnung der Elemente dieser Symbole ist nur insoweit bestimmt, dass ich Symbole, die durch eine gerade Zahl von Transpositionen ineinander übergeführt werden können, als identisch betrachten will.

Sind zwei Punkte a und b gegeben,

$$x_a^{(1)} x_a^{(2)} \dots x_a^{(n)}$$

und

$$x_b^{(1)} x_b^{(2)} \dots x_b^{(n)},$$

so spreche ich von der Linie ($a\bar{b}$) und verstehe darunter die Gesamtheit der Punkte, die $(n-1)$ unabhängigen linearen Gleichungen

$$\begin{aligned} p_1^{(1)} x^{(1)} + p_1^{(2)} x^{(2)} + \dots + p_1^{(n)} x^{(n)} + p_1^{(n+1)} &= 0 \\ p_2^{(1)} x^{(1)} + p_2^{(2)} x^{(2)} + \dots + p_2^{(n)} x^{(n)} + p_2^{(n+1)} &= 0 \\ \vdots & \\ p_{n-1}^{(1)} x^{(1)} + p_{n-1}^{(2)} x^{(2)} + \dots + p_{n-1}^{(n)} x^{(n)} + p_{n-1}^{(n+1)} &= 0 \end{aligned}$$

genügen, wo diese Gleichungen nur der Bedingung unterworfen sind, dass sie von a und b befriedigt werden. Man sieht, dass es auf die Wahl der p dabei nicht ankommt. Denn die $n-1$ Gleichungen definieren die Unbekannten als lineare Funktionen eines Parameters t

$$x^{(1)} = L_1^{(1)} t + L_0^{(1)}$$

$$x^{(2)} = L_1^{(2)} t + L_0^{(2)}$$

⋮

⋮

$$x^{(n)} = L_1^{(n)} t + L_0^{(n)},$$

wo t eine der Variablen x oder eine lineare Funktion derselben bedeuten kann. Nach Definition von t bestimmen sich die L durch

$$\begin{aligned}x_a^{(n)} &= L_1^{(n)}t(a) + L_0^{(n)} \\x_b^{(n)} &= L_1^{(n)}t(b) + L_0^{(n)}\end{aligned}$$

unabhängig von den Werten der p .

Man sieht nun sofort: Habe ich eine Wandung $(1, 2 \dots n)$ und einen Punkt p , der der Bedingung genügt, auf einer gegebenen Linie (ab) zu liegen, so ist die Determinante $|1, 2 \dots n, p|$ eine lineare Funktion des Parameters der Linie (ab) . Hiervon werden wir öfters Gebrauch zu machen haben.

Wenn $n + 1$ Punkte $1, 2, 3 \dots n + 1$ gegeben sind, die

$$|1, 2, 3 \dots n + 1| > 0$$

genügen, so nenne ich das System der Wandungen

$$\begin{aligned}(1, 2 \dots n-1, n), & \quad (-1) (1, 2 \dots n-1, n+1) \\(1, 2 \dots n-2, n+1) & \dots \dots (-1)^n (2, 3 \dots n, n+1)\end{aligned}$$

ein Elementarpolyeder. Das System dieser Wandungen hat die Eigenschaft, dass, wenn ich jede Wandung mit einem Punkte p zur Bildung einer Determinante zusammennehme, die Summe dieser Determinanten von der Wahl des Punktes p unabhängig ist.

$$(1) \quad |wp| + |w'p| + |w''p| + \dots = \text{const.},$$

wo p einen beliebigen Punkt bedeutet. Diese Summe ist nämlich gleich $|1, 2, 3 \dots n + 1|$, sodass die Gleichung besteht:

$$\begin{aligned}(2) \quad & |1, 2, 3 \dots n + 1| + (-1) |1, 2 \dots n, p| + (-1)^2 |1, 2 \dots n-1, n+1, p| \\ & + \dots + (-1)^{n+1} |2, 3 \dots n + 1, p| = 0.\end{aligned}$$

Für alle kleineren Werte von n sei die Gleichung bewiesen. Ich nehme nun auf der linken Seite von (2) den Koeffizienten einer Variablen, etwa den von $x^{(n)}$. In $|1, 2 \dots n + 1|$ kommt $x^{(n)}$ nicht vor; in $(-1) |1, 2 \dots n, p|$ d. h. in

$$\begin{vmatrix} x_1^{(1)} & x_1^{(2)} & \dots & x_1^{(n)} & 1 \\ x_2^{(1)} & x_2^{(2)} & \dots & x_2^{(n)} & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ x_n^{(1)} & x_n^{(2)} & \dots & x_n^{(n)} & 1 \\ x^{(1)} & x^{(2)} & \dots & x^{(n)} & 1 \end{vmatrix}$$

ist der Koeffizient gleich:

$$\begin{vmatrix} x_1^{(1)} & x_1^{(2)} & \dots & x_1^{(n-1)} & 1 \\ x_2^{(1)} & x_2^{(2)} & \dots & x_2^{(n-1)} & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ x_n^{(1)} & x_n^{(2)} & \dots & x_n^{(n-1)} & 1 \end{vmatrix}$$

welche Determinante wir mit $|\bar{1}, \bar{2} \dots \bar{n}|$ bezeichnen wollen, um anzudeuten, dass der Punkt $\bar{1}$ eine Koordinate, nämlich $x^{(n)}$ weniger enthält als 1, also in einem niedriger dimensionalen Raume liegt. Auf diese Weise bestimmt sich der Koeffizient von $x^{(n)}$:

$$|\bar{1}, \bar{2} \dots \bar{n}| + (-1) |\bar{1}, \bar{2} \dots \overline{n-1}, \overline{n+1}| + \dots + (-1)^{n-1} |\bar{2}, \bar{3} \dots \overline{n+1}|$$

dies ist die linke Seite der Gleichung (2) für den niedriger dimensionalen Raum; p hat hier den Wert $\overline{n+1}$. Der Koeffizient von $x^{(n)}$ in (2) verschwindet, und dasselbe können wir von den andern Variablen nachweisen. Zur Betrachtung des absoluten Gliedes setzen wir alle Variablen in (2) gleich Null. Der zweite Term von (2) $(-1) |1, 2 \dots n, p|$ hat dann den Wert

$$(-1) \begin{vmatrix} x_1^{(1)} & x_1^{(2)} & \dots & x_1^{(n)} & 1 \\ x_2^{(1)} & x_2^{(2)} & \dots & x_2^{(n)} & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ x_n^{(1)} & x_n^{(2)} & \dots & x_n^{(n)} & 1 \\ 0 & 0 & \dots & 0 & 1 \end{vmatrix} = (-1) \begin{vmatrix} x_1^{(1)} & x_1^{(2)} & \dots & x_1^{(n)} \\ x_2^{(1)} & x_2^{(2)} & \dots & x_2^{(n)} \\ \vdots & \vdots & & \vdots \\ x_n^{(1)} & x_n^{(2)} & \dots & x_n^{(n)} \end{vmatrix}$$

Das erste Glied von (2) lautet:

$$\begin{vmatrix} x_1^{(1)} & x_1^{(2)} & \dots & x_1^{(n)} & 1 \\ x_2^{(1)} & x_2^{(2)} & \dots & x_2^{(n)} & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ x_{n+1}^{(1)} & x_{n+1}^{(2)} & \dots & x_{n+1}^{(n)} & 1 \end{vmatrix}$$

Entwickeln wir nach den Elementen der letzten Spalte, so liefert das letzte Element eine Determinante, die sich gegen die vorige gerade weghebt, und ebenso heben sich die übrigen Determinanten der Entwicklung von $|1, 2 \dots n+1|$ gegen die folgenden Terme der linken Seite von (2). Da (2) für $n = 1$ gilt, ist sie hierdurch allgemein bewiesen.

Die Bildung der Elementarpolyeder aus $n+1$ Punkten können wir folgendermassen vornehmen: Wir wollen die n Punkte $1, 2 \dots n$ als Wandung auffassen und mit w bezeichnen. $n-1$ Punkte in solcher Anordnung, dass der n te Punkt, hinter sie gesetzt, w ergibt, wollen wir einen Rand nennen, der in w enthalten ist, z. B.

$$\{ 1, 2 \dots n-1 \} = r'.$$

Dagegen $n-1$ Punkte, die, wenn wir den n ten Punkt hinter sie setzen, $-w$ ergeben, wollen wir mit (-1) multiplicieren und alsdann einen Rand nennen, der in w enthalten ist, z. B.

$$- \{ 1, 2 \dots n-2, n \} = r''.$$

Wir können jetzt, wenn wir p statt $(n+1)$ einsetzen, sagen: Ein Elementarpolyeder wird gebildet von den Wandungen

$$w \quad \text{und} \quad -(rp)$$

wo r alle Ränder von w durchlaufen muss, oder von

$$-w \quad \text{und} \quad (rp)$$

je nachdem

$$|wp| > 0 \quad \text{oder} \quad -|wp| > 0.$$

Nehmen wir den ersten Fall an, es sei jetzt q ein Punkt, der $-|wq| > 0$ genügt. Ich bilde das Elementarpolyeder

$-w, (rq)$, nehme das System dieser Wandungen mit dem früheren zusammen und lasse w gegen $-w$ weg. Man sieht, dass dabei die Gleichung (1) erhalten bleibt. Das neue System von Wandungen nennen wir ein Polyeder. Allgemein verstehen wir unter einem Polyeder ein System von Wandungen, das gebildet wird aus den Wandungen einer endlichen Anzahl von Elementarpolyedern unter Weglassung gleicher und entgegengesetzter Wandungen. Ein in den Wandungen vorkommender Punkt 1, 2, 3 . . . heiße ein Eckpunkt.

Für die Polyeder gilt demnach Gleichung (1)

$$|wt| + |w't| + |w''t| + \dots = \text{const.},$$

wo t irgend einen Punkt bedeutet. Die linke Seite heiße der Inhalt des Polyeders; man sieht, dass sich bei dem Aufbau der Polyeder aus Elementarpolyedern der Inhalt additiv zusammensetzt; danach haben alle Polyeder positiven Inhalt.

In jedem Polyeder kommt jeder Rand, der vorkommt, eine gerade Anzahl Male vor, und zwar gleich oft in dem einen und dem entgegengesetzten Sinne. Man sieht zunächst, dass dieser Satz für Elementarpolyeder gilt; gilt er nun für irgend ein Polyeder, so gilt er auch für jedes Polyeder, das aus dem ersten durch Zufügung eines Elementarpolyeders entsteht, sowie für jedes das durch Weglassung zweier gleichen und entgegengesetzten Wandungen entsteht. Somit gilt er allgemein.

§ 2. Konvexe Polyeder.

Unter konvexen Polyedern verstehen wir Polyeder, die der folgenden Bedingung genügen:

Irgend eine Wandung w des Polyeders muss mit jedem in ihr nicht enthaltenen Eckpunkt p zusammengesetzt, eine nicht negative Determinante ergeben.

$$|wp| \geq 0.$$

Unsere Elementarpolyeder mit den Wandungen w und $-(rp)$ genügen der Bedingung, da $|wp| > 0$. Denn sei etwa a der nicht in r vorkommende Punkt von w , so kann die Wandung $-(rp)$ nur noch mit a zusammengesetzt werden, und es ist

$$-|rpa| > 0, \quad \text{da} \quad |rap| = |wp| > 0.$$

Wir führen einen neuen Begriff ein: Wir wollen von einem Punkte i sagen, er liege „innerhalb“ des von den Wandungen $w, w', w'' \dots$ gebildeten konvexen Polyeders wenn

$$|wi| \geq 0 \quad |w'i| \geq 0 \quad |w''i| \geq 0 \dots$$

Wenn nicht immer das Ungleichheitszeichen, sondern ein oder mehrere Male das Gleichheitszeichen gilt, so wollen wir sagen, der Punkt liege auf der Begrenzung des konvexen Polyeders. Ebenso wollen wir sagen, er liege innerhalb der Wandung w , wenn $|wi| = 0$ und $|w'i| \geq 0 \dots$ und auf der Begrenzung der Wandung w , wenn auch ausser $|wi| = 0$ noch Gleichheitszeichen gelten.

