

Werk

Titel: Mathematische Annalen

Jahr: 1928

Kollektion: Mathematica

Digitalisiert: Niedersächsische Staats- und Universitätsbibliothek Göttingen

Werk Id: PPN235181684_0099

PURL: http://resolver.sub.uni-goettingen.de/purl?PPN235181684_0099

Übergeordnetes Werk

Werk Id: PPN235181684

PURL: <http://resolver.sub.uni-goettingen.de/purl?PPN235181684>

OPAC: <http://opac.sub.uni-goettingen.de/DB=1/PPN?PPN=235181684>

Terms and Conditions

The Goettingen State and University Library provides access to digitized documents strictly for noncommercial educational, research and private purposes and makes no warranty with regard to their use for other purposes. Some of our collections are protected by copyright. Publication and/or broadcast in any form (including electronic) requires prior written permission from the Goettingen State- and University Library.

Each copy of any part of this document must contain these Terms and Conditions. With the usage of the library's online system to access or download a digitized document you accept the Terms and Conditions.

Reproductions of material on the web site may not be made for or donated to other repositories, nor may be further reproduced without written permission from the Goettingen State- and University Library.

For reproduction requests and permissions, please contact us. If citing materials, please give proper attribution of the source.

Contact

Niedersächsische Staats- und Universitätsbibliothek Göttingen
Georg-August-Universität Göttingen
Platz der Göttinger Sieben 1
37073 Göttingen
Germany
Email: gdz@sub.uni-goettingen.de

MATHEMATISCHE ANNALEN

BEGÜNDET 1868 DURCH

ALFRED CLEBSCH UND CARL NEUMANN

FORTGEFÜHRT DURCH

FELIX KLEIN

UNTER MITWIRKUNG

VON

LUDWIG BIEBERBACH, HARALD BOHR, L. E. J. BROUWER,
RICHARD COURANT, WALTHER V. DYCK, OTTO HÖLDER,
THEODOR V. KÁRMÁN, ARNOLD SOMMERFELD

GEGENWÄRTIG HERAUSGEGEBEN

VON

DAVID HILBERT

IN GÖTTINGEN

ALBERT EINSTEIN

IN BERLIN

OTTO BLUMENTHAL

IN AACHEN

CONSTANTIN CARATHÉODORY

IN MÜNCHEN

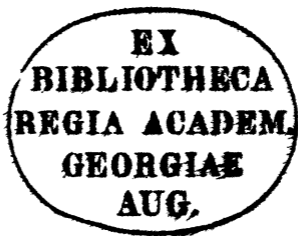
99. BAND



BERLIN

VERLAG VON JULIUS SPRINGER

1928



1426. 7162

Handwritten numbers in a cursive script, likely indicating a library inventory or accession number.

Inhalt des neunundneunzigsten Bandes.

(In alphabetischer Ordnung.)

	Seite
Ackermann, W. , in Göttingen. Zum Hilbertschen Aufbau der reellen Zahlen	118
Bergmann, St. , in Berlin. Zur Theorie der algebraischen Potentialfunktionen des dreidimensionalen Raumes. I.	629
Bernays, P. , in Göttingen, und M. Schönfinkel in Moskau. Zum Entscheidungsproblem der mathematischen Logik	342
Brandt, H. , in Aachen. Idealtheorie in Quaternionenalgebren	1
Cohn-Vossen, St. , in Göttingen. Die parabolische Kurve. Beitrag zur Geometrie der Berührungstransformationen, der partiellen Differentialgleichungen zweiter Ordnung und der Flächenverbiegung	273
Berichtigung dazu	752
Doetsch, G. , in Stuttgart. Die Funktionaldeterminante als Deformationsmaß einer Abbildung und als Kriterium der Abhängigkeit von Funktionen	590
Egerváry, E. , in Budapest. Über gewisse Extremumprobleme der Funktionentheorie	542
Fraenkel, A. , in Marburg. Zusatz zu dem Aufsatz Herrn v. Neumanns auf S. 373	392
Friedrichs, K. , und H. Lewy in Göttingen. Das Anfangswertproblem einer beliebigen nichtlinearen hyperbolischen Differentialgleichung beliebiger Ordnung in zwei Variablen. Existenz, Eindeutigkeit und Abhängigkeitsbereich der Lösung	200
Frommer, M. , in Stuttgart. Die Integralkurven einer gewöhnlichen Differentialgleichung erster Ordnung in der Umgebung rationaler Unbestimmtheitsstellen	222
Furtwängler, Ph. , in Wien. Über die simultane Approximation von Irrationalzahlen. Zweite Mitteilung	71
Geppert, H. , in Gießen. Zur Theorie des arithmetisch-geometrischen Mittels	162
Hoheisel, G. , in Breslau. Eine Illustration zur Riemannschen Vermutung	150
Jonas, H. , in Berlin-Steglitz. Über eine neue geometrische Eigenschaft der Bianchischen Transformation der auf die Mittelpunktsflächen zweiten Grades abwickelbaren Flächen	435
Kamke, E. , in Tübingen. Zur Theorie der Differentialgleichungen	602
Kolmogoroff, A. , in Moskau. Über die Summen durch den Zufall bestimmter unabhängiger Größen	309
Kritikos, N. , in München. Über analytische Abbildungen einer Klasse von vierdimensionalen Gebieten	321
Krull, W. , in Freiburg i. Br. Zur Theorie der allgemeinen Zahlringe	51
Lachmann, K. , in Berlin. Beitrag zum Schwingungsproblem von Duffing	479
Lewy, H. , und K. Friedrichs in Göttingen. Das Anfangswertproblem einer beliebigen nichtlinearen hyperbolischen Differentialgleichung beliebiger Ordnung in zwei Variablen. Existenz, Eindeutigkeit und Abhängigkeitsbereich der Lösung	200
Mehmke, E. , in Stuttgart. Zum Rechnen mit Potenzreihen	616
v. Neumann, J. , in Göttingen. Ein System algebraisch unabhängiger Zahlen	134
v. Neumann, J. , in Göttingen. Über die Definition durch transfinite Induktion und verwandte Fragen der allgemeinen Mengenlehre	373
Newman, M. H. A. , in Cambridge (England). Topological equivalence of complexes	399

	Seite
Ore, Ö. , in Oslo. Newtonsche Polygone in der Theorie der algebraischen Körper	84
Pólya, G. , in Zürich. Über gewisse notwendige Determinantenkriterien für die Fortsetzbarkeit einer Potenzreihe	687
X Saxer, W. , in Zürich. Über quasi-normale Funktionsscharen und eine Verschärfung des Picardschen Satzes	707
Schlesinger, L. , in Gießen. Parallelverschiebung und Krümmungstensor	413
Schönfinkel, M. , in Moskau, und P. Bernays in Göttingen. Zum Entscheidungsproblem der mathematischen Logik	342
Strutt, M. J. O. , in Eindhoven (Niederlande). Der Verlauf der Grenzkurven zwischen labilen und stabilen Lösungsgebieten der Mathieuschen Differentialgleichung	625
Suschkewitsch, A. , in Woronesch. Über die endlichen Gruppen ohne das Gesetz der eindeutigen Umkehrbarkeit	30
Synge, J. L. , in Dublin (Irland). Geodesics in non-holonomic geometry	738
Szegő, G. , in Königsberg. Über Funktionen mit positivem Realteil	142
Tietze, H. , in München. Über den Bereich absoluter Konvergenz von Potenzreihen mehrerer Veränderlichen. (Aus einem Schreiben an Herrn F. Hartogs)	181
Tietze, H. , in München. Eine charakteristische Eigenschaft der abgeschlossenen konvexen Punktmengen	394
Tonelli, L. , in Bologna. Su un problema di Abel	183
Tschebotareff, N. , in Odessa. Über die Realität von Nullstellen ganzer transzendenter Funktionen	660
van der Waerden, B. L. , in Göttingen. Eine Verallgemeinerung des Bézoutschen Theorems	497
Weatherburn, C. E. , in Christchurch (Neu-Seeland). On Families of Surfaces	473
Weitzenböck, R. , in Amsterdam. Berichtigung und Bemerkungen zur Arbeit „Über p -Relationen“	493
Zygmund, A. , in Warschau. Über die Beziehungen der Eindeutigkeitsfragen in den Theorien der geometrischen Reihen und Integrale	562
Berichtigung zu der Abhandlung von H. Brinkmeier Math. Annalen 96, S. 108—118	320

Idealtheorie in Quaternionenalgebren.

Von

H. Brandt in Aachen.

Einleitung.

1. Meine Untersuchungen über die Komposition der quaternären quadratischen Formen¹⁾ haben mich zu einer allgemeinen Idealtheorie der *rationalen Dedekindschen Algebren*²⁾ geführt, die, wie schon die ersten grundlegenden Sätze zeigen, an Schönheit der Idealtheorie algebraischer Zahlkörper kaum nachsteht. An Stelle des hier geltenden Satzes: „Alle Ideale bilden in bezug auf die Multiplikation eine unendliche Gruppe und lassen sich auf Grund einer Äquivalenzdefinition in Klassen einteilen, welche eine endliche Gruppe bilden“ tritt einfach ein allgemeiner, diesen als Spezialfall enthaltender Satz, der sich mit Hilfe des von mir durch Verallgemeinerung des Gruppenbegriffes gewonnenen Gruppoidbegriffes³⁾ sehr einfach formulieren läßt und so lautet: „*Alle Ideale bilden in bezug auf die Multiplikation ein unendliches Gruppoid und lassen sich auf Grund einer Äquivalenzdefinition in Klassen einteilen, die ein endliches Gruppoid bilden.*“

¹⁾ Hier kommen namentlich die folgenden Arbeiten in Betracht:

I. Zur Komposition der quaternären quadratischen Formen (Dissertation, Straßburg 1912), *Journal für reine und angewandte Mathematik* **143** (1913), S. 106.

II. Der Kompositionsbegriff bei den quaternären quadratischen Formen, *Math Annalen* **91** (1924), S. 300.

III. Die Hauptklassen in der Kompositionstheorie der quaternären quadratischen Formen, *Math. Annalen* **94** (1925), S. 166.

IV. Über die Komponierbarkeit quaternärer quadratischer Formen, *Math. Annalen* **94** (1925), S. 179.

²⁾ Dedekindsche Algebra = Algebra ohne Radikal, in Anlehnung an die Frobeniussche Bezeichnung Dedekindsche Gruppe. Vgl. Frobenius, *Theorie der hyperkomplexen Größen*, *Sitzungsberichte der Berliner Akademie* 1903, S. 509.

³⁾ Über eine Verallgemeinerung des Gruppenbegriffes, *Math. Annalen* **96** (1926), S. 360.

Wir geben hier den Beweis dieses Satzes für diejenigen aus Quaternionen gebildeten Algebren, welche der Kompositionstheorie der ganzzahligen quaternären quadratischen Formen in ähnlicher Weise entsprechen wie die algebraischen Zahlkörper einer Kompositionstheorie der zerlegbaren Formen mit ganzzahligen Koeffizienten⁴⁾. (Der allgemeine Beweis folgt in Kürze.)

2. Schon Hurwitz hat die einfachste derartige Algebra zahlentheoretisch behandelt⁵⁾, doch findet sich bei ihm noch nicht die geringste Andeutung für diesen Satz. Unsere Betrachtungen stehen daher auch mit den seinigen kaum in Zusammenhang. Sie sind in ihrem Gegenstand viel umfassender und verfolgen auch ganz andere Ziele. Das soll hier kurz an dem Hurwitzschen Beispiel dargelegt werden.

Die Hurwitzsche Algebra besteht aus allen Hamiltonschen Quaternionen mit rationalen Komponenten. Hurwitz bestimmt für diese Algebra einen größten Integritätsbereich, ohne nach den andern möglichen größten Integritätsbereichen zu fragen⁶⁾. Sind $\iota_0 = 1, \iota_1, \iota_2, \iota_3$ die Hamiltonschen Basiseinheiten der Quaternionen, so ist der Hurwitzsche Integritätsbereich durch die Basis $\frac{1}{2}(1 + \iota_1 + \iota_2 + \iota_3), \iota_1, \iota_2, \iota_3$ festgelegt.

Unterwirft man aber $\iota_1, \iota_2, \iota_3$ einer rationalen eigentlichen ternären orthogonalen Transformation, so kann man aus den transformierten Basiseinheiten $\iota'_1, \iota'_2, \iota'_3$ in derselben Weise einen Integritätsbereich konstruieren. Man sieht auch leicht, daß man auf diese Weise alle möglichen größten Integritätsbereiche der Algebra erhält. Scheinbar ist es nun unnötig, diese neuen Integritätsbereiche, die offenbar in unendlicher Anzahl vorhanden sind, neben dem Hurwitzschen Bereich in Betracht zu ziehen, weil die Produktrelationen zwischen den neuen Basiseinheiten vollständig mit den Produktrelationen zwischen den alten übereinstimmen. Deshalb können alle möglichen größten Integritätsbereiche einander isomorph zugeordnet werden, und es ist daher völlig gleichgültig, welchen von diesen Bereichen man betrachtet, sobald man nur einen für sich allein nimmt.

3. Man wird aber in ganz natürlicher Weise dazu geführt, diese Integritätsbereiche, die auch als Einheitsideale aufgefaßt werden können,

⁴⁾ Über die Resultate dieser Arbeit habe ich bereits am 30. August 1926 in Freiburg (Schweiz) auf der Tagung der Schweizer Mathematischen Gesellschaft vorgetragen. Vgl. Verhandlungen der Schweizerischen Naturforschenden Gesellschaft 1926, II. Teil, S. 155, oder *L'Enseignement Mathématique* 25 (1926), S. 290.

⁵⁾ A. Hurwitz, *Zahlentheorie der Quaternionen*, Nachrichten der Gesellschaft der Wissenschaften Göttingen 1896 und Vorlesungen über die Zahlentheorie der Quaternionen. Berlin 1919.

⁶⁾ Vgl. auch L. E. Dickson, *Algebras and their Arithmetics*, Chicago 1923, S. 148, oder zweite Auflage in deutscher Sprache, herausgegeben von A. Speiser, *Algebren und ihre Zahlentheorie*, Zürich 1927, S. 157.

sämtlich gleichzeitig nebeneinander zu betrachten. Das zeigen die folgenden Erörterungen, denen wir der kürzeren und klareren Ausdrucksweise wegen noch einige Bezeichnungen vorausschicken.