Wir wollen zeigen, dass ein Punkt, der innerhalb der Wandung w liegt, dies unabhängig davon thut, zu welchem Elementarpolyeder w gehört. Es sei wp ein Elementarpolyeder, $|wp| > 0$ und p' irgend ein Punkt, der $|wp'| > 0$ genügen soll. Der Punkt i liege in Bezug auf wp innerhalb w , sodass $|wi| = 0$, die übrigen Determinanten positiv sind. Wir wollen zeigen, dass er auch in Bezug auf das Elementarpolyeder wp' innerhalb w liegt. Sei $-r$ ein in w enthaltener Rand, a der nicht zu r gehörende Eckpunkt von w , $w = -(ra)$, so ist (rp) eine von den andern Wandungen des Polyeders wp ; es sei also $|rpi| > 0$ und es genügt nachzuweisen, dass $|rp'i| > 0$.

Wir verfolgen $|rat|$, während der Punkt t auf der Linie (pp') läuft. Den Grenzfall, wo diese lineare Funktion sich auf eine Konstante reduziert, behandeln wir später, und wir bestimmen p'' durch $|rap''| = 0$, wo p'' auf der geraden Linie (pp') . Da $|rap| < 0$ und $|rap'| < 0$, so ist die Reihenfolge der p

$$p p' p'' \quad \text{oder} \quad p' p p'',$$

p'' kann nicht zwischen p und p' liegen; denn eine lineare Funktion kann nicht zwischen zwei negativen Werten verschwinden.

Wir wollen nun nachweisen:

$$|rp''i| = 0.$$

Wir haben:

$$|rap''| = 0 \quad |rai| = 0.$$

Ich behaupte, die beiden linearen Gleichungen

wo

$$|rat| = 0 \quad \text{und} \quad |rti| = 0,$$

$$t = t(x^{(1)}x^{(2)} \dots x^{(n)})$$

irgend einen Punkt bedeutet, sind identisch. Sie mögen aufgelöst etwa lauten

$$L^{(1)}x^{(1)} + L^{(2)}x^{(2)} + \dots + L^{(n)}x^{(n)} + L^{(0)} = 0$$

und

$$\bar{L}^{(1)}x^{(1)} + \bar{L}^{(2)}x^{(2)} + \dots + \bar{L}^{(n)}x^{(n)} + \bar{L}^{(0)} = 0.$$

In keiner dieser Gleichungen sind alle Koeffizienten Null. Denn das würde bedeuten, dass alle n reihigen Determinanten der Matrix ra bzw. ri verschwänden, was aber nicht der Fall ist, weil $|rap| < 0$ und $|rip| < 0$. Die Gleichungen haben $(n+1)$ Wurzeln gemein, nämlich die $(n-1)$ Punkte von r sowie a und i ;

$$(L^{(1)} - \bar{L}^{(1)})x^{(1)} + (L^{(2)} - \bar{L}^{(2)})x^{(2)} + \dots + (L^{(0)} - \bar{L}^{(0)})$$

verschwindet also an $n+1$ Stellen. Die Determinante dieses Systems von $n+1$ homogenen linearen Gleichungen mit $(n+1)$ Unbekannten $|rai|$ verschwindet zwar, es können aber nicht alle Unterdeterminanten n ten Grades verschwinden, weil sonst nicht

$$|rap| < 0 \quad \text{sondern} \quad |rap| = 0$$

sein müsste.

Hieraus folgt, dass der Quotient zweier der Unbekannten nicht unendlich sein kann, und setze ich, was erlaubt ist, eine der Differenzen gleich 0, so verschwinden auch alle andern. Damit ist die Identität der Gleichungen

$$|rat| = 0 \quad \text{und} \quad |rti| = 0$$

nachgewiesen, und da $|rap''| = 0$, so folgt

$$|rp''i| = 0.$$

Wir werden von dem Schluss, dass aus

$$|rap''| = 0 \quad \text{und} \quad |rai| = 0$$

entweder $|rp''i| = 0$ oder das Verschwinden der n reihigen Determinanten von ra folgt, noch öfters Gebrauch machen.

Da nun p'' nicht zwischen p und p' liegen kann und

$$|rpi| > 0,$$

so folgt

$$|rp'i| > 0.$$

Ist, was wir eben ausschlossen, $|rat|$ konstant, wenn sich t auf (pp') bewegt, so behaupte ich, auch $|rit|$ ist auf dieser Linie konstant. Ist dies nämlich nicht der Fall, so giebt es auf (pp') einen Punkt p'' , für den $|rip''| = 0$, und ich beweise analog wie oben, dass nun auch $|rap''| = 0$, was der Annahme, dass $|rat|$ auf (pp') konstant sei, widerspricht.

Ist

$$|wp'| < 0, \quad \text{während} \quad |wp| > 0,$$

sodass sich $p \dots p'' \dots p'$ folgen, so ist zunächst $|rp'i| < 0$. Wir kehren aber für das Polyeder wp' das Zeichen von w und demnach auch von r um, sodass der Satz bestehen bleibt.

Ein Punkt i , der innerhalb der Wandung w liegt, thut dies unabhängig davon, in welchem Elementarpolyeder w vorkommt.

Es stelle (ab) eine gerade Linie dar, und b liege im Innern eines konvexen Polyeders. Ich behaupte: Wenn ich mit einem variablen Punkte t auf der geraden Linie (ab) von a kommend über b hinaus fortschreite, so komme ich an einen Punkt c , der der Begrenzung des konvexen Polyeders angehört. Bilde ich nämlich die Determinanten $|wt|$, $|w't|$, $|w''t| \dots$ und lasse t längs (ab) laufen, so können diese Determinanten, da ihre Summe konstant bleiben muss, nicht alle wachsen; falls sie nicht alle konstant bleiben — und dass dieses nicht der Fall sein kann, werden wir gleich nachweisen — muss mindestens eine von ihnen abnehmen. Da diese Determinante eine lineare Funktion von t ist, so muss sie, während ich mit t auf (ab) fortschreite, einmal 0 und sodann negativ werden; sie bleibt dann negativ, soweit ich auch über c hinaus fortschreite. Da dasselbe auch gilt, wenn ich von b nach a zu gehe, so erhalten wir sofort die Sätze:

1) Eine gerade Linie kann nur innerhalb eines endlichen Stückes innerhalb eines konvexen Polyeders verlaufen.

2) Enthält eine gerade Linie einen Punkt, der dem Innern, aber nicht der Begrenzung eines konvexen Polyeders angehört, so hat sie mit der Begrenzung dieses Polyeders zwei und nur zwei Punkte gemein.

3) Eine Gerade, die mit dem Innern eines konvexen Polyeders zwei Punkte gemein hat, hat

mit ihm alle Punkte gemein, die auf ihr zwischen diesen beiden liegen.

Wir haben nun nachzuweisen, dass die Determinanten

$$|wt|, \quad |w't|, \quad |w''t| \dots$$

nicht alle konstant bleiben können, wenn sich t auf einer geraden Linie bewegt. Wir weisen dies zunächst für ein Elementarpolyeder mit den Eckpunkten $1, 2, 3 \dots n, n+1$ nach. Was bedeutet es, wenn $|1, 2, \dots, n, t|$ von t unabhängig ist, wenn t sich auf einer geraden Linie bewegt? Die Gerade sei gegeben durch:

$$x^{(1)} = L_1^{(1)}t + L_0^{(1)}$$

$$x^{(2)} = L_1^{(2)}t + L_0^{(2)}$$

$$\vdots$$

$$x^{(n)} = L_1^{(n)}t + L_0^{(n)}$$

Es muss dann

$$\frac{d}{dt} \begin{vmatrix} x_1^{(1)} & x_1^{(2)} & \dots & x_1^{(n)} & 1 \\ x_2^{(1)} & x_2^{(2)} & \dots & x_2^{(n)} & 1 \\ \vdots & \vdots & \dots & \vdots & \vdots \\ x_n^{(1)} & x_n^{(2)} & \dots & x_n^{(n)} & 1 \\ L_1^{(1)}t + L_0^{(1)} & L_1^{(2)}t + L_0^{(2)} & \dots & L_1^{(n)}t + L_0^{(n)} & 1 \end{vmatrix} = 0$$

$$\begin{vmatrix} x_1^{(1)} & x_1^{(2)} & \dots & x_1^{(n)} & 1 \\ x_2^{(1)} & x_2^{(2)} & \dots & x_2^{(n)} & 1 \\ \vdots & \vdots & \dots & \vdots & \vdots \\ x_n^{(1)} & x_n^{(2)} & \dots & x_n^{(n)} & 1 \\ L_1^{(1)} & L_1^{(2)} & \dots & L_1^{(n)} & 1 \end{vmatrix} = 0$$

Dies ist eine homogene lineare Gleichung der L_i , und wir haben gemäss den $n+1$ Wandungen $n+1$ solcher Gleichungen für die n Grössen L_i . Nun ist in der hingeschriebenen Gleichung der Koeffizient von $L_i^{(1)}$ gleich der dem Element $x_{n+1}^{(1)}$ zugeordneten Un-

terdeterminante in $|1, 2, \dots, n+1|$, die wir mit $X_{n+1}^{(i)}$ bezeichnen wollen. Die Matrix der Gleichungen für die L_i ist demnach

$$\begin{array}{cccc} X_1^{(1)} & X_1^{(2)} & \dots & X_1^{(n)} \\ X_2^{(1)} & X_2^{(2)} & \dots & X_2^{(n)} \\ \vdots & \vdots & & \vdots \\ X_{n+1}^{(1)} & X_{n+1}^{(2)} & \dots & X_{n+1}^{(n)} \end{array}$$

Nun müssen entweder die L_i oder alle n reihigen Determinanten dieser Matrix verschwinden. Wäre Letzteres der Fall, so müsste auch die aus den Unterdeterminanten von $|1, 2, 3 \dots, n, n+1|$ gebildete Determinante

$$\begin{vmatrix} X_1^{(1)} & X_1^{(2)} & \dots & X_1^{(n)} & X_1^{(n+1)} \\ X_2^{(1)} & X_2^{(2)} & \dots & X_2^{(n)} & X_2^{(n+1)} \\ \vdots & \vdots & & \vdots & \vdots \\ X_{n+1}^{(1)} & X_{n+1}^{(2)} & \dots & X_{n+1}^{(n)} & X_{n+1}^{(n+1)} \end{vmatrix}$$

verschwinden; diese ist aber nach einem bekannten Determinantensatz gleich $(|1, 2, 3 \dots, n, n+1|)^n$, also sicher von 0 verschieden.

Es können also nicht alle Determinanten $|wt|$, $|w't| \dots$ auf der geraden Linie konstant bleiben; wegen der Bedingung (1) kann nicht eine allein variabel sein, es sind also mindestens 2 Determinanten variabel.

Wir wollen diesen für Elementarpolyeder geführten Beweis auf allgemeine Polyeder ausdehnen. Wir haben dabei nur nachzuweisen, dass nicht alle Wandungen der Elementarpolyeder, die variable Determinanten liefern, bei der Zusammensetzung wegfallen sein können. Es sei eine Linie gegeben, und wir nehmen eine Wandung, die eine auf ihr nicht konstante Determinante liefert. Da es, wie wir sahen, auf die Konstanten L_0 der Linie nicht ankommt, sondern nur auf die L_i , so können wir uns die L_0 so bestimmt denken, dass die Linie die Wandung in ihrem Innern schneidet und demnach an der Schnittstelle in das Innere eines Elementarpolyeders tritt; wir gehen nun auf ihr stets in derselben Richtung weiter, sie muss an irgend einer Stelle aus dem Elementarpolyeder wieder austreten, sagen wir bei α_1 durch

die Wandung w_1 ; wenn die Wandung w_1 wegfallen soll, so muss ein Elementarpolyeder mit der Wandung $-w_1$ existieren, und in dieses tritt, da α_1 auch im Innern von $-w_1$ liegt, die Linie nun ein; sie muss auch aus diesem wieder austreten, etwa bei α_2 durch w_2 ; dann muss auch wieder $-w_1$ existieren, u. s. w. Da wir auf der Linie stets in demselben Sinn fortschreiten, so kommen wir stets zu wirklich voneinander verschiedenen Elementarpolyedern, und wir sehen, dass es bei einer endlichen Anzahl von Elementarpolyedern keine gerade Linie geben kann, auf der die Determinanten aller Wandungen konstant sind.