Ist e einer von diesen Integritätsbereichen und werden alle Quaternionen aus e durch $\varepsilon, \varepsilon_1, \dots$ bezeichnet, so nennen wir ein System a von Quaternionen α, α_1, \dots der Algebra ein *Rechtsideal* von e und ein System b von Quaternionen β, β_1, \dots der Algebra ein *Linksideal* von e , wenn a mit α, α_1, \dots auch $\alpha\varepsilon, \alpha_1\varepsilon_1, \alpha\varepsilon + \alpha_1\varepsilon_1, \dots$ und b mit β, β_1, \dots auch $\varepsilon\beta, \varepsilon_1\beta_1, \varepsilon\beta + \varepsilon_1\beta_1, \dots$ enthält⁷⁾. Wir schreiben dann $a e = a$ und $e b = b$ und bezeichnen e als *rechtes Einheitsideal* von a und als *linkes Einheitsideal* von b . Die Gesamtheit der Quaternionen $\alpha\beta, \alpha_1\beta_1, \alpha\beta + \alpha_1\beta_1, \dots$, wo also die Quaternionen des Rechtsideals a den ersten und die Quaternionen des Linksideals b den zweiten Faktor bilden, werde durch $a b$ bezeichnet und *Idealprodukt* von e genannt. Wegen $a = a e$ und $b = e b$ ist jedes Rechtsideal von e und jedes Linksideal von e gleichzeitig auch Idealprodukt von e .

4. Dann gelten folgende Sätze, deren Beweis wir hier übergehen (vgl. dazu Kap. IV): Jedes Rechtsideal a von e ist gleichzeitig auch Linksideal für einen im allgemeinen von e verschiedenen Integritätsbereich e' , so daß $e' a = a$, und jedes Linksideal b von e ist gleichzeitig auch Rechtsideal für einen im allgemeinen von e, e' verschiedenen Integritätsbereich e'' , so daß $b e'' = b$. Jedes Idealprodukt $a b$ von e ist gleichzeitig auch für jeden andern Integritätsbereich ein Idealprodukt und dabei gerade für einen selbst ein Rechtsideal und gerade für einen selbst ein Linksideal.

Die Menge der Rechtsideale für alle Integritätsbereiche, die Menge der Linksideale für alle Integritätsbereiche und die Menge der Idealprodukte für einen bestimmten Integritätsbereich sind daher miteinander identisch.

5. Da es danach nur eine Frage der Auffassung ist, ob ein System von Quaternionen als Rechtsideal oder als Linksideal oder als Idealprodukt angesehen wird, so scheint es angemessen, von Idealen schlechthin zu sprechen. Die Multiplikation der Ideale zeigt dann ganz gruppenähnliche Eigenschaften, ist aber an die einschränkende Bedingung geknüpft: Für zwei Ideale a, b existiert das Produkt $a b$ dann und nur dann, wenn das rechte Einheitsideal von a zugleich linkes Einheitsideal von b ist. Den präzisen Ausdruck für die Gesetzmäßigkeit in der multiplikativen Verknüpfung der Ideale gibt der Satz: Die Ideale bilden in bezug auf die Multiplikation ein unendliches Gruppoid.

⁷⁾ In derselben Weise werden die Rechts- und Linksideale auch von Herrn Speiser definiert [Allgemeine Zahlentheorie, Vierteljahrsschrift der Naturforschenden Gesellschaft in Zürich 71 (1926) = ⁶⁾ zweite Auflage, Kap. XIII], während bei Hurwitz die Bezeichnungen vertauscht sind.

6. Wir beschränken uns bei den folgenden Betrachtungen nicht auf die Hurwitzsche Algebra, sondern fassen alle (in unendlicher Anzahl vorhandenen) Quaternionenalgebren ins Auge, welche dadurch charakterisiert werden können, daß sie ein *rationales Zentrum*⁸⁾ besitzen. Auch hier gelten genau entsprechende Sätze (Kap. IV).

Die Untersuchung wird gekrönt durch das Ergebnis, daß man nach dem Vorbild der Idealtheorie in algebraischen Zahlkörpern eine Äquivalenzdefinition für unsere Ideale aufstellen kann, wodurch die sämtlichen Ideale sich auf Klassen verteilen, deren Anzahl sich als endlich erweist. Aus der Multiplikation der Ideale entsteht dann eine Komposition dieser Ideal-klassen. Dadurch werden diese in gesetzmäßiger Weise miteinander verknüpft und bilden dabei wiederum ein Gruppoid (Kap. V).

I. Grundbegriffe.

7. Wir bezeichnen gewöhnliche Größen (Zahlen) durch kleine lateinische Buchstaben. Da im folgenden algebraische Zahlen keine Verwendung finden, dürfen ganze rationale Zahlen kurz *ganze Zahlen* genannt werden. Quaternionen bezeichnen wir durch kleine griechische Buchstaben in der Form

$$\xi = x_0 + x_1 \iota_1 + x_2 \iota_2 + x_3 \iota_3.$$

Dabei sind $\iota_0 = 1, \iota_1, \iota_2, \iota_3$ die bekannten *Hamiltonschen Basiseinheiten* und x_0, x_1, x_2, x_3 die *Komponenten des Quaternionen* ξ .

Eine Komponente Null wird nicht hingeschrieben, die Quaternionen $\xi = x_0$, deren drei letzte Komponenten verschwinden, werden nicht von den Zahlen x_0 unterschieden.

8. Das Quaternion

$$\bar{\xi} = x_0 - x_1 \iota_1 - x_2 \iota_2 - x_3 \iota_3$$

wird nach Hurwitz *konjugiert* zu ξ genannt⁹⁾. Offenbar ist auch ξ konjugiert zu $\bar{\xi}$.

Die Ausdrücke

$$\begin{aligned} \xi + \bar{\xi} &= 2x_0 = s(\xi), \\ \xi \bar{\xi} &= x_0^2 + x_1^2 + x_2^2 + x_3^2 = E((x)) = n(\xi) \end{aligned}$$

werden als *Spur* und *Norm von* ξ bezeichnet. Das Quaternion ξ genügt der Gleichung

$$\xi^2 - s(\xi) \cdot \xi + n(\xi) = 0.$$

Wenn $n(\xi) \neq 0$, so wird das Quaternion $\frac{\bar{\xi}}{n(\xi)}$ durch ξ^{-1} bezeichnet und *reziprok zu* ξ genannt. Offenbar ist ξ reziprok zu ξ^{-1} .

⁹⁾ Siehe etwa ⁸⁾, erste Auflage, S. 31.

Wenn $\xi \neq 0$, aber $n(\xi) = 0$, so heißt ξ *Nullteiler*. Die Nullteiler ξ sind durch die Existenz nicht verschwindender Quaternionen η und ζ charakterisiert, für welche $\eta\xi = \xi\zeta = 0$. Dabei sind η, ζ natürlich selbst Nullteiler.

9. Ist neben $\xi = x_0 + x_1 \iota_1 + x_2 \iota_2 + x_3 \iota_3$ ein zweites Quaternion $\eta = y_0 + y_1 \iota_1 + y_2 \iota_2 + y_3 \iota_3$ gegeben, so wird der Ausdruck

$$\xi \bar{\eta} + \eta \bar{\xi} = \bar{\xi} \eta + \bar{\eta} \xi = 2(x_0 y_0 + x_1 y_1 + x_2 y_2 + x_3 y_3) = n(\xi, \eta)$$

als *Zwischennorm* von ξ und η bezeichnet. Man findet

$$n(\xi, \eta) = s(\xi \bar{\eta}) = s(\bar{\xi} \eta), \quad n(\xi, \xi) = 2n(\xi),$$

$$n(\xi, \xi \eta) = n(\xi, \eta \xi) = n(\xi) s(\eta).$$

Sind $\xi_v = \sum_k \iota_k x_{kv}$ ⁹⁾ irgend vier Quaternionen, so wird die aus den Zwischennormen gebildete Determinante

$$|n(\xi_i, \xi_k)| = 16 |x_{ik}|^2 = \Delta(\xi_0, \xi_1, \xi_2, \xi_3)$$

Diskriminante der vier Quaternionen ξ_v genannt. Δ ist dann und nur dann von Null verschieden, wenn die ξ_v linear unabhängig sind.

10. Die Multiplikationsgleichungen

$$\iota_1^2 = -1, \quad \iota_2 \iota_3 = -\iota_3 \iota_2 = \iota_1 \quad \text{usw.}$$

fassen wir in der Formel

$$\iota_\mu \iota_\nu = \sum_k \iota_k v_{k\mu\nu}$$

zusammen. Die Komponenten z_ν des Produktes

$$\xi \eta = \zeta$$

erhält man also durch die Gleichungen

$$z_\nu = \sum_{ik} v_{\nu ik} x_i y_k.$$

Für sie wird identisch

$$n(\xi) n(\eta) = n(\zeta) \quad \text{oder} \quad E((x)) E((y)) = E((z)).$$

11. Das *assoziative Gesetz* der Quaternionenmultiplikation drückt sich durch die Formeln

$$\sum_i v_{\lambda i \mu} v_{i \mu \nu} = \sum_i v_{\mu i \nu} v_{i \lambda \mu}$$

⁹⁾ Σ bedeutet eine Summation über die Indizes 0, 1, 2, 3. Summationsindizes werden durch i, j, k , feste Indizes durch $\alpha, \lambda, \mu, \nu$ bezeichnet.

aus, für die wir auch in einer leicht verständlichen, schon mehrfach benutzten Matrizensymbolik¹⁰⁾

$$V \begin{matrix} \diagup \\ \diagdown \end{matrix} O = V \begin{matrix} \diagup \\ \diagdown \end{matrix} V$$

schreiben. Bedeutet O eine eigentliche orthogonale Substitution mit der ersten Spalte $1, 0, 0, 0$, sonst aber ganz beliebigen Koeffizienten, so geht die bilineare Substitution V in sich selbst über, wenn die $x, y, z,$ gleichzeitig durch O transformiert werden, was wir durch die Formeln

$$O^{-1} \cdot V \begin{matrix} \diagup \\ \diagdown \end{matrix} O = V \quad \text{oder} \quad V \begin{matrix} \diagup \\ \diagdown \end{matrix} O = O \cdot V$$

zum Ausdruck bringen¹¹⁾.

12. Ein Quaternion wird *ganz* genannt, wenn seine Spur und seine Norm ganze Zahlen sind. Zwei ganze Quaternionen heißen *konkordant*, wenn ihre Zwischennorm ganz ist.

Wenn α und β ganze Quaternionen bezeichnen, so ist jedes der Quaternionen

$$\alpha + \beta, \alpha - \beta, \alpha \beta$$

dann und nur dann ganz, wenn α und β *konkordant* sind. Es ist nämlich

$$s(\alpha \pm \beta) = s(\alpha) \pm s(\beta), \quad n(\alpha \pm \beta) = n(\alpha) \pm n(\beta) \pm n(\alpha, \beta), \\ s(\alpha \beta) = s(\alpha) s(\beta) - n(\alpha, \beta), \quad n(\alpha \beta) = n(\alpha) n(\beta),$$

woraus die Behauptung folgt.

Weiter gilt: Wenn α und β ganze und konkordante Quaternionen bezeichnen, so sind alle Quaternionen, die aus den Quaternionen α, β oder auch aus $1, \alpha, \beta$ durch beliebig oft wiederholte Additionen, Subtraktionen und Multiplikationen hergeleitet werden können, ganz und konkordant mit α und β .

13. Ein System von unendlich vielen Quaternionen, in dem Addition, Subtraktion und Multiplikation unbeschränkt und die Division, soweit sie möglich ist, ausgeführt werden können, wird eine *Quaternionenalgebra* genannt. Enthält die Algebra keine Nullteiler, so heißt sie *Divisionsalgebra*¹²⁾.

¹⁰⁾ III, S. 168.

¹¹⁾ Die Richtigkeit der Behauptung ergibt sich leicht direkt durch Ausrechnen daraus, daß die zu O adjungierte Substitution O selbst ist. Vgl. auch IV, S. 192.

¹²⁾ Hurwitz sagt Quaternionenkörper, ich habe selbst diesen Ausdruck in dem unter *) genannten Vortrag gebraucht; doch scheint es mir jetzt zweckmäßiger, den Ausdruck Körper auf Divisionsalgebren mit kommutativer Multiplikation zu beschränken.

Wir betrachten im folgenden solche Quaternionenalgebren \mathfrak{Q} , die nur Quaternionen mit rationalen Normen und Spuren, jedoch vier linear unabhängige Quaternionen enthalten.

Offenbar sind in einer Algebra \mathfrak{Q} wegen

$$n(\alpha, \beta) = s(\alpha) s(\beta) - s(\alpha\beta)$$

auch alle Zwischennormen rational.

14. Sind $\alpha_0, \alpha_1, \alpha_2, \alpha_3$ irgend vier linear unabhängige Quaternionen aus \mathfrak{Q} und ist α ein beliebiges Quaternion, für welches die Zwischennormen $n(\alpha, \alpha_r)$ rational sind, so liegt auch α in \mathfrak{Q} .

Wegen der linearen Unabhängigkeit der α_r gibt es jedenfalls eindeutig bestimmte Zahlen k_r , so daß

$$\alpha = k_0 \alpha_0 + k_1 \alpha_1 + k_2 \alpha_2 + k_3 \alpha_3.$$

Hier sind aber die k_r rational; denn sie ergeben sich durch Auflösen der vier linearen Gleichungen

$$n(\alpha_0, \alpha_r) k_0 + n(\alpha_1, \alpha_r) k_1 + n(\alpha_2, \alpha_r) k_2 + n(\alpha_3, \alpha_r) k_3 = n(\alpha, \alpha_r),$$

deren Koeffizienten rational sind und deren Determinante

$$|n(\alpha_i, \alpha_k)| = \Delta(\alpha_0, \alpha_1, \alpha_2, \alpha_3)$$

nicht verschwindet.

Nimmt man α rational an, so zeigt sich, daß in \mathfrak{Q} alle rationalen Zahlen (im besonderen also auch die Zahl 1) vorkommen.

15. Nach dem Vorhergehenden bilden irgend vier linear unabhängige Quaternionen α_r aus \mathfrak{Q} eine *Basis*, d. h. \mathfrak{Q} enthält alle und nur die Quaternionen von der Form

$$\xi = x_0 \alpha_0 + x_1 \alpha_1 + x_2 \alpha_2 + x_3 \alpha_3,$$

wo die x_r rational sind. Die Quaternionen α_r heißen *Basiselemente*, die x_r *Koordinaten von ξ in bezug auf die Basis*.

Bildet man die Norm von ξ , wobei die x_r als Variable angesehen werden, so erhält man eine quadratische Form, die wir die *Normenform der Basis* nennen. Die Diskriminante dieser Form ist zugleich die Diskriminante der vier Basiselemente und heißt deshalb auch *Diskriminante der Basis*¹³⁾.

Die Koordinaten der Produkte der Basiselemente heißen die *Multiplikationszahlen der Basis*. Sie sind im allgemeinen gebrochen. Doch läßt sich eine Basis so wählen, daß sie ganz werden, z. B. dadurch, daß

¹³⁾ Die allgemeine von Dedekind stammende Definition der Diskriminante liefert den -16-fachen Wert von unserer Diskriminante. Siehe etwa Speiser a. a. O. 7), S. 20.

man die Basiselemente mit dem Generalnenner der Multiplikationszahlen multipliziert.

II. Minimalbasen.

16. Eine Basis heißt *Minimalbasis*, wenn ihre Multiplikationszahlen ganz sind und ihre Diskriminante möglichst klein ist.

Da aus der Ganzzahligkeit der Multiplikationszahlen, wie sich unten zeigen wird, auch die Ganzzahligkeit der Diskriminante folgt (20), so ist klar, daß es unter den Basen mit ganzen Multiplikationszahlen Minimalbasen geben muß.

Offenbar erhält man durch ganzzahlige unimodulare Transformation aus einer Minimalbasis ω_ν wieder eine Minimalbasis. Da jede Algebra \mathfrak{Q} alle rationalen Zahlen enthält (14), läßt sich die neue Basis ω'_ν so wählen, daß $\omega'_0 = r$ eine positive rationale Zahl wird. Wegen der Ganzzahligkeit der Multiplikationszahlen der ω'_ν muß r ganz sein, wegen der Minimalbedingung für die Diskriminante kann dann r aber nur den Wert 1 haben.

Eine Minimalbasis ω_ν , bei der $\omega_0 = 1$ ist, wird *reduziert* genannt.

17. Für eine beliebige Minimalbasis ω_ν setzen wir

$$\omega_\nu = \sum_k t_k t_{k\nu},$$

$$\omega = \omega(x) = x_0 \omega_0 + x_1 \omega_1 + x_2 \omega_2 + x_3 \omega_3,$$

und bezeichnen die Normenform durch

$$n(\omega) = \sum_j \sum_{ik} t_{ji} t_{jk} x_i x_k = \frac{1}{2} \sum_{ik} g_{ik} x_i x_k = G((x)).$$

Die Koordinaten der Zahl 1, die nach dem vorigen ganze teilerfremde Zahlen sind, bezeichnen wir durch e_ν , so daß

$$1 = e_0 \omega_0 + e_1 \omega_1 + e_2 \omega_2 + e_3 \omega_3.$$

Für die Spur von ω gilt dann

$$s(\omega) = n(1, \omega) = \sum_{ik} g_{ik} e_i x_k.$$

18. Die Multiplikationszahlen seien $w_{\lambda\mu\nu}$, so daß

$$\omega_\mu \omega_\nu = \sum_i \omega_i w_{i\mu\nu}.$$

Die drei aus den $w_{\lambda\mu\nu}$ gebildeten Matrizen aus Linearformen bezeichnen wir durch

$$\left\| \sum_j w_{ijk} x_j \right\| = \Omega(x), \quad \left\| \sum_j w_{ikj} y_j \right\| = \Omega'(y), \quad \left\| \sum_j w_{jik} z_j \right\| = \Omega''(z).$$

Durch die Substitution $T = \|t_{ik}\|$, wobei die $t_{\mu\nu}$ die eben eingeführten

Größen sind, stellt sich die bilineare Matrix W bei Benutzung der in 11 eingeführten Symbolik durch die Formel

$$W = T^{-1} \cdot V \begin{matrix} \left\langle T \\ T \right\rangle \end{matrix}$$

dar.

19. Deshalb wird vermöge der bilinearen Substitution

$$z_\nu = \sum_{ik} w_{\nu ik} x_i y_k$$

identisch

$$G((x)) G((y)) = G((z))$$

und

$$\begin{aligned} \left| \frac{\partial z_i}{\partial y_k} \right| &= |\Omega(x)| = [G((x))]^2, \\ \left| \frac{\partial z_i}{\partial x_k} \right| &= |\Omega'(y)| = [G((y))]^2, \\ |\Omega''(z)| &= -[\mathfrak{G}((z))]^2, \end{aligned}$$

wobei \mathfrak{G} eine Form bezeichnet, die wir früher die *Reziproke von G* genannt haben¹⁴⁾. Schreibt man

$$\mathfrak{G}((z)) = \frac{1}{2} \sum_{ik} g_{ik} z_i z_k$$

und

$$|g_{ik}| = |g_{ik}| = d^2,$$

so ist

$$g_{\mu\nu} = \frac{1}{d} \frac{\partial |g_{ik}|}{\partial g_{\mu\nu}}.$$

20. Weil die $w_{\lambda\mu\nu}$ ganz sind, müssen die Koeffizienten von G und \mathfrak{G} ganz sein. Demnach ist auch d ganz. Wir wollen d mit positivem oder negativem Vorzeichen nehmen, je nachdem die Normenform eine definite oder indefinite Form ist¹⁵⁾. Im ersten Fall kann dann G nur positiv sein, im zweiten Fall nur den Trägheitsindex 2 haben¹⁶⁾. Die in dieser Weise bestimmte ganze Zahl d ist wegen der Definition der Minimalbasis unabhängig von der Auswahl der Minimalbasis. Sie wird die *Grundzahl der Algebra* genannt.

Die Diskriminante einer Minimalbasis, die wir auch *Diskriminante der Algebra* nennen und durch Δ bezeichnen, ist das Quadrat der Grundzahl.

Die Grundzahl kann nicht jeden beliebigen Wert haben, wie weiter unten näher angegeben wird (25).

Da die Normenform einer Minimalbasis die Zahl 1 darstellt, ist sie nach früherer Terminologie eine *Hauptform*, welche eine Komposition mit sich selbst zu sich selbst gestattet¹⁷⁾.

¹⁴⁾ I, S. 110 oder II, S. 302.

¹⁵⁾ Vgl. II, S. 302.

¹⁶⁾ I, S. 118.

¹⁷⁾ I, S. 122 und II, S. 309.

21. Wenn $\omega_0 = 1$, die Minimalbasis also reduziert ist, hat die Normenform den ersten Koeffizienten $\frac{1}{2}g_{00} = 1$. Dann wird

$$s(\omega) = 2x_0 + g_{01}x_1 + g_{02}x_2 + g_{03}x_3,$$

deshalb sind die Koordinaten \bar{x}_ν von $\bar{\omega}$

$$\bar{x}_0 = x_0 + g_{01}x_1 + g_{02}x_2 + g_{03}x_3, \quad \bar{x}_1 = -x_1, \quad \bar{x}_2 = -x_2, \quad \bar{x}_3 = -x_3.$$

Sie werden also aus den Koordinaten x_ν von ω durch die Transformation

$$E_* = \begin{vmatrix} 1 & g_{01} & g_{02} & g_{03} \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{vmatrix}$$

erhalten. Wird

$$E_0 = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{vmatrix}$$

gesetzt, so wird

$$E_* = T^{-1} \cdot E_0 \cdot T,$$

wobei $T = \|t_{ik}\|$ die frühere Bedeutung hat (17), aber für eine reduzierte Basis die erste Spalte 1, 0, 0, 0 besitzt.

22. Im Fall der reduzierten Basis sind die Multiplikationszahlen merkwürdigerweise durch die Koeffizienten der Normenform und ihrer Reziproken schon vollständig bestimmt. Man findet nämlich nach Ausführung der Transformation unter 18 die von den $t_{\mu\nu}$ ganz unabhängigen, schon früher an anderer Stelle¹⁸⁾ angegebenen Formeln

$$z_0 = x_0 y_0 - \frac{1}{2} \sum'_{ik} g_{ik} x_i y_k \pm \frac{1}{2} \sum'_i g_{0i} s_i,$$

$$z_\tau = \frac{1}{2} \sum'_i g_{0i} (x_\tau y_i + x_i y_\tau) \pm \frac{1}{2} \sum'_i g_{\tau i} s_i \quad (\tau = 1, 2, 3).$$

Dabei ist zur Abkürzung gesetzt

$$x_2 y_3 - x_3 y_2 = s_1, \quad x_3 y_1 - x_1 y_3 = s_2, \quad x_1 y_2 - x_2 y_1 = s_3,$$

und Σ' bezeichnet eine Summation über die Indizes 1, 2, 3. Hier gelten entweder die oberen oder die unteren Zeichen, und die Zeichen vertauschen sich, wenn die drei Variablenreihen x_ν, y_ν, z_ν gleichzeitig durch die Substitution E_* transformiert werden.

¹⁸⁾ II, S. 308 oder 309.

Die Koeffizienten dieser bilinearen Substitution sind, wie früher gezeigt wurde, stets ganzzahlig, sobald $G = \frac{1}{2} \sum_{i,k} g_{ik} x_i x_k$ eine Form mit einer ganzzahligen Reziproken $\mathfrak{G} = \frac{1}{2} \sum_{i,k} g_{ik} x_i x_k$ und mit dem ersten Koeffizienten $\frac{1}{2} g_{00} = 1$ bezeichnet¹⁹⁾.

23. Die Formen, welche als Normenformen einer beliebigen Basis auftreten können, sind dadurch ausgezeichnet, daß sie sich rational ineinander transformieren lassen und eine rationale Reziproke, also quadratische Diskriminante besitzen. Dabei kann auch jede Form, welche rational in eine von diesen Formen transformiert werden kann, selbst als Normenform auftreten.

Es muß daher unter den möglichen Normenformen auch Stammformen, d. h. solche Formen geben, welche nicht unter einer ganzzahligen Form mit kleinerer Diskriminante enthalten sind. Wie eine einfache Diskussion zeigt, sind solche Formen von zweiter Art²⁰⁾ und besitzen eine ganzzahlige Reziproke. Außerdem gehören sie einem bestimmten Geschlecht an, das wir *Stammgeschlecht* nennen können, da es nur Stammformen enthält. Entsteht aus einer Stammform durch rationale unimodulare Transformation eine ganzzahlige Form, so ist diese auch Stammform.

Hieraus ergeben sich in Verbindung mit der am Schlusse der vorigen Nummer gemachten Bemerkung die folgenden Sätze:

Eine Basis ist dann und nur dann eine Minimalbasis, wenn die Normenform eine Stammform ist und die Koordinaten der Zahl 1 ganz sind.

Aus einer Minimalbasis erhält man alle Minimalbasen, wenn man alle rationalen Transformationen der Determinante ± 1 ausübt, bei denen die Normenform und die Koordinaten der Zahl 1 wieder ganzzahlig werden.

24. Eine Gesamtheit von ganzen Quaternionen, die mit α , β auch $\alpha + \beta$, $\alpha - \beta$ und $\alpha\beta$ enthält, wird *Integritätsbereich* genannt.

Ein Integritätsbereich heißt *größter oder maximaler Integritätsbereich*, wenn er nicht durch Aufnahme weiterer Quaternionen vergrößert werden kann, d. h. mit andern Worten (12), wenn jedes ganze Quaternion, das mit allen Quaternionen des Bereiches konkordant ist, schon in dem Bereich enthalten ist.

Dann gilt der Satz:

Eine Basis ist dann und nur dann eine Minimalbasis, wenn alle Quaternionen mit ganzen Koordinaten einen größten Integritätsbereich bilden.

¹⁹⁾ II, S. 303.

²⁰⁾ D. h. sie sind halb genommene uneigentlich primitive Formen.

Hat eine Basis nämlich diese Eigenschaft, so sind die Multiplikationszahlen ganz. Außerdem sind die Koordinaten der Zahl 1 ganz und die Normenform ist eine Stammform; denn sonst könnte man den Integritätsbereich noch vergrößern. Es liegt also eine Minimalbasis vor.

Umgekehrt ist klar, daß alle Quaternionen, die in bezug auf eine Minimalbasis ganze Koordinaten haben, einen Integritätsbereich bilden. Gäbe es nun ein ganzes mit den Basiselementen konkordantes Quaternion, dessen Koordinaten den Generalnenner r hätten, so könnte man die Basis transformieren durch eine ganzzahlige unimodulare Substitution $\|u_{ik}\|$, für welche die Elemente der ersten Spalte diesen Koordinaten proportional wären. Dann würden aber bei der Normenform G' der neuen Basis die Koeffizienten g'_{00} durch r^2 und $g'_{01}, g'_{02}, g'_{03}$ durch r teilbar werden, was für eine Stammform unmöglich ist.

25. Diskutiert man die möglichen Fälle, für welche eine quaternäre quadratische Form eine quadratische Diskriminante haben und zugleich Stammform sein kann, so ergibt sich, daß die Grundzahl d jeden positiven oder negativen quadratfreien Wert, in dem für $d > 0$ eine ungerade und für $d < 0$ eine gerade Anzahl Primfaktoren aufgehen, jedoch keinen andern Wert haben kann.

Der Fall $d = 2$ entspricht der Hurwitzschen Algebra (siehe Einleitung). Der Fall $d = -1$ (Primzahlanzahl Null) nimmt eine Ausnahmestellung ein, er umfaßt diejenigen Algebren, welche Nullteiler enthalten. Alle andern Algebren sind Divisionsalgebren. Das ergibt sich daraus, daß jede quaternäre quadratische Form mit quadratischer Diskriminante, welche die Null darstellt, unter der Form $x_0 x_3 - x_1 x_2$ enthalten ist.

26. Es fragt sich noch, wie die verschiedenen Algebren, welche dieselbe Grundzahl haben, miteinander zusammenhängen. Da es nur ein Stammgeschlecht der Diskriminante d^2 gibt und Formen eines Geschlechts sich rational ineinander transformieren lassen, können in zwei verschiedenen Algebren der Grundzahl d zwei Minimalbasen so angenommen werden, daß zunächst die Normenformen und wegen der in 22 gemachten Bemerkungen auch die Multiplikationszahlen dieselben sind. Dadurch sind aber beide Algebren isomorph aufeinander bezogen.

Es läßt sich leicht zeigen, daß es zu derselben Grundzahl d unendlich viele Algebren gibt, und ihre Menge ist nicht einmal abzählbar. Die Substitution T der Nummer 18 kann nämlich nach 11, ohne daß W sich ändert, durch OT ersetzt werden, wo O eine ganz beliebige eigentliche orthogonale Substitution mit der ersten Spalte 1, 0, 0, 0 bezeichnet. Offenbar kann man aber in die Koeffizienten von O noch ganz beliebige Irrationalitäten aufnehmen.

Faßt man alle die unendlich vielen Algebren zu derselben Grundzahl als verschiedene Darstellungen einer einzigen abstrakten Algebra auf, so bestimmt jede Grundzahl eine einzige Algebra.

III. Moduln.

27. Zum näheren Studium der Eigenschaften unserer Algebren werde jetzt in einer Algebra \mathfrak{Q} eine Minimalbasis ω_ν mit den Multiplikationszahlen $w_{\lambda,\mu,\nu}$ beliebig aber fest ausgewählt und als *Grundbasis* bezeichnet, die Normenform heie *Grundform*. Zur bequemeren Formulierung der Stze darf ohne Einschrnkung der Allgemeinheit die Basis als *reduziert*, somit $\omega_0 = 1$ angenommen werden.