§ 3. Anordnung zu konvexen Polyedern.

Wir wollen zeigen, dass, wenn eine Anzahl Punkte gegeben ist, die mindestens eine nicht-verschwindende $(n+1)$ reihige Determinante liefern, man stets ein konvexes Polyeder konstruieren kann, dessen Eckpunkte sämtlich zu diesen gegebenen Punkten gehören, und in dessen Innern alle gegebenen Punkte liegen.

Wir nehmen an, wir hätten ein konvexes Polyeder konstruiert, dessen Eckpunkte sämtlich zu den gegebenen Punkten gehören, und in dessen Innern sich möglicherweise noch gegebene Punkte befinden. Wir wollen zeigen, wie dies Polyeder zu erweitern ist, falls noch Punkte sich ausserhalb befinden. Das Polyeder habe die Wandungen $w, w_1, w_2 \dots$, und p sei ein ausserhalb befindlicher Punkt. Es können dann die Determinanten $|wp|$, $|w_1p|$, $|w_2p| \dots$ weder alle positiv (einschliesslich der Null), noch alle negativ sein. Denn im ersten Fall läge der Punkt p innerhalb, im zweiten hätte das Polyeder einen negativen Inhalt (S. 65). Ich teile die Wandungen des Polyeders ein in

$$w'_1, w'_2, w'_3 \dots \quad \text{und} \quad w''_1, w''_2, w''_3 \dots$$

sodass

$$\begin{array}{lll} |w'_1p| \geq 0 & |w'_2p| \geq 0 & |w'_3p| \geq 0 \dots \\ |w''_1p| < 0 & |w''_2p| < 0 & |w''_3p| < 0. \end{array}$$

Die Wandungen w' können in das neu zu bildende konvexe Polyeder mit herüber genommenen werden, die Wandungen w'' nicht. Wir bilden demgemäss die Elementarpolyeder $-w''p$, und ich behaupte, dass das neue Polyeder allen Bedingungen genügt.

Wir teilen zum Beweis die in dem ursprünglicher Polyeder vorkommenden Ränder in drei Arten.

Die Ränder

$$r'_1, r'_2, r'_3 \dots$$

sollen nur in den Wandungen

$$w'_1, w'_2, w'_3 \dots$$

Die Ränder

$$r''_1, r''_2, r''_3 \dots$$

sollen nur in den Wandungen

$$w''_1, w''_2, w''_3 \dots$$

Die Ränder

$$r_1, r_2, r_3 \dots$$

sollen sowohl in den Wandungen w' als auch in w'' enthalten sein.

Hierbei haben wir der Kürze halber „Rand“ im absoluten Sinn, ohne Rücksicht auf das Vorzeichen gebraucht. Wir denken uns nun alle Ränder, die im Polyeder vorkommen, hingeschrieben, und zwar jeden so oft als er vorkommt, unter Beachtung des ihm zukommenden Zeichens. Teilen wir die Ränder in die drei Arten ein, so ist klar, dass auch innerhalb jeder dieser Arten jeder Rand gleich oft mit dem einen und dem entgegengesetzten Zeichen vorkommen muss (S. 65). Für die Richtigkeit unserer Schlussfolgerungen ist es dabei belanglos, ob von jeder Art Ränder existieren oder nicht.

Bilden wir nun sämtliche Elementarpolyeder

$$-w''p,$$

wobei ja $-|w''p| > 0$, so haben diese die Wandungen

$$-w'' \quad \text{und} \quad (\varrho p),$$

wenn ϱ irgend ein Rand von w'' ist. Die ϱ zerfallen in r und r'' , und zwar sind in den ϱ sämtliche r'' , nicht aber sämtliche r enthalten; denn die r kommen, sofern sie Ränder der w' sind, in den ϱ nicht vor. Danach heben sich wohl die Wandungen $(r''p)$, nicht aber alle Wandungen (rp) gegeneinander auf. w'' hebt sich gegen $-w''$ und unser neues Polyeder hat nur Wandungen

$$w' \quad \text{und} \quad (rp).$$

Es ist nun nur noch zu zeigen, dass die Wandungen (rp) mit

jedem der Eckpunkte des alten Polyeders zusammengesetzt, eine positive Determinante ergeben. Zu diesem Zweck verbinde ich p mit irgend einem Punkte a , der auf dem Rande r liegt, d. h. $|w'a| = 0$ und $|w''a| = 0$ genügt, und zu der Begrenzung des ursprünglichen Polyeders gehört, durch eine Gerade (pa) . Schliessen wir dabei den Grenzfall, dass in $|w'p| \geq 0$ das Gleichheitszeichen gelten soll, zunächst aus.

Es ist nun die Determinante $|w''t|$, wo t einen Punkt der Geraden (pa) bedeutet, für $t = a$ gleich Null, und für $t = p$ negativ; sie bleibt also negativ, soweit ich mich auch von a in der Richtung nach p oder über p hinaus entferne. Dagegen ist $|w't|$ für $t = p$ positiv, für $t = a$ Null, und demnach negativ für alle Punkte von (pa) , die von p aus gesehen, jenseits a liegen. Demnach ist für alle Punkte der Geraden ausser a eine der Determinanten des ursprünglichen Polyeders negativ; wir können also sagen, die Gerade (pa) hat mit dem ursprünglichen Polyeder keinen Punkt ausser a gemeinsam. Ist a irgend eine gemeinsame Lösung von $|w't| = 0$ und $|w''t| = 0$, ohne zu der Begrenzung des ursprünglichen Polyeders zu gehören, so hat die Gerade (pa) mit dem Innern des Polyeders überhaupt keinen Punkt gemeinsam.

Ich behaupte nun $|rpt|$ verschwindet für keinen Innenpunkt des Polyeders, abgesehen von den auf dem Rand r liegenden Punkten. Nehmen wir also an, es bestände die Gleichung $|rpi| = 0$, wo i einen Innenpunkt bedeutet. Wir legen eine Gerade durch die Punkte p und i durch $(n-1)$ lineare Gleichungen. Als eine dieser Gleichungen können wir $|rpt| = 0$ wählen, die übrigen Gleichungen seien

$$f_1(t) = 0 \quad f_2(t) = 0 \quad \dots \quad f_{n-1}(t) = 0.$$

Wir schreiben diese Gleichungen zusammen mit $|w't| = 0$. Es sei $w' = (r1)$, so haben wir das System:

$$\begin{aligned} |r1t| &= 0 \\ |rpt| &= 0 \\ f_1(t) &= 0 \\ &\vdots \\ f_{n-1}(t) &= 0. \end{aligned}$$

Die Lösung dieses Systems von n Gleichungen mit n Unbekannten stellt den Schnittpunkt der Geraden (pi) mit $|w't| = 0$ dar.

Es sei nun D die Determinante dieser Gleichungen; ist $D \neq 0$, so haben die Gleichungen eine Lösung, etwa a , und es ist

$$|rpa| = 0 \quad \text{und} \quad |r1a| = 0.$$

Nun folgt nach unserm Beweis S. 67 aus diesen Gleichungen, dass entweder $|rp1| = 0$ oder alle n reihigen Determinanten von ra verschwinden. Da wir die erste Eventualität $|rp1| = |w'p| = 0$ ausgeschlossen haben, so folgt, das Verschwinden der n reihigen Determinanten von ra , und hieraus folgt $|w'a| = 0$. D. h. der Punkt a genügt $|w'a| = 0$ und $|w''a| = 0$, er liegt auf dem Rande r .

(pi) ist dann eine gerade Linie, die von p zu einem Randpunkt a geht, und einen Punkt i des Innern enthält, was wir aber als unmöglich nachgewiesen haben.

Wir haben $D \neq 0$ angenommen. Es sei jetzt $D = 0$. Wir beachten nun: Vom Punkt i war nur $|rpi| = 0$ vorausgesetzt. Wir ziehen nun von i eine Gerade zu einem beliebigen Punkte von r , etwa b , und es sei i' ein beliebiger Punkt der Geraden (ib) zwischen i und b ; i' liegt dann im Innern des Polyeders, und es ist auch $|rpi'| = 0$, weil beide Eigenschaften von den beiden Punkten, i und b gelten. Wir können also in unserer Betrachtung i' statt i benutzen. Wir lassen jetzt i' auf der Geraden (ib) wandern, dadurch werden alle n Koordinaten von i' lineare Funktionen eines Parameters t . Es kann aber D nicht identisch in t verschwinden, weil D für $i' = b$ nicht verschwindet; denn die Gerade (pb) hat mit $|w't| = 0$ sicher einen und nur einen Punkt nämlich b gemeinsam; d. h. die Gleichungen (3) haben eine und nur eine Lösung, und dann kann ihre Determinante nicht verschwinden. Verschwindet aber D nicht identisch in t , so muss es auch ausser $i' = b$ noch Punkte i' geben, für die $D \neq 0$.

Die Annahme $|rpi| = 0$ ist also unzulässig: $|rpt|$ kann für keinen Punkt, der im Innern des ursprünglichen Polyeders liegt, verschwinden. Wir haben aber noch die einschränkende Voraussetzung, dass in $|w'p| \geq 0$ nicht das Gleichheitszeichen gelte; nehmen wir also jetzt $|w'p| = 0$ an. $|r1p| = 0$. Dann verschwindet $|rpt|$ für dieselben Punkte, für die $|r1t|$ verschwindet, und umgekehrt. Denn nach S. 67 bedingt von den beiden Gleichungen

$$|r1t| = 0 \quad \text{und} \quad |rpt| = 0$$

wegen $|r1p| = 0$ die eine die andere. Die n reihigen Deter-

minanten von rp können nicht alle verschwinden, denn r kommt noch in einer Wandung w'' vor, und es müsste $|w''p| = 0$, während $|w''p| < 0$; auch die Determinanten von $r1$ d. h. von w' können nicht verschwinden, wie sich aus der Betrachtung des Elementarpolyeders, in dem w' vorkommt — und in mindestens einem muss es vorkommen —, ergibt.

$|rpt|$ verschwindet also für keinen Punkt des Innern, wo die zur Begrenzung gehörenden Punkte, die $|w't| = 0$ genügen, eventuell auszuschliessen sind. Hieraus folgt aber sofort, dass $|rpt|$ für alle Punkte im Innern dasselbe Vorzeichen haben muss. Wir wissen nun, dass das Vorzeichen für einen Punkt positiv ist, nämlich für den letzten Eckpunkt des Elementarpolyeders, dem die Wandung (rp) angehört.

Hiermit ist alles, was zu beweisen war, bewiesen; wir haben aus einem gegebenen konvexen Polyeder ein neues konstruiert, das einen Punkt p , der ausserhalb des ursprünglichen lag, als Eckpunkt enthält und die Eckpunkte des früheren entweder auch als Eckpunkte, oder im Innern. Damit ist der Eingang dieses Paragraphen angekündigte Satz bewiesen.

§ 4. Trennung zweier konvexer Polyeder.

Es seien P und \bar{P} zwei konvexe Polyeder, die der Bedingung genügen, dass kein Punkt zugleich im Innern von P und im Innern von \bar{P} liegt. Ich behaupte dann:

Es giebt eine lineare Funktion

$$p^{(1)}x^{(1)} + p^{(2)}x^{(2)} + \dots + p^{(n)}x^{(n)} + p^{(n+1)},$$

die an allen im Innern von P liegenden Punkten positive, und an allen im Innern von \bar{P} liegenden Punkten negative Werte annimmt.