Ein System von Quaternionen aus \mathfrak{Q} , die entweder smtlich ganz sind oder doch nach Multiplikation mit einer festen Zahl smtlich ganz werden, wird *Modul* genannt, wenn das System vier linear unabhngige Quaternionen enthlt und mit α, β auch $\alpha + \beta$ und $\alpha - \beta$ darin vorkommen.

Ein Modul heit *ganzer Modul*, wenn er nur ganze Quaternionen enthlt.

28. *Jeder Modul besitzt eine Basis.*

Zum Beweise gengt es, nur die ganzen Moduln ins Auge zu fassen, weil die andern darauf zurckgefhrt werden knnen. In einem ganzen Modul \mathfrak{a} sind aber auch die Zwischennormen und somit die Diskriminanten fr vier linear unabhngige Quaternionen ganz (12). Whlt man nun aus \mathfrak{a} vier linear unabhngige Quaternionen α_ν aus, fr welche die Diskriminante einen Minimalwert annimmt, so enthlt \mathfrak{a} alle aber auch nur die Quaternionen, welche linear mit ganzen Koeffizienten durch die α_ν darstellbar sind. Andernfalls knnte man nmlich, wie man leicht erkennt, vier Quaternionen mit noch kleinerer Diskriminante finden.

29. Ist \mathfrak{a} ein beliebiger Modul und bilden die Quaternionen

$$\alpha_\nu = \sum_i \omega_i a_{i\nu},$$

eine Basis, so wird die Matrix $A = \|a_{i,k}\|$ *Basismatrix des Moduls* genannt. Die Reihenfolge der Basiselemente werde immer so gewhlt, da die Determinante dieser Matrix, welche wir auch *Determinante des Moduls* nennen und durch $|\alpha_{i,k}| = |\mathfrak{a}|$ bezeichnen, positiv wird.

Die Norm der Linearform

$$\alpha(x) = x_0 \alpha_0 + x_1 \alpha_1 + x_2 \alpha_2 + x_3 \alpha_3$$

ist eine quaternre quadratische Form mit im allgemeinen gebrochenen Koeffizienten. Sie kann aber stets in der Gestalt

$$n(\alpha(x)) = a F((x))$$

geschrieben werden, wo F eine ganzzahlige primitive Form bezeichnet und a positiv ist. Wir nennen dann den Koeffizienten a die *Norm des Moduls* und schreiben $a = n(a)$. Die Form F aber werde die zu der Basis gehörige *Normenform des Moduls* genannt.

30. Offenbar sind Norm und Determinante eines Moduls gänzlich unabhängig sowohl von der Basis des Moduls wie von der Grundbasis der Algebra, dagegen ist für die Normenform nur die Klasse bestimmt. (Würde man für die Determinante eines Moduls auch negative Werte zulassen, so würden die Normenformen zwei im allgemeinen verschiedenen Klassen angehören.)

Weil die Grundform G durch die lineare Substitution A in aF übergeht und G Stammform ist, so gilt immer

$$a^2 \leq |a|, \text{ d. h. in Worten:}$$

Für jeden Modul ist das Quadrat der Norm kleiner oder höchstens so groß wie die Determinante.

31. Ersetzt man alle Quaternionen des Moduls a durch die konjugierten Quaternionen, so erhält man wieder einen Modul (21), den wir durch \bar{a} bezeichnen und den *konjugierten Modul* nennen.

Bilden die Quaternionen α_r eine Basis von a , so ist klar, daß durch die konjugierten Quaternionen $\bar{\alpha}_r$ jedes Quaternion von \bar{a} linear mit ganzzahligen Koeffizienten dargestellt werden kann. Die $\bar{\alpha}_r$ liefern daher eine Basis, wenn sie (wegen der Forderung der positiven Determinante) noch durch eine beliebige ganzzahlige Substitution mit der Determinante -1 z. B. durch E_0 transformiert werden. Daher gilt:

Aus der Basismatrix A von a erhält man durch die Formel

$$\bar{A} = E_* A E_0$$

eine Basismatrix von \bar{a} .

32. Wenn man alle Quaternionen des Moduls a links mit ϱ und rechts mit σ multipliziert, wobei ϱ, σ irgend zwei Quaternionen aus \mathfrak{D} mit positivem Normenprodukt bezeichnen, so erhält man wieder einen Modul, den wir durch $a' = \varrho a \sigma$ bezeichnen. Wenn die Quaternionen α_r eine Basis von a bilden, so bilden die Quaternionen $\alpha'_r = \varrho \alpha_r \sigma$ offenbar eine Basis von a' . Werden die Koordinaten von ϱ durch r_r und von σ durch s_r bezeichnet, so ergibt eine einfache Rechnung für die Basismatrix dieser Basis von a'

$$A' = \Omega(r) \Omega'(s) A.$$

Deshalb ist

$$n(a') = n(\varrho) n(\sigma) n(a), \quad |a'| = [n(\varrho)]^2 [n(\sigma)]^2 |a|,$$

und die Normenformen der Moduln a, a' gehören derselben Formenklasse an.

33. Sind zwei Moduln α und β gegeben und durchläuft α alle Quaternionen von α , β alle Quaternionen von β , so bilden die Produkte $\alpha\beta$ und ihre Summen wieder einen Modul, den wir durch

$$c = \alpha \times \beta$$

bezeichnen und das *Produkt der Moduln α und β* nennen. Werden nämlich die $\alpha\beta$ und ihre Summen durch γ bezeichnet, so ist klar, daß unter den γ vier linear unabhängige Quaternionen und mit γ_1, γ_2 auch $\gamma_1 + \gamma_2$ und $\gamma_1 - \gamma_2$ vorkommen, die γ bilden daher einen Modul, wenn wir noch zeigen können, daß sie nach Multiplikation mit einer festen ganzen Zahl sämtlich ganz werden. Ein solcher Multiplikator wird aber erhalten, wenn man den Generalnenner der Koeffizienten einer Basismatrix von α mit dem Generalnenner der Koeffizienten einer Basismatrix von β multipliziert. Sind nämlich die Basismatrizen von α und β ganzzahlig, so sind alle Quaternionen der beiden Moduln α, β ganz und untereinander konkordant, daher sind auch alle γ ganz (12).

Für die Multiplikation der Moduln gilt das assoziative, aber nicht das kommutative Gesetz.

Das folgt aus den Multiplikationsgesetzen der Quaternionen.

34. Sind A, B Basismatrizen der Moduln α, β , so genügt jede Basismatrix C des Produktes $c = \alpha \times \beta$ den Bedingungen:

Die bilineare Substitution

$$C^{-1} \cdot W \begin{matrix} A \\ B \end{matrix} = M$$

ist ganzzahlig, und die Determinanten der rechteckigen Matrix

$$\|m_{\lambda\mu\nu}\| \quad (\lambda = 0, 1, 2, 3; \mu\nu = 00, 01, \dots, 33)$$

haben keinen gemeinsamen Teiler.

Sind umgekehrt für eine Matrix C mit nicht verschwindender positiver Determinante diese beiden Bedingungen erfüllt, so ist C Basismatrix des Produktes $c = \alpha \times \beta$.

35. Werden die Normen von α, β, c durch a, b, c und die durch A, B, C bestimmten Normenformen durch F_a, F_b, F_c bezeichnet, so transformiert die bilineare Substitution M die Form cF_c in das Produkt $aF_a \cdot bF_b$, d. h. F_c ganzzahlig in $\frac{ab}{c} F_a F_b$.

Hier ist $\frac{ab}{c}$ ganz.

Erteilt man nämlich den Variablen von F_a und F_b ganzzahlige Werte, so müssen wegen der Ganzzahligkeit von M die Variablen von F_c und somit auch der Wert von F_c ganzzahlig ausfallen. Da aber F_a und F_b

primitive Formen sind, so können ihre Werte zu jeder vorgegebenen Zahl prim gemacht werden, also auch zu dem etwaigen Nenner von $\frac{ab}{c}$. Dann könnte aber F_c nicht ganzzahlig sein.

Demnach gilt der Satz:

*Das Produkt der Normen zweier Moduln ist stets durch die Norm ihres Produktes teilbar, d. h. $n(a)n(b)$ ist teilbar durch $n(a \times b)$.*²¹⁾

36. Bildet man endlich die beiden Determinanten

$$\left| \sum_j m_{ijk} x_j \right|, \quad \left| \sum_j m_{ikj} y_j \right|,$$

so haben diese wegen der Matrizengleichung aus 34 die Werte

$$a^2 \frac{|b|}{|c|} F_a^2, \quad b^2 \frac{|a|}{|c|} F_b^2.$$

Also gilt:

Wenn $a \times b = c$, so sind $a^2 |b|$ und $b^2 |a|$ beide durch $|c|$ teilbar.

IV. Ideale.

37. Wir stellen jetzt folgende Definition auf:

*Ein Modul wird Ideal genannt, wenn seine Determinante das Quadrat seiner Norm ist*²²⁾.

Da jedes Ideal ein Modul ist, können die Begriffe der Basis, der Determinante, der Norm, der Normenform ohne weiteres auf Ideale übertragen werden.

Der zu einem Ideal a konjugierte Modul \bar{a} ist offenbar auch ein Ideal und heißt *konjugiertes Ideal*. Dividiert man die sämtlichen Quaternionen des konjugierten Ideals \bar{a} durch die Norm a , so entsteht ein Ideal $\frac{1}{a} \bar{a} = a^{-1}$ mit der Norm $\frac{1}{a}$, welches *reziprok zu a* heißen soll. Offenbar ist a zu a^{-1} reziprok.

Ein Ideal heißt *ganz*, wenn es nur aus ganzen Quaternionen besteht; ein ganzes Ideal a heißt *primitiv*, wenn es keine ganze Zahl $t > 1$ gibt, so daß $\frac{1}{t} a$ ganz ist.

²¹⁾ Dabei heißt eine rationale Zahl a durch eine rationale Zahl b teilbar, wenn der Quotient $\frac{a}{b}$ eine ganze Zahl ist.

²²⁾ Diese Definition mag vielleicht auf den ersten Blick befremden, doch werden sich bald weitere charakteristische Eigenschaften unserer Ideale ergeben, die einen Zusammenhang mit den üblichen Definitionen eines Ideals in einem algebraischen Zahlkörper herstellen. (Übrigens wird auch in einem algebraischen Zahlkörper ein Modul durch eine ähnliche Forderung, nämlich die Gleichheit von Determinante und Norm als Ideal charakterisiert.)

Die sämtlichen Quaternionen, die in bezug auf die Grundbasis ganze Koordinaten haben, bilden ein Ideal von der Norm 1, das wir *Grundideal* nennen und durch \mathfrak{o} bezeichnen.

38. Ist \mathfrak{a} ein beliebiges Ideal mit der Norm a , bilden die Quaternionen $\alpha_\nu = \sum_i \omega_i a_{i\nu}$ eine Basis von \mathfrak{a} und ist F die zugehörige Normenform, so ist nach einer früheren Terminologie die Basismatrix A eine zur Grundform G gehörige, die Form F erzeugende Substitution mit der Norm a .²³⁾ Diese Erzeugende ist zudem eigentlich, weil $|a| > 0$. Umgekehrt ist auch jede eigentliche Erzeugende zur Grundform G Basismatrix eines Ideals. Der einzige Unterschied gegen früher besteht darin, daß wir jetzt auch gebrochene Werte für die Koeffizienten der Matrix A zulassen, während wir uns früher auf ganzzahlige Werte beschränken konnten.

Wegen der Bedeutung der *Erzeugenden von positivem und negativem Charakter* vgl. man weiter unten Nr. 55.

39. In jedem Ideal \mathfrak{a} gibt es Quaternionen α , für welche der Quotient $n(\alpha)/n(\mathfrak{a})$ unter einer festen, nur von der Grundzahl d der Algebra abhängigen Schranke liegt.

Stellt man nämlich die Quaternionen des Ideals \mathfrak{a} durch eine Basis dar, so zeigt sich, daß dieser Quotient eine durch die Normenform F von \mathfrak{a} darstellbare Zahl ist. Bekanntlich stellt aber eine quadratische Form Zahlen unterhalb einer endlichen Schranke dar, für welche man den Wert $2|\sqrt{d}|$ nehmen kann, wo d die Grundzahl der Algebra bezeichnet.

Da eine primitive quadratische Form stets Zahlen darstellt, die zu einer beliebig vorgegebenen Zahl prim sind, so gilt weiter:

In jedem Ideal \mathfrak{a} gibt es Quaternionen α , für welche der Quotient $n(\alpha)/n(\mathfrak{a})$ zu einer beliebig vorgegebenen Zahl prim ist.

40. Ein Ideal von der Norm 1, das die Zahl 1 enthält, wird *Einheitsideal* genannt.

Ein Einheitsideal \mathfrak{e} enthält nur ganze Quaternionen; denn wenn ε in \mathfrak{e} liegt, so sind $n(\varepsilon)$ und $s(\varepsilon) = n(1, \varepsilon)$ ganz (29). Deshalb ist auch $\bar{\varepsilon} = s(\varepsilon) - \varepsilon$ in \mathfrak{e} enthalten. Ein Einheitsideal ist also dem konjugierten und, weil die Norm gleich 1 ist, auch dem reziproken Ideal gleich.

Jede Basis eines Einheitsideals ist zugleich Minimalbasis der Algebra, und umgekehrt ist jede Minimalbasis zugleich Basis eines Einheitsideals (weshalb man den Begriff der reduzierten Basis (16) auf die Einheitsideale anwenden kann); im besonderen ist also das Grundideal \mathfrak{o} ein Einheitsideal.

²³⁾ IV, S. 180.

Die Einheitsideale sind daher mit den größten Integritätsbereichen der Algebra identisch.

41. Wir betrachten im besonderen diejenigen Einheitsideale, deren Normenformen der Grundform äquivalent sind. Für ein solches Ideal e kann man eine reduzierte Basis so auswählen, daß die Normenform der Grundform gleich wird. Die Basismatrix U ist dann eine eigentliche automorphe Substitution der Grundform mit der ersten Spalte $1, 0, 0, 0$.

Aus der bekannten Eulerschen Darstellung der ternären orthogonalen Substitutionen²⁴⁾ folgt dann für U die Darstellung

$$U = \Omega(r)^{-1} \Omega'(r).$$

(Diese Formel geht nämlich in die Eulersche über, wenn $W = V$ (10, 11), und folgt aus dieser durch lineare Transformation (18).)

Schreibt man sie in der Gestalt $\Omega(r)U = \Omega'(r)$, so erkennt man, daß die r_i als Auflösung linearer homogener Gleichungen mit rationalen Koeffizienten selbst rational, daher auch ganz angenommen werden können. Umgekehrt ist klar, daß jedes U von dieser Form mit ganzen r_i Basismatrix eines Einheitsideals ist. Also gilt:

Die Formel

$$U = \Omega(r)^{-1} \Omega'(r),$$

wobei die r_i ganz sind, liefert als Basismatrix alle und nur die Einheitsideale, deren Normenformen der Grundform äquivalent sind.

Wählt man die r_i im besonderen so, daß $G((r)) = m$ eine zu 2d prime Zahl ist, so hat U den Generalnenner m .²⁵⁾ Da m unendlich viele verschiedene Werte haben kann und verschiedenen Werten von m natürlich auch verschiedene Einheitsideale entsprechen, so gilt weiter:

Es gibt unendlich viele verschiedene Einheitsideale.

42. Sind a und b zwei Ideale, für welche das Modulprodukt $c = a \times b$ der Bedingung $n(c) = n(a)n(b)$ genügt, so ist c ebenfalls ein Ideal.

Daß die Annahmen dieses Satzes möglich sind, ergibt sich schon aus früheren Betrachtungen über die Komponierbarkeit der quaternären quadratischen Formen²⁶⁾.

Zum Beweise beachte man, daß wegen 30 die Determinante des Moduls $c = a \times b$ der Bedingung

$$|c| \geq [n(c)]^2$$

²⁴⁾ L. Euler, Opera omnia, I. Serie, 6, S. 309.

²⁵⁾ IV, S. 186 ff.

²⁶⁾ IV, S. 195.

und wegen 36 der Bedingung

$$|c| \leq [n(a)n(b)]^2$$

genügt, woraus wegen $n(a)n(b) = n(c)$

$$|c| = [n(c)]^2$$

folgt, d. h. der Modul c ist ein Ideal.

Wir definieren daher:

Wenn für zwei Ideale a, b das Modulprodukt $c = a \times b$ der Bedingung $n(c) = n(a)n(b)$ genügt, so daß c wieder ein Ideal ist, so schreiben wir $c = ab$ und nennen c Idealprodukt oder, wenn keine Verwechslung mit dem allgemeinen Modulprodukt möglich ist, kurz Produkt aus a und b .

43. Für die Multiplikation der Ideale gilt das *assoziative Gesetz* in folgendem Sinne:

Wenn für eine Anzahl in fester Reihenfolge gegebene Ideale für eine bestimmte Art der Klammersetzung ein Idealprodukt existiert, so existiert dasselbe Idealprodukt auch für jede andere Art der Klammersetzung, sobald nur die Reihenfolge erhalten bleibt.

Das folgt aus der Gültigkeit des assoziativen Gesetzes für die Modulmultiplikation. Wendet man dies Gesetz auf Ideale an und ergibt sich bei einer Anordnung der Klammern, daß die sämtlichen auftretenden Modulprodukte Idealprodukte sind, so muß das auch für jede andere Anordnung der Klammern der Fall sein, weil sonst das Endprodukt nicht der Idealbedingung genügen könnte.

Demnach können die Klammern bei Idealprodukten ganz entbehrt werden²⁷⁾.

44. Gilt für die Ideale a, b, c die Produktgleichung

$$ab = c$$

und sind A, B, C Basismatrizen dieser Ideale, so ist die bilineare Substitution

$$C^{-1} \cdot W \begin{matrix} \langle A \\ B \end{matrix} = M$$

wie bei der Modulmultiplikation ganzzahlig. Sie vermittelt gleichzeitig die Komposition der Normenformen. Hält man zwei der drei Matrizen A, B, C fest, so kann, wenn M ganzzahlig und $|A||B| = |C|$ bleiben soll, die dritte immer nur durch rechts äquivalente ersetzt werden, d. h. durch AU, BU, CU , wo U eine ganzzahlige Matrix der Determinante 1 bezeichnet.

²⁷⁾ Vgl. hierzu Brandt, Über das assoziative Gesetz bei der Komposition der quaternären quadratischen Formen, *Math. Annalen* 96 (1926), S. 353.

Das folgt daraus, daß die Determinanten der drei rechteckigen Matrizen mit den Elementen $m_{\lambda\mu\nu}$, $m_{\nu\lambda\mu}$, $m_{\mu\nu\lambda}$ ($\lambda = 0, 1, 2, 3$; $\mu\nu = 00, 01, \dots, 33$) wegen der Primitivität der Normenformen von a, b, c keinen gemeinsamen Teiler haben können.

Das besagt für die Ideale:

Wenn das Produkt $ab = c$ existiert, so ist jedes der drei Ideale a, b, c durch die beiden andern eindeutig bestimmt.

45. Weil ein Einheitsideal e ein Integritätsbereich ist, so gilt $ee = e$. Zwei verschiedene Einheitsideale können aber nicht multipliziert werden. Existierte nämlich für zwei verschiedene Einheitsideale e_1, e_2 das Idealprodukt $e_1 e_2 = e_3$, so müßte e_3 Einheitsideal und somit auch größter Integritätsbereich sein. Weil die Zahl 1 in e_2 und e_1 vorkommt, sind in e_3 wegen $e_1 e_2 = e_3$ alle Quaternionen aus e_1 und alle Quaternionen aus e_2 enthalten. Das ist aber unmöglich, weil e_1 und e_2 selbst schon größte Integritätsbereiche sind.

46. Für jedes Ideal a gibt es zwei eindeutig bestimmte Einheitsideale e_1 und e_2 , so daß $e_1 a = a$ und $a e_2 = a$. Wir nennen e_1 das links zugehörige oder linke und e_2 das rechts zugehörige oder rechte Einheitsideal von a und a Linksideal von e_1 und Rechtsideal von e_2 .

Ist nämlich A eine Basismatrix und F die zugehörige Normenform, so gibt es nach meinen früheren Ergebnissen²⁸⁾ zu F links und rechts zugehörige Hauptformen H_1 und H_2 , welche die Kompositionen $H_1 F = F$ und $F H_2 = F$ gestatten. Zugleich müssen zwei diese Kompositionen vermittelnde bilineare Substitutionen von der Form

$$A^{-1} \cdot W \begin{matrix} \swarrow E_1 \\ \searrow A \end{matrix} \quad \text{und} \quad A^{-1} \cdot W \begin{matrix} \swarrow A \\ \searrow E_2 \end{matrix}$$

existieren, wobei E_1 und E_2 unimodulare Substitutionen bezeichnen, welche die Grundform G bzw. in H_1 und H_2 transformieren²⁹⁾.

Daraus folgt aber die Behauptung³⁰⁾.

Ist das Ideal a im besonderen ganz, so sind alle seine Quaternionen sowohl in dem linken wie in dem rechten Einheitsideal von a enthalten (12, 40).

47. Aus den früheren Betrachtungen über koordinierte bilineare Sub-

²⁸⁾ II, S. 308 ff.

²⁹⁾ III, S. 171.

³⁰⁾ Die links und rechts zugehörigen Einheitsideale können in Anlehnung an den Dedekindschen Begriff der Ordnung eines Moduls auch als Links- und Rechtsordnung von a aufgefaßt und hergeleitet werden.

stitutionen³¹⁾ ergibt sich, daß mit den eben angegebenen bilinearen Substitutionen auch diese

$$\tilde{A}^{-1} \cdot W \begin{matrix} \tilde{A} \\ E_1 \end{matrix}, \quad \tilde{A}^{-1} \cdot W \begin{matrix} E_2 \\ \tilde{A} \end{matrix}, \quad E_1^{-1} \cdot W \begin{matrix} A \\ \tilde{A} \end{matrix}, \quad E_2^{-1} \cdot W \begin{matrix} \tilde{A} \\ A \end{matrix}$$

ganzzahlig sind. Dabei ist

$$\tilde{A} = \frac{1}{a} E_* A E_0$$

Basismatrix des reziproken Ideals. Das besagt für die Ideale

$$a^{-1} e_1 = a^{-1}, \quad e_2 a^{-1} = a^{-1}, \quad a a^{-1} = e_1, \quad a^{-1} a = e_2.$$

48. Zwei Ideale a und b lassen sich in dieser Reihenfolge dann und nur dann multiplizieren, wenn das rechte Einheitsideal von a zugleich linkes Einheitsideal von b ist.

Wenn nämlich $ab = c$ und e das rechte Einheitsideal von a ist, so gilt $c = (ae)b = a(eb) = ab$, also $eb = b$ (43, 44).

Und wenn $ae = a$ und $eb = b$, also auch $bb^{-1} = e$, so existiert $a(bb^{-1})$, also auch ab (47, 43).

Daraus folgt: Wenn zwei Ideale a, a' das rechte oder zwei Ideale b, b' das linke Einheitsideal gemeinsam haben, so gibt es Ideale p, q , so daß $a' = pa$, $b' = bq$.

49. Sind e_1 und e_2 irgend zwei Einheitsideale, so gibt es stets Ideale a , die Linksideale für e_1 und Rechtsideale für e_2 sind.

Zum Beweise bilde man das Modulprodukt $e_1 \times e_2 = a$. Dann bestehen die Modulgleichungen

$$e_1 \times a = a, \quad a \times e_2 = a,$$

welche nach 36 erkennen lassen, daß a ein Ideal ist³²⁾. Deshalb kann man auch schreiben

$$e_1 a = a \quad \text{und} \quad a e_2 = a.$$

Als wichtige Folgerung ergibt sich:

Ist c ein beliebiges Ideal und e ein beliebiges Einheitsideal, so gibt es ein Rechtsideal a von e und ein Linksideal b von e , so daß $c = ab$ ist.

50. Die Sätze dieses Kapitels zeigen, daß die Ideale von \mathfrak{Q} in bezug auf die Multiplikation ein unendliches Gruppoid bilden³⁾.

³¹⁾ I, S. 111, oder Bilineare Transformation quadratischer Formen, Math. Zeitschr. 20 (1924), S. 153.

³²⁾ Sind e_1 und e_2 Einheitsideale, so wird also ein beliebiger Modul a durch jede dieser beiden Modulgleichungen als Ideal charakterisiert (vgl. 3).

Die Einheitselemente sind die Einheitsideale, die inversen Elemente sind die reziproken Ideale. Die links bzw. rechts einander zugehörigen Elemente sind die Ideale, welche das linke bzw. das rechte Einheitsideal gemeinsam haben. Doppelt zugehörig sind also diejenigen Ideale, welche sowohl das linke wie das rechte Einheitsideal gemeinsam haben.

Die einem Einheitsideal e doppelt zugehörigen Ideale sind die zweiseitigen Ideale des Integritätsbereichs e , d. h. diejenigen Ideale, die sowohl Rechts- wie Linksideale von e sind. Weil dazu alle Ideale re gehören, wo r eine beliebige rationale Zahl ist, so ist die Anzahl dieser Ideale, d. h. die Ordnung des Gruppoids unendlich. Daß auch der Rang, d. h. die Anzahl der Einheitselemente unendlich ist, wurde bereits gezeigt (41).

V. Idealklassen.

51. Wenn a ein beliebiges Ideal bezeichnet und ϱ, σ irgendwelche Quaternionen aus \mathfrak{Q} mit nicht verschwindenden Normen sind, so ist der Modul $a' = \varrho a \sigma$ ebenfalls ein Ideal.

Ist e_1 das linke und e_2 das rechte Einheitsideal von a , so ist $e'_1 = \varrho e_1 \varrho^{-1}$ das linke. und $e'_2 = \sigma^{-1} e_2 \sigma$ das rechte Einheitsideal von a' .

Die erste Behauptung ergibt sich nach unserer Idealdefinition (37) aus 32. Daraus folgt auch, daß e'_1, e'_2 Einheitsideale sind. e'_1, e'_2 genügen aber den Gleichungen $e'_1 a' = a'$ und $a' e'_2 = a'$, woraus sich die Richtigkeit der letzten Behauptung ergibt.

52. Auf Grund dieses Satzes stellen wir folgende Definitionen auf:

Zwei Ideale a und a' heißen äquivalent, wenn es zwei Quaternionen ϱ, σ gibt, so daß $n(\varrho \sigma) \neq 0$ und $a' = \varrho a \sigma$. Alle äquivalenten Ideale bilden eine Idealklasse.

Ist im besonderen $\sigma = 1$ bzw. $\varrho = 1$, so wird die Äquivalenz als rechtsseitige bzw. linksseitige Äquivalenz bezeichnet. Alle rechts bzw. links äquivalenten Ideale bilden eine Rechts- bzw. Linksidealklasse.

Offenbar haben rechts äquivalente Ideale dasselbe rechte und links äquivalente Ideale dasselbe linke Einheitsideal. Vgl. 59.

(Für Einheitsideale besagt die Äquivalenz dasselbe wie die Möglichkeit der isomorphen Zuordnung der von ihnen gebildeten Integritätsbereiche.)

53. In jeder Rechts- und in jeder Linksidealklasse gibt es ganze Ideale, deren Norm unter einer endlichen Schranke liegt.

Ist a ein beliebiges Ideal mit der Norm a , und e_1 das linke, e_2 das rechte Einheitsideal von a , so wähle man aus dem reziproken Ideal a^{-1} nach 39 ein Quaternion ϱ so aus, daß

$$0 < n(\varrho) \leq \frac{1}{a} 2 |\sqrt{a}|.$$

Bildet man dann die beiden Ideale $\alpha_1 = a \varrho$ und $\alpha_2 = \varrho a$, von denen α_1 der Linksidealklasse von a und α_2 der Rechtsidealklasse von a angehört, so ist wegen $a a^{-1} = e_1$ und $a^{-1} a = e_2$ α_1 in e_1 und α_2 in e_2 enthalten, beide Ideale bestehen also aus lauter ganzen Quaternionen. Man findet aber

$$n(\alpha_1) \leq 2 |\sqrt{d}|, \quad n(\alpha_2) \leq 2 |\sqrt{d}|.$$

54. Für jedes Einheitsideal hat die Anzahl der Rechts- und Linksidealklassen, in die die rechts bzw. links äquivalenten Ideale zerfallen, denselben endlichen von der ausgewählten Einheitsklasse unabhängigen Wert h , der als Klassenzahl der Algebra bezeichnet wird.

Wählt man als Einheitsideal zunächst das Grundideal \mathfrak{o} aus, so kann man die Rechtsideale von \mathfrak{o} auf Rechtsidealklassen verteilen und jede Klasse nach dem vorigen Satze durch ein ganzes Ideal, dessen Norm kleiner als $2 |\sqrt{d}|$ ist, repräsentieren.

Die Basismatrizen dieser Ideale sind nach der am Schlusse von 46 gemachten Bemerkung ganzzahlig. Weil zwei Basismatrizen A und AU , wo U eine ganzzahlige Matrix mit der Determinante $+1$ bezeichnet, dasselbe Ideal liefern, so ist die Anzahl der repräsentierenden Ideale gewiß kleiner als die Anzahl der Klassen von Substitutionen mit einer positiven Determinante $< 4 |d|$, also endlich.

Wird die Anzahl der Rechtsidealklassen von \mathfrak{o} durch h bezeichnet, so ist klar, daß die Anzahl der Linksidealklassen von \mathfrak{o} ebenso groß ist. Durchläuft nämlich a alle Ideale der Rechtsidealklassen von \mathfrak{o} , so durchläuft entsprechend \bar{a} alle Ideale der Linksidealklassen von \mathfrak{o} .