Es sei eine Funktion zweier Punkte

$$p = p(x^{(1)}, x^{(2)}, \dots, x^{(n)})$$

und

$$\bar{p} = \bar{p}(\bar{x}^{(1)}, \bar{x}^{(2)}, \dots, \bar{x}^{(n)})$$

definiert durch

$$f = + \sqrt{(x^{(1)} - \bar{x}^{(1)})^2 + (x^{(2)} - \bar{x}^{(2)})^2 + \dots + (x^{(n)} - \bar{x}^{(n)})^2}.$$

Wir suchen das Minimum von f unter der Bedingung, dass p im Polyeder P und \bar{p} im Polyeder \bar{P} liege. Dieses Minimum muss existieren und von 0 verschieden sein, weil nicht alle Quadrate verschwinden können. Es werde etwa an den Punkten p^* und \bar{p}^* angenommen. Wir legen durch diese beiden Punkte eine gerade Linie, d. h. wir stellen $(n-1)$ Gleichungen auf, denen die Koordinaten von p^* und \bar{p}^* genügen. Diese Gleichungen seien die folgenden:

$$a_1^{(1)} x^{(1)} + a_1^{(2)} x^{(2)} + \dots + a_1^{(n)} x^{(n)} + a_1^{(n+1)} = 0$$

$$a_2^{(1)} x^{(1)} + a_2^{(2)} x^{(2)} + \dots + a_2^{(n)} x^{(n)} + a_2^{(n+1)} = 0$$

⋮

$$a_{n-1}^{(1)} x^{(1)} + a_{n-1}^{(2)} x^{(2)} + \dots + a_{n-1}^{(n)} x^{(n)} + a_{n-1}^{(n+1)} = 0.$$

Zu diesen Gleichungen nehmen wir noch eine, die für p^* und \bar{p}^* nicht erfüllt ist:

$$a_n^{(1)} x^{(1)} + a_n^{(2)} x^{(2)} + \dots + a_n^{(n)} x^{(n)} + a_n^{(n+1)} = 0,$$

und unterwerfen die Koeffizienten a den Bedingungen:

$$\sum_{\nu=1}^{\nu=n} (a_{\mu}^{(\nu)})^2 = 1$$

für alle μ und

$$\sum_{\nu=1}^{\nu=n} (a_{\mu}^{(\nu)} a_{\mu'}^{(\nu)}) = 0$$

für $\mu \neq \mu'$.

Man überzeugt sich leicht, dass dies stets auf unendliche vielfache Weise möglich ist, und es ist bekannt, dass alsdann auch

$$\sum_{\mu=1}^{\mu=n} (a_{\nu}^{(\mu)})^2 = 1$$

für alle ν und

$$\sum_{\mu=1}^{\mu=n} (a_{\nu}^{(\mu)} a_{\nu'}^{(\mu)}) = 0$$

für $\nu \neq \nu'$.

Wir gehen jetzt zu folgendem Koordinatensystem über:

$$\begin{aligned}
a_1^{(1)}x^{(1)} + a_1^{(2)}x^{(2)} + \dots + a_1^{(n)}x^{(n)} + a_1^{(n+1)} &= \xi_1 \\
a_2^{(1)}x^{(1)} + a_2^{(2)}x^{(2)} + \dots + a_2^{(n)}x^{(n)} + a_2^{(n+1)} &= \xi_2 \\
&\vdots \\
a_n^{(1)}x^{(1)} + a_n^{(2)}x^{(2)} + \dots + a_n^{(n)}x^{(n)} + a_n^{(n+1)} &= \xi_n.
\end{aligned}$$

Man sieht sofort, dass bei dieser Transformation die Form der Funktion f erhalten bleibt, sodass

$$\begin{aligned}
f &= \sqrt{(x^{(1)} - \bar{x}^{(1)})^2 + (x^{(2)} - \bar{x}^{(2)})^2 + \dots + (x^{(n)} - \bar{x}^{(n)})^2} \\
&= \sqrt{(\xi^{(1)} - \bar{\xi}^{(1)})^2 + (\xi^{(2)} - \bar{\xi}^{(2)})^2 + \dots + (\xi^{(n)} - \bar{\xi}^{(n)})^2}.
\end{aligned}$$

Es hat also auch

$$\sqrt{(\xi^{(1)} - \bar{\xi}^{(1)})^2 + (\xi^{(2)} - \bar{\xi}^{(2)})^2 + \dots + (\xi^{(n)} - \bar{\xi}^{(n)})^2},$$

wenn p in P und \bar{p} in \bar{P} liegen soll, bei $p = p^*$ und $\bar{p} = \bar{p}^*$ ein Minimum. Die neuen Koordinaten von p^* und \bar{p}^* sind alle gleich Null mit Ausnahme von $\xi^{(n)*}$ und $\bar{\xi}^{(n)*}$. Es sei etwa $\xi^{(n)*} > \bar{\xi}^{(n)*}$. Ich behaupte: Kein Punkt im Innern von P kann ein $\xi^{(n)}$ haben, das kleiner ist als $\xi^{(n)*}$. Nehmen wir an, dies sei bei einem Punkte p^{**} der Fall. Wir ziehen eine gerade Linie von p^* nach p^{**} , und zwar denken wir uns die Koordinaten der Punkte dieser Linie dargestellt durch einen Parameter t , der zunimmt, wenn ich von p^* nach p^{**} gehe. Dann ist $\frac{d\xi^{(n)}}{dt}$ für $\xi^{(n)} = \xi^{(n)*}$ negativ. Ich bilde $\frac{df}{dt}$ bei festgehaltenem $\bar{p} = \bar{p}^*$ für $p = p^*$

$$\frac{df}{dt} = \frac{\xi^{(n)*} - \bar{\xi}^{(n)*}}{\sqrt{(\xi^{(n)*} - \bar{\xi}^{(n)*})^2}} \cdot \frac{d\xi^{(n)}}{dt}.$$

Diese Grösse ist negativ, wenn $\frac{d\xi^{(n)}}{dt} < 0$. Da nun jeder Punkt der Geraden von p^* nach p^{**} im Innern von P liegt, so widerspricht dies der Minimalbedingung der Punkte p^* und \bar{p}^* . Ebenso weisen wir nach, dass kein Punkt von \bar{P} ein grösseres $\bar{\xi}^{(n)}$ haben kann als $\bar{\xi}^{(n)*}$. Nehmen wir daher eine Konstante c an, sodass

$$\xi^{(n)*} > c > \bar{\xi}^{(n)*},$$

so haben wir die verlangte lineare Funktion, die für alle Punkte von P positive und für alle Punkte von \bar{P} negative Werte annimmt, dargestellt in:

$$\begin{aligned} & \xi^{(n)} - c \\ &= a_n^{(1)}x^{(1)} + a_n^{(2)}x^{(2)} + \dots + a_n^{(n)}x^{(n)} + a_n^{(n+1)} - c. \end{aligned}$$

§ 5. Zerlegung in Elementarpolyeder.

Wir wollen den Satz beweisen: Wenn ein Punkt i in einem konvexen Polyeder \mathfrak{P} liegt, so kann man stets ein Elementarpolyeder P , dessen Eckpunkte zu den Eckpunkten von \mathfrak{P} gehören, angeben, in dem i liegt.

Der Satz gelte bei niedriger dimensionalen Räumen für bewiesen. Für $n = 1$ ist er trivial, da in diesem Fall jedes konvexe Polyeder ein Elementarpolyeder ist.

Von einem Eckpunkt a von \mathfrak{P} ziehe ich eine gerade Linie nach i , die bei a aus dem Innern von \mathfrak{P} austreten möge. Wenn ich ein Elementarpolyeder angeben kann, das a und i als Innenpunkte (wobei a speziell Eckpunkt sein möge, und nur Eckpunkte von \mathfrak{P} als Eckpunkte enthält, so ist der Satz bewiesen.

Der Punkt a muss mindestens einer Determinantengleichung $|w't| = 0$ genügen. Wir nehmen eine lineare Transformation vor, die die Koordinaten $x^{(1)}, x^{(2)}, x^{(3)} \dots x^{(n)}$ in $\xi^{(1)}, \xi^{(2)} \dots \xi^{(n)}$ überführt, und zwar sei

$$|w't| = \pm \xi^{(n)}.$$

Die Substitutionsdeterminante sei positiv; es bleiben dann alle Eigenschaften der Polyeder, da sich alle Determinanten nur mit der Substitutionsdeterminante multiplicieren, in dem neuen Raum erhalten.

Es können ausser w' noch andere Wandungen von \mathfrak{P} Determinantengleichungen liefern, die mit $\xi^{(n)} = 0$ identisch sind. Es seien dies

$$w', w'_1, w'_2 \dots,$$

so dass in allen Eckpunkten dieser Wandungen die $\xi^{(n)}$ -Koordinate verschwindet.

Dagegen sollen

$$w'', w''_1, w''_2 \dots,$$

hiervon verschiedene Determinantengleichungen liefern, sodass keine dieser Wandungen nur Eckpunkte mit verschwindender $\xi^{(n)}$ -Koordinate enthält.

Es sei das Vorzeichen von $|w't| = \pm \xi^{(n)}$ so bestimmt, dass ein Eckpunkt von \mathfrak{B} , etwa b ein negatives $\xi^{(n)}$ habe. Da \mathfrak{B} konvex, muss $|w'b| \geq 0$ d. h. wenn etwa $w' = (1, 2 \dots n)$

$$\begin{vmatrix} \xi_1^{(1)} & \xi_1^{(2)} & \dots & \xi_1^{(n-1)} & 0 & 1 \\ \xi_2^{(1)} & \xi_2^{(2)} & \dots & \xi_2^{(n-1)} & 0 & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \xi_n^{(1)} & \xi_n^{(2)} & \dots & \xi_n^{(n-1)} & 0 & 1 \\ \xi_b^{(1)} & \xi_b^{(2)} & \dots & \xi_b^{(n-1)} & \xi_b^{(n)} & 1 \end{vmatrix} \geq 0.$$

Vertauschen wir die letzten Spalten und berücksichtigen, dass $\xi_b^{(n)} < 0$, so folgt

$$\begin{vmatrix} \xi_1^{(1)} & \xi_1^{(2)} & \dots & \xi_1^{(n-1)} & 1 \\ \xi_2^{(1)} & \xi_2^{(2)} & \dots & \xi_2^{(n-1)} & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \xi_n^{(1)} & \xi_n^{(2)} & \dots & \xi_n^{(n-1)} & 1 \end{vmatrix} \geq 0.$$

Das Gleichheitszeichen kann nicht gelten, weil sonst w' mit jedem Punkt eine verschwindende Determinante liefern müsste. Lassen wir die $\xi^{(n)}$ -Koordinate weg, was wir durch $\overline{w}, \overline{r} \dots$ statt $w, r \dots$ ausdrücken wollen, so kann in dem $(n-1)$ -dimensionalen Raum \overline{w}' als Elementarpolyeder betrachtet werden. Dasselbe gilt von $\overline{w}'_1 \overline{w}'_2 \dots$, da ja alle mit b positive Determinanten liefern müssen. Ebenso sehen wir, dass jeder Punkt von \mathfrak{B} ein negatives oder verschwindendes $\xi^{(n)}$ haben muss; denn ein Punkt mit positivem $\xi^{(n)}$ würde mit w' eine negative Determinante liefern.

Wir teilen die Ränder der Wandungen $w', w'_1, w'_2 \dots$ in zwei Teile:

$$r, r_1, r_2 \dots \quad \text{mögen ausser in} \quad w', w'_1, w'_2 \dots$$

noch in Wandungen w'' vorkommen,

$$r', r'_1, r'_2 \dots \quad \text{mögen nur in} \quad w', w'_1, w'_2 \dots$$

vorkommen. (Vergl. S. 72).

Die Ränder

$r, r_1, r_2 \dots$ können in den Wandungen $w', w'_1, w'_2 \dots$

nicht mit demselben Vorzeichen vorkommen, wie in $w'', w''_1 \dots$.
Denn wäre etwa

$$(rb) = w'' \quad (rc) = w',$$

so müsste

$$\begin{array}{|l} |w'b| \geq 0 \\ |rcb| \geq 0 \end{array} \quad \begin{array}{|l} |w''c| \geq 0 \\ |rbc| \geq 0 \end{array}$$

woraus $|rbc| = 0$ folgen würde, was aber nicht möglich ist, weil sonst die Gleichungen

$$|rbt| = 0 \quad \text{und} \quad |rct| = 0$$

nach S. 67 identisch wären. Denn alle n reihigen Determinanten von rb oder rc d. h. von w'' und w' können nicht verschwinden. Die Ränder $r, r_1, r_2 \dots$ mögen als $r, r_1, r_2 \dots$ in $w', w'_1, w'_2 \dots$ und als $-r, -r_1, -r_2 \dots$ in $w'', w''_1, w''_2 \dots$ enthalten sein.

Fassen wir $\bar{w}, \bar{w}'_1, \bar{w}'_2 \dots$ als Elementarpolyeder, die

$$\bar{r}, \bar{r}_1, \bar{r}_2 \dots \bar{r}', \bar{r}'_1, \bar{r}'_2, \dots$$

als ihre Wandungen auf, so heben sich die $\bar{r}, \bar{r}'_1, \bar{r}'_2 \dots$ gegenseitig heraus. (Vergl. S. 72). Ich behaupte, die Wandungen $\bar{r}, \bar{r}_1, \bar{r}_2 \dots$ bilden ein konvexes Polyeder. Sei c ein nicht in \bar{r} vorkommender Eckpunkt, so ist $|\bar{r}c| \geq 0$ nachzuweisen. Es sei etwa $\bar{r} = (1, 2 \dots n-1)$, so soll

$$\begin{vmatrix} \xi_1^{(1)} & \xi_1^{(2)} & \dots & \xi_1^{(n-1)} & 1 \\ \xi_2^{(1)} & \xi_2^{(2)} & \dots & \xi_2^{(n-1)} & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ \xi_{n-1}^{(1)} & \xi_{n-1}^{(2)} & \dots & \xi_{n-1}^{(n-1)} & 1 \\ \xi_c^{(1)} & \xi_c^{(2)} & \dots & \xi_c^{(n-1)} & 1 \end{vmatrix} \geq 0.$$

$-r$ kommt in einer Wandung $w'', w''_1 \dots$ vor, und zwar etwa als $-(rb)$; b muss dann ein negatives $\xi^{(n)}$ haben. Da \mathfrak{P} konvex ist, muss $-|rbc| \geq 0$; d. h. es muss

$$\begin{vmatrix} \xi_1^{(1)} & \xi_1^{(2)} & \dots & \xi_1^{(n-1)} & 0 & 1 \\ \xi_2^{(1)} & \xi_2^{(2)} & \dots & \xi_2^{(n-1)} & 0 & 1 \\ \vdots & \vdots & & \vdots & & \vdots \\ \xi_b^{(1)} & \xi_b^{(2)} & \dots & \xi_b^{(n-1)} & \xi_b^{(n)} & 1 \\ \xi_c^{(1)} & \xi_c^{(2)} & \dots & \xi_c^{(n-1)} & 0 & 1 \end{vmatrix} \leq 0,$$

woraus nach Vertauschung der letzten Zeilen und Spalten unter Beachtung von $\xi_b^{(n)} < 0$ die obige Ungleichung folgt. Die \bar{r} bilden also ein konvexes Polyeder.

Der Punkt α muss innerhalb dieses konvexen Polyeders liegen. Gäbe es nämlich eine Wandung \bar{r} , für die $|\bar{r}\alpha| < 0$, so ist analog wie oben zu sehen, dass im n dimensionalen Raum diejenige Wandung w'' , die $-\bar{r}$ enthält, auch $|w''\alpha| < 0$ liefern müsste, d. h. es läge α nicht in \mathfrak{P} .

In dem $(n-1)$ dimensionalen Polyeder $\bar{r}, \bar{r}_1, \bar{r}_2, \dots$ gibt es nach Voraussetzung ein Elementarpolyeder, in dessen Innern α liegt. Sind $\bar{r}, \bar{r}_1, \dots, \bar{r}_{n-1}$ seine Wandungen, so liegt, wie nunmehr leicht zu sehen, α auch innerhalb des n dimensionalen Elementarpolyeders mit den Wandungen

$$-(ra), \quad -(r_1a) \quad \dots \quad -(r_{n-1}a),$$

die $(n+1)$ te Wandung wird von den n Eckpunkten von r, r_1, \dots, r_{n-1} gebildet. Sie liefert, da alle $\xi^{(n)} = 0$ für α eine verschwindende Determinante.

Damit ist der angekündigte Satz bewiesen. Sind nun \mathfrak{P} und $\bar{\mathfrak{P}}$ zwei konvexe Polyeder, und tritt der Fall des vorigen Paragraphen, dass es keinen gemeinsamen Innenpunkt geben möge, nicht ein, so können wir zwei Elementarpolyeder P und \bar{P} , wo P nur Eckpunkte von \mathfrak{P} , \bar{P} nur Eckpunkte von $\bar{\mathfrak{P}}$ enthält, in solcher Weise angeben, dass auch P und \bar{P} einen gemeinsamen Innenpunkt haben.

§ 6. Auswahl der $(n+2)$ Punkte bei Unmöglichkeit der Trennung.

Unter einem Begrenzungselement eines konvexen Polyeders verstehen wir die Gesamtheit der Punkte p , bei denen von

den Bedingungen $|wp| \leq 0$, $|w'p| \leq 0$, $|w''p| \leq 0 \dots$ eine bestimmte Anzahl Male das Gleichheitszeichen gilt. Ist das gegebene konvexe Polyeder ein Elementarpolyeder, so unterscheiden wir nach dieser Anzahl die Ordnung des Begrenzungselementes. Es mögen die Punkte

$$1, 2, 3 \dots n + 1$$

ein Elementarpolyeder bilden. Es sei dann das Begrenzungselement

$$(1, 2 \dots \nu)$$

die Gesamtheit der Punkte, bei denen diejenigen Determinanten verschwinden, in denen die Punkte $1, 2 \dots \nu$ alle vorkommen. Dies sind, da wir unter den $n + 1 - \nu$ nicht in $(1, 2 \dots \nu)$ enthaltenen Punkten jeden zur Bildung einer Determinante weglassen können, $n + 1 - \nu$ Gleichungen. Ist μ die Ordnung eines Begrenzungselementes und ν die Anzahl seiner Eckpunkte, so gilt

$$\mu = n + 1 - \nu.$$

Verbinden wir zwei in einem Begrenzungselement liegende Punkte durch eine Gerade, so ist klar, dass für dieses Begrenzungselement dieselben Sätze gelten, die wir S. 68 für das ganze Polyeder ableiteten. Lassen wir den Punkt t auf dieser Geraden wandern, so muss nach einem endlichen Stück eine der noch positiven Determinanten des Elementarpolyeders verschwinden. Jede dieser Determinanten enthält die übrigen $n + 1 - \nu$ Punkte und $\nu - 1$ aus der Zahl der $1, 2, \dots \nu$. Möge etwa die den Punkt ν nicht enthaltende verschwinden, so heisst dies, wir gelangen in das Begrenzungselement $(\mu + 1)$ ter Ordnung $(1, 2, 3 \dots \nu - 1)$. Wir können daher sagen: Ein Begrenzungselement n ter Ordnung wird von lauter Begrenzungselementen $(\mu + 1)$ ter Ordnung begrenzt.

Wir beweisen nun den Satz: Eine lineare Funktion

$$p^{(1)}x^{(1)} + p^{(2)}x^2 + \dots + p^{(n)}x^{(n)} + p^{(n+1)},$$

die an den Eckpunkten $1, 2, \dots \nu$ positiv ist, ist dies auch überall im Innern des Begrenzungselementes. Wir betrachten zunächst die Begrenzungselemente mit 2 Eckpunkten

$$(1, 2) \quad (1, 3) \quad (1, 4) \dots (\nu - 1, \nu).$$

Eine lineare Funktion, die bei 1 und 2 positiv ist, kann im Innern von $(1, 2)$ weder negativ noch 0 sein, weil lineare Funktionen

kein Minimum haben. Wir gehen nun zu den Begrenzungselementen der nächst niedrigeren Ordnung d. h. zu

$$(1, 2, 3) \quad (1, 2, 4) \dots \dots \quad (\nu-2, \nu-1, \nu).$$

Es liege p im Innern von $(1, 2, 3)$. Ich ziehe von 1 eine Linie nach p . Diese muss nach dem, was wir eben sahen, ein Begrenzungselement höherer Ordnung schneiden, und daher sowohl an dem Schnittpunkt als auch bei 1 positiv sein, sie kann demnach bei p , das zwischen 1 und dem Schnittpunkt liegt, nicht anders als gleichfalls positiv sein. Und in dieser Weise fahren wir fort.

Es seien nun P und \bar{P} zwei Elementar-Polyeder, und wir wollen annehmen, es gäbe einen Punkt i , der sowohl im Innern von P als auch im Innern von \bar{P} liege. Ich behaupte, es lassen sich von den Eckpunkten von P und \bar{P} $(n+2)$ von der Art auslesen, dass es unmöglich ist, eine lineare Funktion

$$p^{(1)}x^{(1)} + p^{(2)}x^{(2)} + \dots \dots p^{(n)}x^{(n)} + p^{(n+1)}$$

anzugeben, die bei denjenigen der ausgewählten Punkte, die Eckpunkte von P sind, positiv, und bei denjenigen, die Eckpunkte von \bar{P} sind, negativ ist.

Wir ziehen durch i irgend eine Gerade und verfolgen sie von i aus in einer beliebigen Richtung. Die Gerade kann nicht immer im Innern der beiden Polyeder verlaufen. Sie möge etwa aus P zuerst austreten und zwar bei dem Punkte i' ; i' liegt dann auf einem Begrenzungselement erster Ordnung. Zu der Gleichung dieses Begrenzungselementes nehmen wir noch $n-2$ beliebige Gleichungen, die durch i' befriedigt werden, hinzu und erhalten hierdurch eine Gerade. Verfolgen wir wieder diese von i' an, so sei i'' der erste Punkt, bei dem sie aus einem der Polyeder austritt. i'' kann entweder auf einem Begrenzungselement zweiter Ordnung von P oder auf einem Begrenzungselement erster Ordnung von \bar{P} liegen. Jedenfalls genügt i'' zwei Gleichungen, zu denen wir nun noch $(n-3)$ hinzunehmen, um so wieder eine Gerade zu bilden und zu einem Punkt i''' zu gelangen. Das Weitergehen auf einer Geraden ist erst dann unmöglich, wenn der Punkt n Gleichungen genügt.

Wir können also einen Punkt i^* bestimmen, der auf zwei Begrenzungselementen der Ordnung μ und $\bar{\mu}$ liegt, wo

$$\mu + \bar{\mu} = n$$

und da

$$\mu = n + 1 - \nu \qquad \bar{\mu} = n + 1 - \bar{\nu},$$

so ist

$$n + 1 - \nu + n + 1 - \bar{\nu} = n$$

$$\nu + \bar{\nu} = n + 2.$$

Eine lineare Funktion, die an diesen ν Punkten positiv und an diesen $\bar{\nu}$ negativ wäre, müsste bei i^* zugleich positiv und negativ sein, sie kann also nicht existieren.