Endlich ist die Anzahl auch von der ausgewählten Einheitsklasse ganz unabhängig. Ist nämlich e irgendein Einheitsideal, so kann man ein Ideal p finden, das \mathfrak{o} als linkes und e als rechtes Einheitsideal hat (49). Durchläuft dann a alle Rechtsideale von \mathfrak{o} , so durchläuft $a' = ap$ alle Rechtsideale von e . Zwei Ideale a' sind aber dann und nur dann rechts äquivalent, wenn die entsprechenden Ideale a rechts äquivalent sind.

55. Für das Grundideal \mathfrak{o} sind die Basismatrizen der ganzen und primitiven Rechts- und Linksideale mit zu d primen Norm mit den erzeugenden Substitutionen von positivem bzw. negativem Charakter³³⁾ identisch.

Es genügt zu zeigen, daß die Rechtsideale mit den genannten Eigenschaften vollständig den Erzeugenden von positivem Charakter entsprechen, weil der zweite Teil der Behauptung genau entsprechend folgt. Das ergibt sich aber aus dem Satz:

³³⁾ IV, S. 180.

Von den ganzzahligen primitiven erzeugenden Substitutionen A mit der zu d primen Norm a liefern alle und nur diejenigen eine ganzzahlige bilineare Substitution

$$A^{-1} \cdot W \begin{matrix} \diagup \\ A \end{matrix} = M,$$

welche die Elementarteiler $1, 1, a, a$ und positiven Charakter haben.

Der erste Teil dieses Satzes ergibt sich aus früheren Betrachtungen³⁴⁾. (Diese erfordern, wenn d ungrade und a grade, nur eine kleine Ergänzung.) Für den zweiten Teil beachte man, daß mit M nach 47 auch die bilinearen Substitutionen

$$W \begin{matrix} \diagup \\ \tilde{A} \\ \diagdown \\ A \end{matrix} = N$$

ganz ausfallen³⁵⁾. Deshalb gestattet A die Parameterdarstellung

$$A = \left\| \sum_j n_{ijk} p_j \right\|,$$

welche, wie früher gezeigt, nur solche ganzzahligen primitiven Substitutionen A mit einer zu d primen Determinante a^2 liefern kann, welche die Elementarteiler $1, 1, a, a$ und positiven Charakter haben³⁶⁾.

56. *Es gibt kein ganzes und primitives Ideal mit zu d primen Norm, das rechts und links zu demselben Einheitsideal e gehört außer e selbst.*

Würde man nämlich eine Basis von e als Grundbasis annehmen, so würde nach dem vorigen Satz die Basismatrix eines solchen Ideals eine erzeugende Substitution sein, die sowohl positiven wie negativen Charakter hat, was unmöglich ist.

Untersucht man bei Stammformen die ganzen und primitiven erzeugenden Substitutionen, deren Norm nur Primteiler von d enthält, so ergeben sich die Sätze:

Die Norm eines ganzen und primitiven Ideals kann einen Diskriminantenteiler stets nur einfach, nicht im Quadrat enthalten.

Ein ganzes und primitives Ideal ist dann und nur dann gleichseitig, d. h. gehört links und rechts zu demselben Einheitsideal, wenn seine Norm ein Teiler t von d ist, und für jeden Teiler t gibt es zu jedem Einheitsideal e gerade ein einziges derartiges Ideal.

57. Aus diesen Sätzen ergeben sich wichtige Folgerungen für die aus gleichseitigen Idealen gebildeten Gruppen. Dabei soll eine solche Gruppe mit einem Ideal i auch alle rationalen Multipla ri enthalten.

³⁴⁾ IV, S. 192.

³⁵⁾ Die sämtlichen A , welche dasselbe M liefern, ergeben auch dasselbe N und umgekehrt.

³⁶⁾ IV, S. 187.

Eine derartige Gruppe \mathfrak{h} wird dann vollständig charakterisiert durch die darin enthaltenen ganzen und primitiven gleichseitigen Ideale oder nach 56 auch durch die Normen dieser Ideale, d. h. durch ein System von Teilern t von d . Diese Teiler bilden dann selbst eine multiplikative Gruppe, wenn man die bei der Multiplikation etwa auftretenden quadratischen Faktoren jedesmal entfernt. Die Anzahl dieser Teiler und somit auch die der ganzen und primitiven in \mathfrak{h} enthaltenen Ideale ist somit eine Potenz von 2.

Daraus schließt man, daß auch der Index der Gruppe \mathfrak{h} in der Gruppe \mathfrak{g} aller zu demselben Einheitsideal e gehörigen gleichseitigen Ideale eine Potenz von 2 sein muß.

58. Ist umgekehrt ein System (t) von Teilern t von d so beschaffen, daß es in dem eben genannten Sinne eine multiplikative Gruppe bildet, so werden dadurch für irgend zwei Einheitsideale e_1 und e_2 in eindeutiger Weise Gruppen \mathfrak{h}_1 und \mathfrak{h}_2 von gleichseitigen Idealen bestimmt.

Diese Gruppen sind isomorph aufeinander bezogen, wenn man die demselben Teiler t von d entspringenden, bzw. in \mathfrak{h}_1 und \mathfrak{h}_2 liegenden ganzen Ideale i_1 und i_2 und ihre entsprechenden rationalen Multiplika ri_1 und ri_2 einander zuordnet.

Ist a ein beliebiges Ideal, das links zu e_1 und rechts zu e_2 gehört, und sind j_1 und j_2 irgend zwei einander in dieser Weise entsprechende Ideale aus den Gruppen \mathfrak{h}_1 und \mathfrak{h}_2 , so gilt stets $j_1 a = a j_2$.

59. Hält man a fest und läßt j_1 die Gruppe \mathfrak{h}_1 oder j_2 die Gruppe \mathfrak{h}_2 durchlaufen, so werde die Gesamtheit der Ideale $j_1 a = a j_2$ durch (a) bezeichnet und die *durch das Teilersystem (t) bestimmte Schar* von a genannt.

Sind a, b Ideale, für welche das Produkt $ab = c$ existiert, und durchläuft a die Schar (a) , b die Schar (b) , so durchläuft auch c die Schar (c) , die wir daher durch $(a)(b) = (c)$ bezeichnen und das *Produkt der Scharen* (a) und (b) nennen können.

Aus diesem Satz ergibt sich, daß auch die Scharen ebenso wie die Ideale *durch die Multiplikation zu einem Gruppoid verknüpft* sind. Die Ordnung dieses Gruppoids ist aber endlich, nämlich gleich dem Quotienten der Anzahl aller Teiler von d durch die Anzahl der in dem System (t) enthaltenen Teiler oder gleich dem Index der für irgendein Einheitsideal e gebildeten Gruppe \mathfrak{h} in \mathfrak{g} (57).

60. Die sämtlichen ganzen Quaternionen μ , für welche $n(\mu) \neq 0$ und der größte gemeinsame Teiler $(n(\mu), d)$ dem Teilersystem (t) angehört, bilden zusammen mit den rationalen Multipla $r\mu$ eine multiplikative Gruppe, welche wir die *zu dem Teilersystem (t) gehörige Multiplikatorengruppe* nennen wollen.

Ist α ein beliebiges Ideal, das links zu dem Einheitsideal e_1 und rechts zu dem Einheitsideal e_2 gehört, sind ferner τ_1 und τ_2 zwei Quaternionen der Multiplikatorengruppe, die den Bedingungen $\tau_1 e_1 = e_1 \tau_1$, $\tau_2 e_2 = e_2 \tau_2$ genügen, so ist das Ideal $\tau_1 \alpha \tau_2$ in der durch das Teilersystem (t) bestimmten Schar (α) enthalten (58).

Sind ϱ, σ irgend zwei Quaternionen der Multiplikatorengruppe und durchläuft das Ideal α die ganze Schar (α) , so durchläuft das Ideal $\beta = \varrho \alpha \sigma$ ebenfalls die Schar (β) , die wir daher durch $(\beta) = \varrho(\alpha)\sigma$ bezeichnen und der Schar (α) *in bezug auf das Teilersystem (t) äquivalent* nennen können.

Alle äquivalenten Scharen werden zu einer *Scharenklasse* zusammengefaßt, und die in einer Scharenklasse enthaltenen Ideale heißen *in bezug auf das Teilersystem (t) äquivalent* und bilden eine *Idealklasse in bezug auf das Teilersystem (t)* .

61. Wir nennen eine solche Ideal- oder Scharenklasse \mathfrak{R}_1 *komponierbar* mit einer Klasse \mathfrak{R}_2 , wenn es in \mathfrak{R}_1 Scharen (α) und in \mathfrak{R}_2 Scharen (β) gibt, für welche das Produkt $(\alpha)(\beta) = (\gamma)$ existiert. Bildet man auf alle möglichen Weisen solche Produkte, so ergeben die Scharen (γ) wieder eine Scharenklasse \mathfrak{R}_3 , die wir *komponiert aus \mathfrak{R}_1 und \mathfrak{R}_2* nennen.

Bezeichnet man nämlich durch e das rechts zu α und links zu β gehörige Einheitsideal, so kann die Schar $\varrho(\alpha)\varrho_1$ nur mit der Schar $\sigma_1(\beta)\sigma$ multipliziert werden, wenn $\varrho_1 \sigma_1 = \tau$ der Bedingung $\tau e = e \tau$ genügt (48). Daher ist nach dem zweiten Satz des vorigen Paragraphen $(\alpha)\tau = (\alpha)$ oder $\tau(\beta) = (\beta)$ und somit $\varrho(\alpha)\tau(\beta)\sigma = \varrho(\alpha)(\beta)\sigma = \varrho(\gamma)\sigma$.

Deshalb bilden die *Idealklassen in bezug auf das Teilersystem (t) wieder ein Gruppoid*. Dies Gruppoid kann aufgefaßt werden als Faktorgruppoid³⁷⁾ aus dem Gruppoid aller Ideale und dem Gruppoid der Ideale einer *Einheitsklasse*, d. h. einer Klasse $\varrho(e)\sigma$, wo e ein beliebiges Einheitsideal bezeichnet. Statt der Ideale kann man auch die Scharen als die Elemente der beiden Gruppoiden ansehen, ohne daß das Faktorgruppoid sich ändert.

Der Rang des Gruppoids der Klassen ist endlich, nämlich gewiß kleiner als das Produkt $h2^e$, wo h dieselbe Bedeutung hat wie in 54 und e die Anzahl der Primteiler von d bezeichnet.

62. Wenn das Teilersystem (t) nur aus der Zahl 1 besteht, erhält man den engsten Klassenbegriff. Stellt man nun für das Gruppoid \mathcal{G} dieser Klassen eine Kompositionstafel her, so kann man daran auch die Kompositionsgesetze für die Idealklassen in bezug auf ein beliebiges Teilersystem (t) veranschaulichen.

³⁷⁾ Vgl. F. K. Schmidt, Bemerkungen zum Brandtschen Gruppoid. Sitzungsberichte der Heidelberger Akademie 1927, 8. Abh., S. 94.

Man erhält nämlich die Idealklassen in bezug auf das Teilersystem (t), wenn man die Elemente dieses Gruppoids in geeigneter Weise zu Komplexen zusammenfaßt. Ein Komplex \mathfrak{k} , der ein Einheitselement E des Gruppoids \mathfrak{G} enthält, ist eine Gruppe, und das gesuchte Klassengruppoid kann als Faktorgruppoid $\mathfrak{G}/\mathfrak{k}$ aufgefaßt werden. Zugleich kann man durch geeignete Anordnung der Zeilen und Spalten der Tafel von \mathfrak{G} erreichen, daß die Komplexe sich auf quadratische Bezirke verteilen, die zugleich die Felder einer Kompositionstafel für das gesuchte Klassengruppoid darstellen³⁾.

63. Aber auch die Idealklassen im früheren Sinne (52) können aus den Elementen des Gruppoids \mathfrak{G} aufgebaut werden.

Durchläuft nämlich t_s alle Teiler von d , so kann man stets ganze Quaternionen μ_s in der Algebra \mathfrak{Q} finden, für welche der größte gemeinsame Teiler, den die Norm von μ_s mit d gemeinsam hat, gerade t_s ist. Ist dann \mathfrak{A} eine beliebige Idealklasse, in der sich das Element A befindet, so enthält \mathfrak{A} alle und nur die Elemente $\mu_s A \mu_s$. Diese Elemente sind aber nicht sämtlich verschieden voneinander.

Um die Anzahl der verschiedenen Elemente zu bestimmen, verstehen wir unter dem *Index eines Ideals* α die Anzahl der ganzen und primitiven, in der Form $\nu_1 \alpha = \alpha \nu_2$ darstellbaren Ideale. Ähnliche Betrachtungen wie in 57 zeigen, daß der Index endlich und eine Potenz von 2 ist.

Der Index hat für alle Ideale derselben Idealklasse für ein beliebiges Teilersystem und auch für alle Ideale einer Idealklasse schlechthin stets denselben Wert.

Wird die Anzahl der Primfaktoren von d mit e bezeichnet, so daß es 2^e Teiler von d gibt, so finden sich in der Idealklasse \mathfrak{A} vom Index 2^i gerade 2^{e-i} verschiedene Elemente.

64. Ist $\mathfrak{A} = \mathfrak{G}$ speziell eine Einheitsklasse, d. h. eine solche, die ein Einheitsideal e oder auch ein Einheitselement E des Gruppoids \mathfrak{G} enthält, so bilden die 2^{2e-i} Elemente von \mathfrak{G} selbst ein Gruppoid vom Range 2^{e-i} und von der Ordnung 2^i , das ein Teilgruppoid von \mathfrak{G} ist.

Wir wollen nun die Kompositionstafel für \mathfrak{G} so anordnen, daß in quadratischen Bezirken von 2^{2e} Feldern entlang der von links oben nach rechts unten verlaufenden Diagonale gerade diese Teilgruppoiden der Einheitsklassen \mathfrak{G} erscheinen, wobei die Einheitselemente von \mathfrak{G} zweckmäßig auf der Diagonale selbst stehen. Ist E eins von ihnen, so besteht die Zeile des Bezirkes von E aus den Elementen $\mu_s E$ und die Spalte aus den Elementen $E \mu_s$.

Durch diese Anordnung ist nun zugleich eine Einteilung der ganzen Tafel in horizontale und vertikale Streifen von der Breite 2^e und damit auch eine Einteilung in quadratische Bezirke von 2^{2e} Feldern gegeben.

Ist A ein Element eines solchen Bezirkes und trifft die Spalte von A die Diagonale im Feld E_1 , die Zeile im Feld E_2 , so enthält der Bezirk wegen der Grundeigenschaft der Tafel³⁾ alle und auch nur die Elemente

$$\mu_s E_1 \cdot A \cdot E_2 \mu_{s'} = \mu_s A \mu_{s'},$$

umfaßt also gerade eine Idealklasse \mathfrak{A} .