Wir haben bisher stets vorausgesetzt, dass sich sowohl unter den gegebenen Punkten $p_1, p_2, p_3 \dots$ als auch unter $\bar{p}_1, \bar{p}_2, \bar{p}_3 \dots$ ($n + 1$) finden mögen, die eine nichtverschwindende ($n + 1$) reihige Determinante liefern. Dies braucht aber nicht immer der Fall zu sein. Es können etwa alle ($\nu + 1$) reihigen und höheren Determinanten der p verschwinden. Wir nehmen dann eine Koordinaten-Transformation vor und ordnen die Punkte p in dem $-\nu$ dimensionalen Raum zu einem convexen Polyeder. Dieses ergänzen wir nach der Methode des § 5 durch successive Hinzunahme von Punkten aus den übrigen Dimensionen zu einem n dimensionalen convexen Polyeder. In diesem giebt es dann Begrenzungselemente, deren Eckpunkte nur aus gegebenen Punkten bestehen, und in deren Innern alle gegebenen Punkte liegen. Mit den Punkten \bar{p} verfahren wir ebenso. Nun stellen wir die Alternative der Existenz oder Nichtexistenz eines gemeinsamen Innenpunktes nicht bezüglich der ganzen Polyeder, sondern bezüglich der ausgezeichneten Begrenzungselemente. Man sieht sofort, dass man auch auf diese die Methoden von § 4 und § 6 anwenden kann.

Kapitel V. Das Interpolationsproblem bei zwei Variablen.

§ 1. Die allgemeinen Bedingungen.

Unsere Bemerkung S. 30, dass bei der annäherungsweisen Darstellung der Funktionen einer Variablen sich die wesentlichsten Resultate schon bei dem einfacheren Fall des Interpolationsproblem es ergeben, lässt es gerechtfertigt erscheinen, dass wir uns bei dem Fall zweier Variablen auf dieses beschränken.

Es sei eine Anzahl Punkte

$$p_1 = (x_1 y_1 z_1), \quad p_2 = (x_2 y_2 z_2) \dots p_v = (x_v y_v z_v)$$

gegeben und eine ganze rationale Funktion

$$\begin{aligned} z = f(x, y) = & a_1 x^n + a_2 x^{n-1} y + \dots + a_{n+1} y^n \\ & + a_{n+2} x^{n-1} + a_{n+3} x^{n-2} y + \dots \\ & \dots + a_{m-2} x + a_{m-1} y + a_m \end{aligned}$$

gesucht, von der Art, dass wenn $p_\mu = (x_\mu y_\mu z_\mu)$ einer der gegebenen Punkte ist, die grösste der Differenzen

$$| f(x_\mu y_\mu) - z_\mu |$$

möglichst klein sei. Den Existenzbeweis für $f(x, y)$ haben wir bereits S. 10—11 geführt. Wir wollen vorläufig, um eine einfachere Schreibweise zu haben,

$$n = 2$$

und

$$f(x, y) = a_1 x^2 + a_2 xy + a_3 y^2 + a_4 x + a_5 y + a_6$$

annehmen. Es bezeichne L wieder die Abweichung, d. h. die grösste der Differenzen $| f(x_\mu y_\mu) - z_\mu |$; sie werde an den Punkten

$$(x_1 y_1), (x_2 y_2) \dots (x_v y_v)$$

angenommen; wir haben dann, wenn

$$\varepsilon_2, \varepsilon_3 \dots \varepsilon_v$$

gleich ± 1 sind, nach Fixierung des Vorzeichens von L die Gleichungen:

$$(1) \quad \begin{aligned} a_1 x_1^2 + a_2 x_1 y_1 + a_3 y_1^2 + a_4 x_1 + a_5 y_1 + a_6 - z_1^2 &= L \\ a_1 x_2^2 + a_2 x_2 y_2 + a_3 y_2^2 + a_4 x_2 + a_5 y_2 + a_6 - z_2 &= \varepsilon_2 L \\ \vdots & \\ a_1 x_v^2 + a_2 x_v y_v + a_3 y_v^2 + a_4 x_v + a_5 y_v + a_6 - z_v &= \varepsilon_v L. \end{aligned}$$

Es habe nun

$$\begin{array}{llll} s_1 & \text{das Vorzeichen von } L & \text{und sei} & \neq 0 \\ s_2 & \text{''} & \text{''} & \varepsilon_2 L \text{ '' '' } \neq 0 \\ \vdots & & & \\ \vdots & & & \\ s_v & \text{''} & \text{''} & \varepsilon_v L \text{ '' '' } \neq 0. \end{array}$$

Im Uebrigen seien diese Grössen s unbestimmt. Wenn es eine Funktion $g(x,y)$ von kleinerer Abweichung als $f(x,y)$ giebt, so muss $|g(x,y) - z|$ an allen Punkten $x_1y_1, x_2y_2, \dots, x_vy_v$ kleiner sein als $|f(x,y) - z|$ d. h. es muss

$$\begin{aligned} f(x_1, y_1) - g(x_1, y_1) &= s_1 \\ f(x_2, y_2) - g(x_2, y_2) &= s_2 \\ &\vdots \\ f(x_v, y_v) - g(x_v, y_v) &= s_v. \end{aligned}$$

Und umgekehrt: Nehmen wir an, es gäbe eine Funktion $g(x,y)$, die diese Gleichungen für irgend ein Wertsystem s befriedigte, so giebt es auch eine Funktion von kleinerer Abweichung als $f(x,y)$. Denn

$$f(x,y) - \varepsilon(f(x,y) - g(x,y))$$

erfüllt offenbar für genügend kleines ε diese Bedingung, wobei natürlich L von 0 verschieden vorausgesetzt wird. Sei

$$f(x,y) - g(x,y) = b_1x^2 + b_2xy + b_3y^2 + b_4x + b_5y + b_6,$$

so haben wir den Satz:

Notwendige und hinreichende Bedingung dafür, dass es keine Funktion kleinerer Abweichung giebt als $f(x,y)$, ist die Unauflösbarkeit der Gleichungen

$$(2) \quad \begin{aligned} b_1x_1^2 + b_2x_1y_1 + b_3y_1^2 + b_4x_1 + b_5y_1 + b_6 &= s_1 \\ b_1x_2^2 + b_2x_2y_2 + b_3y_2^2 + b_4x_2 + b_5y_2 + b_6 &= s_2 \\ &\vdots \\ b_1x_v^2 + b_2x_vy_v + b_3y_v^2 + b_4x_v + b_5y_v + b_6 &= s_v \end{aligned}$$

für jedes mögliche Wertsystem s .

Die Zahl der Unbekannten ist 6, die Zahl der Gleichungen ν . Im allgemeinen Fall können diese Gleichungen nur dann unauflösbar sein, wenn $\nu > 6$. Die speziellen Fälle, wo $\nu \leq 6$ werden wir später untersuchen und nehmen nun $\nu = 7$. Es wird alsdann Unauflösbarkeit gefordert von dem System:

$$\begin{aligned}
b_1 x_1^2 + b_2 x_1 y_1 + b_3 y_1^2 + b_4 x_1 + b_5 y_1 + b_6 &= s_1 \\
b_1 x_2^2 + b_2 x_2 y_2 + b_3 y_2^2 + b_4 x_2 + b_5 y_2 + b_6 &= s_2 \\
\vdots & \\
b_1 x_7^2 + b_2 x_7 y_7 + b_3 y_7^2 + b_4 x_7 + b_5 y_7 + b_6 &= s_7.
\end{aligned}$$

Die Bedingung der Unauflösbarkeit für dieses System lautet aber :

$$D = \begin{vmatrix} x_1^2 & x_1 y_1 & y_1^2 & x_1 & y_1 & 1 & s_1 \\ x_2^2 & x_2 y_2 & y_2^2 & x_2 & y_2 & 1 & s_2 \\ \cdot & & & & & & \\ \cdot & & & & & & \\ x_7^2 & x_7 y_7 & y_7^2 & x_7 & y_7 & 1 & s_7 \end{vmatrix} \neq 0.$$

Entwickeln wir nach den Elementen der letzten Spalte, so haben wir, abgekürzt geschrieben :

$$D = s_1 \cdot (2, 3, \dots 7) - s_2 \cdot (1, 3, \dots 7) + \dots + s_7 \cdot (1, 2, \dots 6).$$

Da die s bis auf das Vorzeichen vollkommen willkürlich sind, die Determinante aber für alle mit der Vorzeichenbedingung verträglichen s von 0 verschieden sein soll, so sieht man leicht die Bedingung: Die Grössen

$$\begin{aligned}
& s_1 \cdot (2, 3 \dots 7), \\
& -s_2 \cdot (1, 3 \dots 7), \\
& \cdot \\
& \cdot \\
& s_7 \cdot (1, 2 \dots 6)
\end{aligned}$$

müssen gleiches Vorzeichen haben. Nehmen wir s_1 etwa positiv, so bestimmt sich hieraus das Vorzeichen der übrigen s .

Sind 7 Punkte bekannt, an denen die Abweichung angenommen wird, und wird der Sinn der Annahme bei einem Punkt willkürlich als positiv bezeichnet, so liefert uns $D \neq 0$ den Sinn, in dem die Abweichung an allen andern Stellen angenommen wird. Und zwar bestimmt sich dieser Sinn aus dem Vorzeichen einer Determinante, die aus den Projektionen der übrigen 6 Punkte gebildet wird. Die

z Koordinate ist ohne Einfluss darauf. — Ist eine der Determinanten z. B. $(1, 3 \dots 7) = 0$, so kann s_2 beliebig positiv oder negativ angenommen werden. Nunmehr können wir, nachdem die Vorzeichen der s und damit die ε bekannt sind, die Annäherungsfunktion $f(x, y)$ leicht bestimmen; wir haben dazu nur das System der linearen Gleichungen (1) zu lösen.

$$\begin{aligned} a_1x_1^2 + a_2x_1y_1 + a_3y_1^2 + a_4x_1 + a_5y_1 + a_6 - \varepsilon_1L &= z_1 \\ a_1x_2^2 + a_2x_2y_2 + a_3y_2^2 + a_4x_2 + a_5y_2 + a_6 - \varepsilon_2L &= z_2 \\ &\vdots \\ a_1x_7^2 + a_2x_7y_7 + a_3y_7^2 + a_4x_7 + a_5y_7 + a_6 - \varepsilon_7L &= z_7. \end{aligned}$$

Das System dieser 7 Gleichungen mit den 7 Unbekannten $a_1 \dots a_6$ und L ist sicher eindeutig lösbar. Denn die Determinante des Systems unterscheidet sich von D nur dadurch, dass statt der s hier die ε auftreten; sie ist also sicher von 0 verschieden.

Wir haben nachgewiesen, dass es eine Funktion kleinerer Abweichung als die so konstruierte Funktion $f(x, y)$ nicht geben kann. Es fragt sich, ob und wann es eine von $f(x, y)$ verschiedene Funktion von ebensogrosser Abweichung geben kann. Nehmen wir an, $g(x, y)$ sei eine solche Funktion, so muss, gleichviel an welchen Punkten $g(x, y)$ seinerseits die Abweichung annimmt, an den 7 Punkten, an denen $f(x, y)$ seine Abweichung annimmt,

$$|f(x_\mu y_\mu) - z_\mu| \leq |g(x_\mu y_\mu) - z_\mu|,$$

d. h. es darf $f(x, y) - g(x, y)$ an den Punkten $x_\mu y_\mu$ nicht entgegengesetztes Vorzeichen haben wie s_μ , sondern es muss entweder gleiches Vorzeichen haben oder verschwinden, was wir durch s'_μ andeuten wollen. Dies ergibt für die Koeffizienten der Differenz $f(x, y) - g(x, y)$ die Bedingungen:

$$\begin{aligned} b_1x_1^2 + b_2x_1y_1 + b_3y_1^2 + b_4x_1 + b_5y_1 + b_6 &= s'_1 \\ b_1x_2^2 + b_2x_2y_2 + b_3y_2^2 + b_4x_2 + b_5y_2 + b_6 &= s'_2 \\ &\vdots \\ b_1x_7^2 + b_2x_7y_7 + b_3y_7^2 + b_4x_7 + b_5y_7 + b_6 &= s'_7. \end{aligned}$$

Die Bedingung für die Auflösbarkeit ist das Verschwinden der Determinante. Ist keine der Unterdeterminanten gleich Null,

so kann dies nur eintreten, wenn $s'_1 = s'_2 = \dots = s'_7 = 0$ und alsdann müssen in den übrig bleibenden homogenen Gleichungen alle b verschwinden, d. h. $f(x, y)$, ist eindeutig bestimmt.

Ist aber z. B. $(1, 2 \dots 6) = 0$, so kann s'_7 von 0 verschieden sein, und es brauchen nur $s'_1 = s'_2 = \dots = s'_6 = 0$. Die hieraus resultierenden Gleichungen sind nicht homogen und liefern nach willkürlicher Festsetzung von s'_7 ein bestimmtes System der Koeffizienten von $f(x, y) - g(x, y)$. An allen Punkten, an denen $f(x, y)$ die Abweichung annimmt, nimmt sie auch $g(x, y)$ an, nur an dem Punkte x, y_7 braucht dies, wie aus der willkürlichen Festsetzung von s'_7 hervorgeht, nicht der Fall zu sein.

§ 2. Der Fall verschwindender Determinante.

Wir werden also auf den oben ausgeschlossenen Fall geführt, dass die Abweichung an weniger als 7 Punkten, also etwa 6 angenommen wird. Wie man sieht, kann dies nur dann der Fall sein, wenn die Projektionen der 6 Punkte auf einem Kegelschnitt liegen. Alsdann ist die Annäherungsfunktion nicht eindeutig bestimmt. Man sieht leicht, innerhalb welcher Grenzen die Annäherungsfunktion willkürlich ist. Es sei $f(x, y)$ die Annäherungsfunktion, die die Abweichung an 6 Punkten annehme. $g(x, y) = 0$ die Gleichung des Kegelschnittes, dem die Projektionen genügen:

$$| f(x, y) + \lambda \cdot g(x, y) - \varepsilon |$$

ist dann für $\lambda = 0$ an allen Punkten, an denen die Abweichung nicht angenommen wird, kleiner als L . Wenn nun nicht alle Projektionen der gegebenen Punkte auf dem Kegelschnitt $g(x, y) = 0$ liegen, so ergeben sich Grenzen für λ , die nicht überschritten werden dürfen, ohne dass $| f(x, y) + \lambda g(x, y) - \varepsilon |$ bei mindestens einem der gegebenen Punkte grösser wird als L und damit

$$f(x, y) + \lambda g(x, y)$$

aufhört Annäherungsfunktion zu sein. Setze ich für λ einen dieser Grenzwerte ein, so wird die Abweichung an 7 Punkten angenommen. Wir können also stets erreichen, dass die Abweichung an mindestens 7 Punkten angenommen wird, mit Ausnahme des Falles, wo die Projektionen aller gegebenen Punkte auf einem

Kegelschnitt liegen. — Analog verhält es sich, wenn auch noch niedrigere Determinanten der Projektionen verschwinden, was geometrisch bedeutet, dass alle Projektionen mit Ausnahme einer einzigen auf einer Geraden liegen, oder dass sie alle auf einer Geraden liegen.

Wir wollen zum Fall $\nu = 6$ noch einige Bemerkungen machen. Das System der Gleichungen

$$(2) \quad \begin{aligned} b_1 x_1^2 + b_2 x_1 y_1 + b_3 y_1^2 + b_4 x_1 + b_5 y_1 + b_6 &= s_1 \\ b_1 x_2^2 + b_2 x_2 y_2 + b_3 y_2^2 + b_4 x_2 + b_5 y_2 + b_6 &= s_2 \\ &\vdots \\ b_1 x_6^2 + b_2 x_6 y_6 + b_3 y_6^2 + b_4 x_6 + b_5 y_6 + b_6 &= s_6 \end{aligned}$$

kann nur dann unauflösbar sein, wenn die Determinante verschwindet. Aber diese Bedingung ist zur Unauflösbarkeit noch nicht hinreichend; es muss sich vielmehr aus der Matrix:

$$\begin{array}{ccccccc} x_1^2 & x_1 y_1 & y_1^2 & x_1 & y_1 & 1 & s_1 \\ x_2^2 & x_2 y_2 & y_2^2 & x_2 & y_2 & 1 & s_2 \\ & \vdots & & & & & \\ & \vdots & & & & & \\ x_6^2 & x_6 y_6 & y_6^2 & x_6 & y_6 & 1 & s_6 \end{array}$$

eine nichtverschwindende 6 reihige Determinante bilden lassen; hieraus bestimmen sich die s . Zunächst sieht man, dass es gleichgültig ist, welche der 6 reihigen Determinanten wir dabei zur Untersuchung heranziehen. Verschwindet eine, so verschwinden auch alle andern. Es möge z. B. die durch Weglassung der 5ten Vertikalreihe entstehende Determinante verschwinden. Unsere Behauptung lautet dann in abgekürzter Schreibweise, dass wir aus

$$| x^2, xy, y^2, x, y, 1 | = 0$$

und

$$| x^2, xy, y^2, x, y, s | = 0$$

z. B. auf

$$| x^2, xy, y^2, x, 1, s |$$

schliessen können. Dies ist ein spezieller Fall des auf S. 67 bewiesenen allgemeinen Satzes. Nur dürfen die 5 ausgezeichneten Vertikalreihen, die mit den s zusammen die zu untersuchende Determinante bilden sollen, nicht gerade so gewählt sein, dass

alle 5reihigen Determinanten verschwinden. Dies ist möglich, denn wir setzen voraus, dass nicht alle 5reihigen Determinanten der Matrix verschwinden, in welchem Fall wir $\nu = 5$ setzen müssten.

Es genügt also, das Vorzeichen der s so zu bestimmen, dass für dies einmal bestimmte Vorzeichensystem für alle s

$$\begin{vmatrix} x_1^2 & x_1 y_1 & y_1^2 & x_1 & y_1 & s \\ x_2^2 & x_2 y_2 & y_2^2 & x_2 & y_2 & s \\ \cdot & & & & & \\ \cdot & & & & & \\ x_6^2 & x_6 y_6 & y_6^2 & x_6 & y_6 & s \end{vmatrix} \neq 0.$$

Man sieht leicht, dass die Einschränkung, es sollen nicht alle 5reihigen Determinanten $|x_i^2, xy, y^2, x, y|$ verschwinden in unserm Fall des nichtzerfallenden Kegelschnittes nichts anderes bedeutet, als dass der Punkt $x = 0 \quad y = 0$ nicht auf dem Kegelschnitt liegen soll.

Wir wollen eine einfache Regel über das Vorzeichen der s ableiten. Die Bedingung

$$\begin{vmatrix} x_1^2 & x_1 y_1 & y_1^2 & x_1 & y_1 & s_1 \\ x_2^2 & x_2 y_2 & y_2^2 & x_2 & y_2 & s_2 \\ \cdot & & & & & \\ \cdot & & & & & \\ x_6^2 & x_6 y_6 & y_6^2 & x_6 & y_6 & s_6 \end{vmatrix} \neq 0$$

lautet entwickelt:

$$s_1(2, 3 \dots 6) - s_2(1, 3 \dots 6) + s_3(1, 2, 4 \dots 6) - \dots \neq 0,$$

wo die Abkürzungen sofort verständlich sind. Wir denken uns nun die Punkte 1, 2, . . . 6 auf dem Kegelschnitt so angeordnet, dass wir sie bei einer Umlaufung des ganzen Kegelschnittes in dieser cyklischen Reihenfolge passieren. Ich behaupte, alsdann haben die Determinanten

$$(2, 3, \dots 6), \quad (1, 3, \dots 6) \dots$$

alle dasselbe Vorzeichen. Wir betrachten, um dies zu beweisen, die Determinante

$$\begin{vmatrix} x_2^2 & x_2 y_2 & y_2^2 & x_2 & y_2 & 1 \\ x_3^2 & x_3 y_3 & y_3^2 & x_3 & y_3 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_6^2 & x_6 y_6 & y_6^2 & x_6 & y_6 & 1 \\ x^2 & xy & y^2 & x & y & 1 \end{vmatrix}$$

Für den betrachteten Punkt xy möge sie nicht verschwinden. Wir lassen nun bei festgehaltenem xy den Punkt $x_2 y_2$ nach $x_1 y_1$ hinwandern, sodass er bei der Wanderung immer auf dem Kegelschnitt bleibt. Hierbei kann die Determinante nicht verschwinden; denn dies würde bedeuten, dass der Punkt xy auf dem von den 5 andern bestimmten Kegelschnitt läge, während er dies in der ursprünglichen Lage nicht that und der Kegelschnitt doch derselbe geblieben ist. Es haben also

$$\begin{vmatrix} x_1^2 & x_1 y_1 & y_1^2 & x_1 & y_1 & 1 \\ x_2^2 & x_2 y_2 & y_2^2 & x_2 & y_2 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_6^2 & x_6 y_6 & y_6^2 & x_6 & y_6 & 1 \\ x^2 & xy & y^2 & x & y & 1 \end{vmatrix} \quad \text{und} \quad \begin{vmatrix} x_2^2 & x_2 y_2 & y_2^2 & x_2 & y_2 & 1 \\ x_3^2 & x_3 y_3 & y_3^2 & x_3 & y_3 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_6^2 & x_6 y_6 & y_6^2 & x_6 & y_6 & 1 \\ x^2 & xy & y^2 & x & y & 1 \end{vmatrix}$$

dasselbe Vorzeichen. Für $x = 0$ $y = 0$ heisst dies nichts anderes, als dass (1, 3 . . . 6) und (2, 3 . . . 6) gleiches Vorzeichen haben, und zwar ohne 0 zu sein, da der Punkt $x = 0$ $y = 0$ nicht auf dem Kegelschnitt liegen soll.

Aus der Bedingung

$$s_1(2, 3 \dots 6) - s_2(1, 3 \dots 6) + s_3(1, 2, 4 \dots 6) - \dots \neq 0$$

können wir daher schliessen: die s müssen abwechselndes Vorzeichen haben.

Liegen die Projektionen von 6 Annahmestellen der Abweichung auf einem Kegelschnitt, so wechselt der Sinn der Abweichung ab, wenn wir den Kegelschnitt durchlaufen.

Dies Resultat gilt für $n > 2$ unverändert, wenn die betreffende Kurve eine Unikursalkurve

ist; andernfalls gilt für jeden einzelnen Ast ein entsprechender Satz.

§ 3. Allgemeine Lösung des Problems.

Es sei jetzt $\nu > 7$. Es sei xy die Projektion eines Punktes, an dem die Abweichung angenommen wird, und zwar wollen wir xy mit p bezeichnen, wenn $s > 0$ und mit \bar{p} wenn $s < 0$. Unsere Bedingung (2) S. 86 können wir dann auch so fassen: Es soll unmöglich sein, eine Funktion

$$b_1x^2 + b_2xy + b_3y^2 + b_4x + b_5y + b_6$$

anzugeben, die an allen Punkten p positive und an allen Punkten \bar{p} negative Werte annimmt.

Es folgt nun aus dem Satze des vorigen Kapitels: Wenn die Punkte p und \bar{p} so gelegen sind, dass sie die Existenz einer derartigen Funktion

$$b_1x^2 + b_2xy + b_3y^2 + b_4x + b_5y + b_6$$

ausschliessen, so kann man aus ihnen 7 derartige Punkte auswählen, dass schon sie eine Funktion der verlangten Form und Eigenschaft unmöglich machen.