65. Wenn dasselbe Element A außerhalb des betrachteten Bezirkes noch an einer anderen Stelle der Tafel vorkommt, so enthält der neue Bezirk wieder die ganze Klasse \mathfrak{A} und keine anderen Elemente. Betrachtet man nun einen von diesen Bezirken, welche die Idealklasse \mathfrak{A} enthalten und die beiden Streifen von der Breite 2^e , welche sich in diesem Bezirk kreuzen, so möge die Idealklasse, welche sich in dem Diagonalfeld des vertikalen Streifens befindet, durch \mathfrak{C}_1 und die Idealklasse, welche sich in dem Diagonalfeld des horizontalen Streifens befindet, durch \mathfrak{C}_2 bezeichnet werden. Wir nennen \mathfrak{C}_1 *die links* und \mathfrak{C}_2 *die rechts zu \mathfrak{A} gehörige Einheitsklasse*.

Sind dann die Indizes der drei Klassen \mathfrak{A} , \mathfrak{C}_1 , \mathfrak{C}_2 bzw. 2^i , 2^{i_1} , 2^{i_2} , wobei, wie man leicht zeigt, i höchstens so groß ist wie die kleinere der beiden Zahlen i_1 und i_2 , so enthalten diese Klassen 2^{2e-i} , 2^{2e-i_1} , 2^{2e-i_2} verschiedene Elemente. In einem Streifen der Kompositionstafel eines Gruppoids kommt aber jedes Element und somit auch jeder Komplex von Elementen gleich oft vor. Der Komplex \mathfrak{C}_1 kommt in seinem Bezirk 2^{i_1} -mal und in dem vertikalen Streifen sonst nicht vor, ebenso kommt der Komplex \mathfrak{C}_2 in seinem Bezirk 2^{i_2} -mal und in dem horizontalen Streifen sonst nicht vor.

Da nun der Komplex \mathfrak{A} in jedem Bezirk, wo er auftritt, 2^i -mal vorkommt und ein Bezirk aus dem Komplex \mathfrak{A} entweder alle oder keine Elemente enthält, so muß der vertikale Streifen 2^{i_1-i} und der horizontale Streifen 2^{i_2-i} Bezirke enthalten, in denen sich jedesmal die Klasse \mathfrak{A} befindet.

66. Die Einteilung in Bezirke kann zugleich als Kompositionstafel für die Idealklassen in dem früheren Sinn angesehen werden. Für den Gebrauch der Tafel muß man nur die für ein Gruppoid³⁾ oder auch für die Formenklassen³⁸⁾ gegebene Regel anwenden. Die bei der Komposition der Klassen i. a. auftretenden Mehrdeutigkeiten werden durch die Tafel gerade vollständig zum Ausdruck gebracht.

Alle Klassen \mathfrak{A} , für die die Indizes 2^i , 2^{i_1} , 2^{i_2} denselben Wert haben, schließen sich zu einem oder auch zu mehreren Gruppoiden zusammen.

³⁸⁾ Verhandlungen der Schweizerischen Naturforschenden Gesellschaft 1924, II. Teil, S. 102.

Alle Klassen, die derselben Einheitsklasse links und rechts zugehören, bilden eine Idealklasse in bezug auf das System aller Teiler von d , haben daher denselben Index (63) und bilden eine Gruppe.

Die Idealklassen entsprechen vollständig den von den Normenformen gebildeten Formenklassen, wenn $d > 0$ und wenn $d < 0$ bei denjenigen Algebren \mathfrak{Q} , bei denen es ganze Quaternionen der Norm -1 gibt. Ist das nicht der Fall, so besteht ein Unterschied, der sich aber beheben läßt, wenn man den Äquivalenzbegriff verengt.

67. Ist α ein beliebiges Ideal der Klasse \mathfrak{A} , e_1 das linke und e_2 das rechte Einheitsideal von α , so betrachten wir die für die Einheitsideale e_1 und e_2 gebildeten Rechts- und Linksideale und deren Klassen und fragen, wieviel von ihnen auf die Klasse \mathfrak{A} entfallen. Man findet für e_1 2^{i_1-i} und für e_2 2^{i_2-i} Klassen, also dieselben Anzahlen³⁹⁾, welche angeben, wie oft die Klasse \mathfrak{A} in den Spalten und Zeilen der Kompositionstafel vorkommt (66).

Daraus ergibt sich, daß die Anzahl der Zeilen und Spalten in der Kompositionstafel gleich der früheren Klassenzahl h ist (54).

Wendet man zur Bestimmung von h die bekannten transzendenten Methoden von Dirichlet an, so ergibt sich für $d > 3$ die Formel

$$h_1 + \frac{1}{2}h_2 + \frac{1}{3}h_3 = \frac{1}{12}\varphi(d),$$

wobei $h_1 + h_2 + h_3 = h$ und φ die Eulersche Funktion bezeichnet⁴⁰⁾.

³⁹⁾ Siehe auch III, S. 172 Fußnote.

⁴⁰⁾ Die §§ 57–67 haben bei der Korrektur im Dezember 1927 eine vollständige Umarbeitung erfahren wegen eines irrümlichen Diskriminantensatzes in der früheren Darstellung, der mich veranlaßt hatte, meine bereits vor mehreren Jahren auf dem Begriff der Relativkomposition beruhenden Untersuchungen über die Kompositionstafeln der Formenklassen (vgl. meine Vorträge in Nauheim 1920 und Luzern 1924) bei maximalen Ordnungen (Stammformen) für entbehrlich zu halten. Diese Untersuchungen erscheinen hier in den §§ 63–67 in ganz neuer Form, so daß zugleich ihre Verallgemeinerung auf beliebige Dedekindsche Algebren ersichtlich ist. Die §§ 57–62 sind auch inhaltlich neu und aus dem Gedanken entstanden, den Faktorgruppoidbegriff in allgemeinerer Auffassung für die Klassenbildung zu benutzen.

Den Hinweis auf das Versehen verdanke ich Herrn Artin, der übrigens in seiner im März 1927 erschienenen, mir aber erst nach Abschluß der Fahnenkorrektur bekannt gewordenen, daher oben nicht berücksichtigten Arbeit „Zur Arithmetik hyperkomplexer Größen“, Hamburger Abhandlungen 1927, S. 282 einen wesentlichen Teil des in der Einleitung genannten allgemeinen Satzes vom Gruppoid der Ideale bewiesen hat.

Über die endlichen Gruppen ohne das Gesetz der eindeutigen Umkehrbarkeit.

Von

Anton Suschkewitsch in Woronesch (Rußland).

Einleitung.

In der vorliegenden Abhandlung habe ich den Versuch gemacht eine abstrakte Theorie der endlichen Gruppen, deren Operation nicht eindeutig umkehrbar ist, zu konstruieren. Freilich sind in der mathematischen Literatur solche Gruppen in konkreter Form schon betrachtet worden. Als Beispiel solcher konkreten Gruppen kann man die Theorie der nicht-kommutativen Ringe, speziell auch die Theorie der hyperkomplexen Zahlen anführen, wobei auch das Analogon zu der Aufspaltung des besonderen Teiles, den ich als „Kern“ bezeichne, durchgeführt ist¹⁾. Dabei werden aber zugleich zwei Operationen betrachtet: die „Addition“ und die „Multiplikation“. Es entsteht nun die Frage nach der Verallgemeinerung, die man erhält, wenn man die eine Operation — nämlich die Addition — wegläßt und bloß die andere — die Multiplikation — beibehält, die als eindeutig, assoziativ, aber nicht eindeutig umkehrbar vorausgesetzt wird.

Die Darstellung, die ich im folgenden einführe, ist von diesen konkreten Fällen völlig unabhängig. Ich bleibe fortwährend im Gebiete der reinen Gruppentheorie, betrachte also nur eine einzige Operation in einer völlig abstrakten Form und beschränke meine Betrachtungen ausschließlich auf *endliche* Gruppen mit einer eindeutigen assoziativen Operation.

Bekanntlich hat das Gesetz der eindeutigen Umkehrbarkeit zwei Seiten: das „linke“ Gesetz: „Aus der Gleichung $BA = CA$ folgt $B = C$ “; das „rechte“ Gesetz: „Aus $AB = AC$ folgt $B = C$ “. Sofern über die Kommutativität der Operation keine Voraussetzung gemacht wird, sind diese beiden Seiten völlig unabhängig voneinander.

¹⁾ Vgl. MacLagan-Wedderburn, On hypercomplex numbers, Proceedings of the London Math. Soc. (2) 6 (1908). Ich bin auf diese Arbeit erst nach Fertigstellung der meinigen durch einen freundlichen Hinweis von Fr. E. Noether aufmerksam geworden.

Bei der Eindeutigkeit der Operation und der Endlichkeit der Gruppe ist jede Seite des Gesetzes der eindeutigen Umkehrbarkeit gleichbedeutend mit der entsprechenden Seite des Gesetzes der unbeschränkten Umkehrbarkeit; die linke Seite dieses Gesetzes besagt, daß die Gleichung $XA = B$, wo A und B zwei beliebige Elemente der Gruppe sind, immer wenigstens eine Lösung X aus derselben Gruppe hat; entsprechend die rechte Seite. Es möge bemerkt werden, daß diese Gleichwertigkeit bei den endlichen Gruppen zu ihrer Herstellung keines Assoziativgesetzes bedarf.

Zunächst fasse ich die Eigenschaften der Gruppen, bei denen die eine Seite des Gesetzes der eindeutigen Umkehrbarkeit gilt, zusammen. Danach gehe ich zum allgemeinen Fall über, mit folgenden Resultaten:

I. Es stellt sich heraus, daß jede Gruppe ohne das Gesetz der eindeutigen Umkehrbarkeit eine besondere Untergruppe \mathfrak{K} , die ich „Kerngruppe“ nenne, hat. Diese Kerngruppe wird im folgenden untersucht. Sie hat folgenden Bau:

$$\mathfrak{K} = \sum_{\kappa=1}^r \mathfrak{A}_\kappa = \sum_{\lambda=1}^s \mathfrak{B}_\lambda = \sum_{\kappa=1}^r \sum_{\lambda=1}^s \mathfrak{C}_{\kappa\lambda}; \quad \mathfrak{A}_\kappa = \sum_{\lambda=1}^s \mathfrak{C}_{\kappa\lambda}; \quad \mathfrak{B}_\lambda = \sum_{\kappa=1}^r \mathfrak{C}_{\kappa\lambda};$$

dabei sind die \mathfrak{A}_κ Gruppen, bei denen die linke Seite des Gesetzes der eindeutigen Umkehrbarkeit erfüllt ist („Linksgruppen“), die \mathfrak{B}_λ Gruppen mit der rechten Seite desselben Gesetzes („Rechtsgruppen“); die $\mathfrak{C}_{\kappa\lambda}$ sind gewöhnliche Gruppen. Alle \mathfrak{A}_κ sind zueinander einstufig isomorph; ebenso alle \mathfrak{B}_λ und alle $\mathfrak{C}_{\kappa\lambda}$. Keine zwei der Gruppen $\mathfrak{C}_{\kappa\lambda}$ haben ein Element gemeinsam (auch ihre Einheiten sind voneinander verschieden). Schematisch kann \mathfrak{K} folgendermaßen dargestellt werden (s. Schema I):

$$\begin{array}{l} \mathfrak{K} = \mathfrak{A}_1 + \mathfrak{A}_2 + \cdots + \mathfrak{A}_r \\ \parallel \qquad \parallel \qquad \parallel \qquad \parallel \\ \mathfrak{B}_1 = \left[\begin{array}{cccc} \mathfrak{C}_{11} + & \mathfrak{C}_{21} + & \cdots + & \mathfrak{C}_{r1} \\ + & + & & + \\ \mathfrak{B}_2 = & \mathfrak{C}_{12} + & \mathfrak{C}_{22} + & \cdots + & \mathfrak{C}_{r2} \\ + & + & & & + \\ \vdots & \vdots & & \vdots & \vdots \\ + & + & & + & + \\ \mathfrak{B}_s = & \mathfrak{C}_{1s} + & \mathfrak{C}_{2s} + & \cdots + & \mathfrak{C}_{rs} \end{array} \right. \end{array}$$

Schema I.

II. Es wird ferner gezeigt, daß die Kerngruppe \mathfrak{K} vollständig bestimmt ist, sofern gegeben sind:

1. Die Struktur der gewöhnlichen Gruppen $\mathfrak{C}_{\kappa\lambda}$;
2. die Zahlen r und s ;
3. die $(r - 1)(s - 1)$ Produkte $E_{11}E_{\kappa\lambda}$ ($\kappa = 2, \dots, r$; $\lambda = 2, \dots, s$); dabei bedeutet allgemein $E_{\kappa\lambda}$ die Einheit der gewöhnlichen Gruppe $\mathfrak{C}_{\kappa\lambda}$.

III. Es entsteht nun die Frage, ob diese drei Bedingungen auch willkürlich gewählt werden können; diese Frage wird im folgenden bejahend beantwortet. Zu dem Zwecke beweise ich, daß man bei diesen drei willkürlich gegebenen Bedingungen immer eine ihnen genügende Kerngruppe konstruieren kann, und zwar in konkreter Form als Gruppe verallgemeinerter Substitutionen. Darunter verstehe ich Substitutionen, bei denen mehrere verschiedene Symbole in ein und dasselbe Symbol übergehen können, die also, in gewöhnlicher Substitutionsform dargestellt, in ihrer unteren Zeile nicht alle Symbole zu haben brauchen, dagegen ein und dasselbe Symbol auch mehr als einmal haben können. Die Anzahl aller solcher Substitutionen der n Symbole ist n^n (gewöhnliche Substitutionen sind hier als Spezialfall mitbegriffen); diese Substitutionen können auch miteinander komponiert werden; es ist leicht zu sehen, daß bei dieser Komposition das assoziative Gesetz gilt; dagegen gilt das Gesetz der eindeutigen Umkehrbarkeit im allgemeinen nicht. Aus diesen Substitutionen können also endliche Gruppen der von uns betrachteten Art konstruiert werden, speziell auch Kerngruppen, und zwar *jede beliebige* Kerngruppe, wie ich im folgenden zeige. Als Anhang gebe ich ein Beispiel einer Kerngruppe, die auf diese Art konstruiert ist. Das Beispiel ist möglichst allgemein gewählt, um die Theorie in allen ihren Teilen zu veranschaulichen.

Bei meinen Bezeichnungen lehne ich mich an die Abhandlungen von Frobenius an³⁾: mit großen lateinischen Buchstaben bezeichne ich die Elemente, mit großen deutschen Buchstaben die Mengen (Komplexe) der Elemente, speziell auch Gruppen. Den Komplex bezeichne ich als „Summe“ der Elemente: $\mathfrak{A} = A + B + C + \dots$, oder $\mathfrak{A} = \sum_{\kappa=1}^n A_{\kappa}$. Ist $\mathfrak{A} = \sum_{\kappa} A_{\kappa}$, $\mathfrak{B} = \sum_{\lambda} B_{\lambda}$, so ist $\mathfrak{A}\mathfrak{B} = \sum_{\kappa, \lambda} A_{\kappa}B_{\lambda}$; dabei werden die einander gleichen Elemente als ein einziges Element gezählt. Gehört das Element A dem Komplex \mathfrak{B} an, so wird das folgendermaßen bezeichnet: $A < \mathfrak{B}$; ebenso bedeutet $\mathfrak{A} < \mathfrak{B}$, daß alle Elemente von \mathfrak{A} dem \mathfrak{B} angehören.

Wird für ein Elementenaggregat eine neue abgekürzte Bezeichnung eingeführt, so gebrauche ich statt des Gleichheitszeichens = das Zeichen \simeq : z. B. bedeutet $C \simeq AB$ oder $AB \simeq C$, daß man einfach das Produkt AB mit einem einzigen Buchstaben C bezeichnet hat.

§ 1.

Es sei eine Menge (Komplex) von Elementen gegeben; außerdem sei auch eine Verknüpfungsart (Operation) dieser Elemente definiert, mittelst der zwei beliebige (in bestimmter Ordnung genommene) Elemente mit-

³⁾ Vgl. z. B. Frobenius, Über endliche Gruppen, Sitzungsber. der Berl. Akad. 1895.

einander komponiert werden können, so daß sich ein bestimmtes Element als Resultat („Produkt“) dieser Komposition herausstellt. Diese Elementenmenge insgesamt mit der Operation soll nun folgenden Grundpostulaten unterliegen:

- I. Die Operation ist eindeutig und unbeschränkt anwendbar.
- II. Es gilt für sie das assoziative Gesetz.
- III. Die Menge enthält nur eine endliche Anzahl von Elementen.

Eine solche Menge \mathcal{G} nennen wir *Gruppe*. Es kann vorkommen, daß schon ein Teil \mathcal{A} der Elemente von \mathcal{G} auch eine Gruppe bilden, eine „Untergruppe“ von \mathcal{G} . Um sich zu überzeugen, daß \mathcal{A} eine Gruppe ist, braucht man bloß aufzustellen, daß das Produkt zweier beliebiger Elemente von \mathcal{A} auch dem \mathcal{A} angehört, oder symbolisch ausgedrückt, daß $\mathcal{A}^2 < \mathcal{A}$ ist.

Ist \mathcal{A} eine Gruppe und das Element $P < \mathcal{A}$, so sind $\mathcal{A}P$ und $P\mathcal{A}$ auch Gruppen, und es ist im allgemeinen $\mathcal{A}P < \mathcal{A}$, $P\mathcal{A} < \mathcal{A}$.

Betrachten wir nun die Potenzen eines beliebigen Elementes A ; aus II folgt leicht, daß die Gleichung

$$(1) \quad A^x A^\lambda = A^{x+\lambda}$$

für beliebige ganze positive x und λ gilt. Ferner folgt nach III, daß alle diese Potenzen nicht ohne Ende voneinander verschieden sein können; es wird vorkommen, daß

$$(2) \quad A^{k+m} = A^k$$

ist; die kleinsten (positiven) Werte von k und m , für die die Gleichung (2) erfüllt ist, nennen wir: k die *Art*, m die *Ordnung* des Elementes A . Aus (2) folgt:

$$A^{k+m+\lambda} = A^{k+\lambda}$$

für jedes ganze $\lambda > 0$. Die Elemente $A, A^2, \dots, A^{k-1}, A^k, \dots, A^{k+m-1}$ sind alle voneinander verschieden; dabei kommen A, A^2, \dots, A^{k-1} nur *einmal* vor; A^k, \dots, A^{k+m-1} dagegen wiederholen sich weiter periodisch; es ist dann und nur dann $A^x = A^\lambda$, wenn x und $\lambda \geq k$ sind, und $x \equiv \lambda \pmod{m}$ ist.

Aus (1) und (2) folgt nun, daß alle Potenzen eines Elementes A eine Gruppe von der Ordnung $k + m - 1$ bilden; für diese Gruppe gilt auch das kommutative Gesetz.

§ 2.

Betrachten wir jetzt den speziellen Fall, wo außer den Postulaten I–III noch das folgende Postulat erfüllt ist:

IV₁. Es gilt das *linke* Gesetz der eindeutigen (also nach III auch der unbeschränkten) Umkehrbarkeit.

(Die rechte Seite dieses Gesetzes bezeichnen wir mit IV_7 ; im Falle, den wir betrachten, wird sie im allgemeinen nicht erfüllt sein.)

Es ist leicht zu sehen, daß IV_7 mit folgendem Satze gleichbedeutend ist: Ist P ein beliebiges Element der Gruppe \mathfrak{A} , so ist $\mathfrak{A}P = \mathfrak{A}$ (dagegen ist $P\mathfrak{A} < \mathfrak{A}$). Daraus folgt übrigens, daß jede Gruppe \mathfrak{A} der Gleichung $\mathfrak{A}^2 = \mathfrak{A}$ genügt.

Nun stellen wir die Eigenschaften unserer „Linksgruppen“ zusammen – die Beweise folgen:

1. Es existieren in einer Linksgruppe \mathfrak{A} immer rechte Einheiten, und zwar sind das alle Elemente, die der Gleichung $E^2 = E$ genügen (die wir als „Hauptelemente“ bezeichnen). Jedes Element aus \mathfrak{A} hat seine eigene linke Einheit (zu der dieses Element „gehört“), die zugleich rechte Einheit für alle Elemente aus \mathfrak{A} ist. Hat die Gruppe \mathfrak{A} ein und dieselbe linke Einheit für alle Elemente oder hat sie bloß eine einzige rechte Einheit, so ist sie eine gewöhnliche Gruppe.

2. Die Art jedes Elementes von \mathfrak{A} ist $= 1$; also bilden die Potenzen eines Elementes eine gewöhnliche zyklische Gruppe. Wir können in diesem Falle auch nullte und negative Potenzen definieren: $A^0 \simeq A^m \simeq E$; $A^{-\lambda} \simeq A^{m-\lambda}$; dabei ist m die Ordnung von A ; $A^m \simeq E$ ist die Einheit, zu der A und alle Potenzen von A gehören. Die Gleichung (1) gilt für beliebige ganze λ und λ' ; speziell ist $A^\lambda A^{-\lambda} = E$, also $A^{-\lambda}$ invers zu A^λ .

3. Sind E_1, E_2, \dots, E_s alle rechte Einheiten der Gruppe \mathfrak{A} , so ist:

$$(3) \quad \mathfrak{A} = E_1 \mathfrak{A} + E_2 \mathfrak{A} + \dots + E_s \mathfrak{A};$$

dabei sind $E_n \mathfrak{A} \simeq \mathfrak{C}_n$ zueinander einstufig isomorphe gewöhnliche Gruppen, von denen keine zwei ein Element gemeinsam haben. Ist n die Ordnung von \mathfrak{C}_n , so ist ns die Ordnung von \mathfrak{A} . Alle Einheiten von \mathfrak{A} bilden auch eine Gruppe, – „die linke Hauptgruppe“ $\mathfrak{E} \simeq E_1 + E_2 + \dots + E_s$, deren Struktur besonders einfach ist: $E_n E_\lambda = E_n$. Die allgemeinste Untergruppe von \mathfrak{A} hat die Form: $\mathfrak{E}_1 \mathfrak{D}_n$, wo \mathfrak{D}_n irgendeine Untergruppe von \mathfrak{C}_n und $\mathfrak{E}_1 \simeq E_n + E_\lambda + E_\mu + \dots$ ist; dabei sind $E_n, E_\lambda, E_\mu, \dots$ einige von den Einheiten von \mathfrak{A} .

4. Sind A und B Elemente der Gruppe \mathfrak{A} , so hat die Gleichung

$$(4) \quad AX = B$$

dann und nur dann Lösungen, wenn A und B zur selben Einheit gehören, und zwar genau s verschiedene Lösungen, wo s die Anzahl der Hauptelemente von \mathfrak{A} ist. Ist P irgendeine Lösung von (4), so haben alle Lösungen die Gestalt: $E_1 P, E_2 P, \dots, E_s P$.

Wir wollen die Beweise aller dieser Sätze kurz skizzieren.

1. Sei E die Lösung der Gleichung $EP = P$; ist X ein beliebiges Element, so folgt nach II und IV₁: $X(EP) = (XE)P = XP$; $XE = X$. D. h. es ist E die linke Einheit für P und die rechte Einheit für alle Elemente X . Speziell für $X = E$ haben wir $E^2 = E$. Sei umgekehrt $E^2 = E$; dann ist $XE^2 = XE$, also nach II und IV₁: $XE = X$.

Sei E_l linke Einheit für alle Elemente und E beliebige rechte Einheit von \mathfrak{A} , dann ist $E_l E = E_l = E$; also ist in diesem Falle E die einzige rechte und linke Einheit von \mathfrak{A} . Ist $A_1 A = E$ und $AB = AC$, so ist $A_1 A B = A_1 A C$, $EB = EC$, $B = C$, d. h. auch IV_r ist erfüllt, \mathfrak{A} ist also eine gewöhnliche Gruppe.

Zeigen wir noch, daß, wenn $PE = P$ für ein bestimmtes Element P ist, dann auch für jedes Element X $XE = X$ wird. Nach IV₁ finden wir das Element Q so, daß $X = QP$ wird; dann ist nach II: $QPE = QP$, d. h. $XE = X$.

2. Ist $A^* = A^1$, so folgt daraus nach IV₁ auch $A^{*-1} = A^{1-1}$; daraus folgt, daß die Art von A gleich 1 sein muß.

3. Nach den eingeführten Bezeichnungen, — sind A und B Elemente aus \mathfrak{A} , so sind $E_1 A$ und $E_1 B$ aus \mathfrak{C}_1 und $E_1 A \cdot E_1 B = E_1(AB)$, also ist \mathfrak{C}_1 verallgemeinert isomorph zu \mathfrak{A} . Die Gruppe \mathfrak{C}_1 besteht aus allen Elementen aus \mathfrak{A} , die zur Einheit E_1 gehören; sie ist gewöhnlich, da sie eine einzige linke und rechte Einheit hat.

Da $E_1 E_2 = E_1$, $E_2 E_1 = E_2$ ist, so ist $\mathfrak{C}_2 = E_2 \mathfrak{A} = E_2(E_1 \mathfrak{A}) = E_2 \mathfrak{C}_1$; ebenso auch $\mathfrak{C}_1 = E_1 \mathfrak{C}_2$. Also haben \mathfrak{C}_1 und \mathfrak{C}_2 dieselbe Ordnung; überdies sind sie isomorph: dem Element $E_1 A < \mathfrak{C}_1$ entspricht das Element $E_2 A < \mathfrak{C}_2$; sie sind also einstufig isomorph. Daß \mathfrak{C}_1 und \mathfrak{C}_2 kein Element gemeinsam haben, folgt daraus, daß sie verschiedene Einheiten haben. Damit ist die Darstellung (3) bewiesen.

4. Es gehören A und AX immer zur selben Einheit; also muß auch B zu dieser Einheit gehören. Sei diese Einheit $= E_1$, also seien A und B aus der Gruppe \mathfrak{C}_1 , dann hat (4) eine und nur eine Lösung $P < \mathfrak{C}_1$. Sei nun X eine beliebige Lösung von (4); sei $X < \mathfrak{C}_x$; dann ist $E_1 X < \mathfrak{C}_1$ auch eine Lösung von (4), d. h. $E_1 X = P$; also $X = E_x X = E_x E_1 X = E_x P$.

§ 3.

Es entsteht nun die Frage: Wie wird der Isomorphismus zwischen den Gruppen \mathfrak{C}_x hergestellt? Wie bezeichnen wir die Elemente dieser Gruppen? Sei

$$\mathfrak{C}_1 = E_1 + A_1 + B_1 + C_1 + \dots;$$

dabei ist E_1 die Einheit; die anderen Elemente sind irgendwie bezeichnet. Wir haben

$$\mathfrak{C}_x = E_x \mathfrak{C}_1 = E_x E_1 + E_x A_1 + E_x B_1 + E_x C_1 + \dots \quad (x = 2, 3, \dots, s).$$

Es ist $E_\kappa E_1 = E_\kappa$; und nun *bezeichnen* wir: $E_\kappa A_1 \simeq A_\kappa$, $E_\kappa B_1 \simeq B_\kappa$, $E_\kappa C_1 \simeq C_\kappa, \dots$; dann stellt sich zwischen den Gruppen \mathfrak{G}_1 und \mathfrak{G}_κ ein Isomorphismus her, in dem den Elementen E_1, A_1, B_1, \dots die Elemente $E_\kappa, A_\kappa, B_\kappa, \dots$ entsprechen. Nun sind aber auch \mathfrak{G}_κ und \mathfrak{G}_λ isomorph, und es ist $E_\lambda \mathfrak{G}_\kappa = \mathfrak{G}_\lambda$; die Frage ist die, ob auch hier $E_\lambda A_\kappa = A_\lambda$, $E_\lambda B_\kappa = B_\lambda, \dots$ sein wird (daß $E_\lambda E_\kappa = E_\lambda$ ist, wissen wir schon). Nun in der Tat:

$$E_\lambda A_\kappa = E_\lambda (E_\kappa A_1) = (E_\lambda E_\kappa) A_1 = E_\lambda A_1 = A_\lambda;$$

ebenso auch für die anderen Elemente. Drücken wir diese Tatsache aus, indem wir sagen: Bei der oben eingeführten Bezeichnung sind je zwei der Gruppen $\mathfrak{G}_1, \mathfrak{G}_2, \dots, \mathfrak{G}_s$ gegeneinander *gleichgeordnet*.

Das Postulat IV_1 kann durch folgendes Postulat ersetzt werden:

IV'_1 . Jedes Hauptelement ist auch rechte Einheit für alle Elemente der Gruppe.

Dieses Postulat kommt in einem Postulatsystem von Huntington³⁾ vor, das zur Definition der gewöhnlichen Gruppen dient. Dieses System besteht aus den Postulaten, die wir mit I bis III bezeichnet haben, dann aus IV'_1 und noch aus einem Postulate, daß es nämlich nur ein einziges Hauptelement gibt. Unsere Gruppen erhält man also, wenn man in diesem System das letzte Postulat wegfällen läßt. Man kann leicht beweisen, daß es in jeder Gruppe mit den Postulaten I bis III wenigstens ein Hauptelement gibt⁴⁾.

Bemerkung. Führen wir statt des Postulates IV_1 das Postulat IV_2 ein, so bleiben natürlich alle obigen