Setzen wir

$$x^2 = \xi^{(1)} \quad xy = \xi^{(2)} \quad y^2 = \xi^{(3)} \quad x = \xi^{(4)} \quad y = \xi^{(5)}.$$

Bezeichnet p einen Punkt x_1y_1 , so bezeichne π den entsprechenden Punkt des 5-dimensionalen Raums

$$\pi_1 = \pi(\xi_1^{(1)} \xi_1^{(2)} \xi_1^{(3)} \xi_1^{(4)} \xi_1^{(5)})$$

und wir unterscheiden Punkte π und $\bar{\pi}$. Es kann nun keine Funktion

$$\beta_1\xi^{(1)} + \beta_2\xi^{(2)} + \beta_3\xi^{(3)} + \beta_4\xi^{(4)} + \beta_5\xi^{(5)} + \beta_6 \quad \Gamma \xi^{(5)}$$

geben, die an allen Punkten π positive und an allen Punkten $\bar{\pi}$ negative Werte annähme; denn es würde alsdann die Funktion

$$\beta_1x^2 + \beta_2xy + \beta_3y^2 + \beta_4x + \beta_5y + \beta_6$$

auch an allen Punkten p positive und an allen Punkten \bar{p} negative Werte annehmen. Ich wähle unter den Punkten π und $\bar{\pi}$

nach § 6 des vorigen Kapitels 7 aus und nehme dann die diesen 7 Punkten entsprechenden Punkte p und \bar{p} und behaupte: Es kann keine Funktion

$$b_1x^2 + b_2xy + b_3y^2 + b_4x + b_5y + b_6$$

geben, die an den ausgewählten Punkten p positive und an den ausgewählten Punkten \bar{p} negative Werte annimmt, denn sonst würde sich

$$b_1\xi^{(1)} + b_2\xi^{(2)} + b_3\xi^{(3)} + b_4\xi^{(4)} + b_5\xi^{(5)} + b_6$$

bei den entsprechenden Punkten π und $\bar{\pi}$ ebenso verhalten.

Daraus folgt: Nimmt die Annäherungsfunktion die Abweichung mehr als 7mal an, so lassen sich aus der Zahl der Annahmestellen 7 so auswählen, dass die Annäherungsfunktion auch Annäherungsfunktion dieser 7 Punkte allein ist.

Hieraus ergibt sich die Lösung des Problems: Man greift aus der Zahl der gegebenen Punkte 7 heraus und bestimmt ihre Annäherungsfunktion nach § 1. Man untersucht sodann, ob die so gefundene Funktion die Annäherungsfunktion für die Gesamtheit der gegebenen Punkte ist, indem man untersucht, ob $|f(x_\mu, y_\mu) - z_\mu| \leq L$ für alle gegebenen Punkte erfüllt ist. Dies ist, wenn man auf alle möglichen Arten 7 Punkte herausgreift einmal und im Wesentlichen auch nur einmal der Fall. Liegen die Projektionen aller gegebenen Punkte auf einem Kegelschnitt, so wählen wir nur 6 Punkte aus, und entsprechend, wenn die Unterdeterminanten von noch niedrigerem Grade verschwinden. Man sieht dies leicht, wenn man bedenkt, dass alsdann der den xy entsprechende ξ -Raum in Wahrheit um eine bzw. mehrere Dimensionen niedriger ist, als im allgemeinen Fall.

Der spezielle Fall, wo 6 Punkte auf einem Kegelschnitt liegen und die Annäherungsfunktion nicht völlig bestimmt ist, hat noch eine weitere Bedeutung. Man sieht nämlich sofort, dass die Annäherungsfunktion sich nicht unbedingt stetig ändern kann, wenn sich einer der gegebenen Punkte bewegt. Denn dann könnten wir von der Annäherungsfunktion, die zu irgend welchen 7 Punkten gehört durch stetige Änderung zu der Annäherungsfunktion gelangen, die zu irgend einer andern Lage der Punkte gehört. Ein Punkt, bei dem die Abweichung in positivem Sinn angenommen würde, könnte hierbei diese Eigenschaft nicht ver-

lieren und die Anzahl der positiven Annahmen und ebenso die der negativen Annahmen müsste von der Lage der Punkte unabhängig sein. Dies ist aber, wie wir wissen, nicht der Fall. Man sieht leicht: Die Annäherungsfunktion ändert sich im Allgemeinen stetig mit den anzunähernden Punkten; sie macht aber einen Sprung, wenn sich die Projektion eines Punktes durch den Kegelschnitt, den die Projektionen von 5 andern bilden, hindurchbewegt.

Den Hauptsatz des vorigen Kapitels können wir geometrisch für einen speziellen Fall so aussprechen: Es seien in der Ebene Punkte $p_1, p_2, p_3 \dots$ und $\bar{p}_1, \bar{p}_2, \bar{p}_3 \dots$ gegeben und ein Kegelschnitt gesucht, der die p von den \bar{p} trennt. Die Aufgabe ist entweder möglich, oder man kann 7 Punkte aus den gegebenen auswählen, die auch schon die Trennung durch einen Kegelschnitt unmöglich machen. Entsprechende Sätze gelten für die gerade Linie und für höhere Kurven.

Wir haben unsere Theorie für den Fall $n = 2$ durchgeführt, man sieht aber sofort, dass sie allgemein gültig ist. Die Anzahl der Annahmestellen der Abweichung ist im Allgemeinen um 1 grösser als die der zur Verfügung stehenden Parameter. Unsere Determinante D S. 87 lautet für $n = 3$:

$$\begin{vmatrix} x_1^3 & x_1^2 y_1 & x_1 y_1^2 & y_1^3 & x_1^2 x_1 y_1 & y_1^2 x_1 y_1 & y_1^2 x_1 y_1 & 1 & s_1 \\ x_2^3 & x_2^2 y_2 & x_2 y_2^2 & y_2^3 & x_2^2 x_2 y_2 & y_2^2 x_2 y_2 & y_2^2 x_2 y_2 & 1 & s_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{11}^3 & x_{11}^2 y_{11} & x_{11} y_{11}^2 & y_{11}^3 & x_{11}^2 x_{11} y_{11} & y_{11}^2 x_{11} y_{11} & y_{11}^2 x_{11} y_{11} & 1 & s_{11} \end{vmatrix}$$

und die s bestimmen sich aus $D \neq 0$. Ebenso sieht man leicht, dass die Theorie auch für mehr als zwei unabhängige Variable gilt; auch hier ändert sich nichts gegenüber dem eben behandelten Fall. Für 3 Variable und $n = 2$ lautet die Determinante:

$$\begin{vmatrix} x_1^2 & x_1 y_1 & x_1 z_1 & y_1^2 & y_1 z_1 & z_1^2 & x_1 y_1 z_1 & 1 & s_1 \\ x_2^2 & x_2 y_2 & x_2 z_2 & y_2^2 & y_2 z_2 & z_2^2 & x_2 y_2 z_2 & 1 & s_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{11}^2 & x_{11} y_{11} & x_{11} z_{11} & y_{11}^2 & y_{11} z_{11} & z_{11}^2 & x_{11} y_{11} z_{11} & 1 & s_{11} \end{vmatrix}$$

Ferner lässt die Theorie eine Uebertragung auf den Fall zu, wo die Form der Annäherungsfunktion weiteren Bedingungen unterworfen ist, z. B. homogen sein soll u. dergl.

Für alle diese Fälle können wir die folgenden Sätze aussprechen:

1) Die Abweichung wird im allgemeinen Fall einmal öfter angenommen, als die Zahl der Parameter der Annäherungsfunktion beträgt. Spezielle Fälle treten ein, wenn die Projektionen der gegebenen Punkte auf die Ebene der unabhängigen Variablen gewissen Determinantengleichungen genügen.

2) Der Sinn, in dem die Abweichung angenommen wird, wird durch eine Determinantenungleichung geliefert.

3) Alle diese Probleme sind durch Auflösung eines Systems linearer Gleichungen lösbar. Dieses System selbst wird aus einer endlichen Anzahl von Systemen durch Versuche bestimmt.

Wir haben uns auf das Interpolationsproblem beschränkt; man sieht aber leicht, wie die Theorie für das allgemeinere Annäherungsproblem zu erweitern ist: Soll eine ganze Funktion etwa $\varphi(x,y)$ in einem gegebenen Intervall etwa $-1 \leq x \leq 1$; $-1 \leq y \leq 1$ durch $f_n(v,y)$ angenähert dargestellt werden, so braucht die Abweichung nicht an einer endlichen Anzahl diskreter Punkte angenommen zu werden, sondern wir können hier Abweichungskurven haben. Wir projizieren diese in die xy -Ebene und ergänzen sie, falls sie noch nicht konvex sind, durch Hinzunahme geradliniger Strecken zu zwei konvexen Figuren. Für diese ist der Satz des vorigen Kapitels nachzuweisen, was hier zu weit führen dürfte. Ist m die Anzahl der verfügbaren Parameter der gesuchten Annäherungsfunktion, so wird die Abweichung an mindestens $m+1$ Stellen angenommen. Es lassen sich $m+1$ Punkte der gegebenen Funktion so bestimmen, dass die Annäherungsfunktion dieser $m+1$ Punkte die gesuchte Annäherungsfunktion der gegebenen Funktion ist. Dieser Satz stellt die genaue Verallgemeinerung des von Tchebychef für den Fall einer Variablen ausgesprochenen Resultates dar. Die Frage nach einer Methode der Auswahl der $m+1$ Punkte, die wir beim Interpolationsproblem durch eine endliche Anzahl von Versuchen erledigen konnten, bedarf einer weitem Untersuchung.

I n h a l t.

Einleitung	5
Kapitel I. Allgemeine Theorie bei einer reellen Veränderlichen.	
§ 1. Definition	6
§ 2. Existenzbeweis	7
§ 3. Notwendige Bedingungen	11
§ 4. Diese Bedingungen sind eindeutig und hinreichend	16
§ 5. Stetigkeit	18
§ 6. Exkurs über die Darstellung durch rationale Funktionen	21
§ 7. Interpolationsprobleme	26
Kapitel II. Aufstellung der Annäherungsfunktion.	
§ 1. Analytische Formulierung	30
§ 2. Aufstellung für den einfachsten Fall	32
§ 3. Die Annäherungsfunktion beliebiger analytischer Funktionen	34
Kapitel III. Rand- und Nebenbedingungen.	
§ 1. Eine Randbedingung	49
§ 2. Eine Nebenbedingung	51
Kapitel IV. Von den konvexen Polyedern.	
§ 1. Polyeder	60
§ 2. Konvexe Polyeder	65
§ 3. Anordnung zu konvexen Polyedern	71
§ 4. Trennung zweier konvexer Polyeder	75
§ 5. Zerlegung in Elementarpolyeder	78
§ 6. Auswahl der $(n + 2)$ Punkte bei Unmöglichkeit der Trennung	81
Kapitel V. Das Interpolationsproblem bei zwei Variablen.	
§ 1. Die allgemeinen Bedingungen	84
§ 2. Der Fall verschwindender Determinante	89
§ 3. Allgemeine Lösung des Problems	98

Lebenslauf

Ich, Paul Kirchberger, wurde am 23sten Juni 1878 zu Niederlahnstein am Rhein als Sohn des Kaufmanns Theodor Kirchberger geboren. In Weilburg, wohin meine Eltern alsbald verzogen, besuchte ich die Volksschule und das Königliche Gymnasium, das ich Ostern 1897 mit dem Zeugnis der Reife verliess. Um Mathematik und Naturwissenschaften zu studieren, bezog ich zunächst die Berliner Universität und besuchte Vorlesungen und Uebungen u. A. bei den Herren Professoren und Dozenten: Blasius, Fischer, Frobenius, Fuchs (†), Hensel, Knoblauch, Planck, Rosenheim, H. A. Schwarz, Warburg. Von Berlin, wo mich insbesondere die Vorlesungen des Herrn Professors Frobenius so lange festgehalten hatten, wandte ich mich Ostern 1900 nach Göttingen. Hier hörte ich Vorlesungen und Uebungen bei den Herren Professoren und Dozenten: Abraham, Baumann, Hilbert, Klein, v. Koenen, Liebisch, G. E. Müller, Rehnisch (†), Riecke, Schilling, Voigt, Wallach. Mein Staatsexamen für das Lehramt an höheren Schulen bestand ich am 24. Januar 1902.

Allen meinen verehrten Lehrern schulde ich herzlichsten Dank; namentlich hat mich Herr Professor Hilbert durch sein ständiges förderndes Interesse, sowie durch manchen wertvollen Rat zu tiefer und dauernder Dankbarkeit verpflichtet. Auch einigen Kollegen vom Göttinger Mathematischen Verein bleibe ich dankbar für mancherlei Anregung und Aufklärung.
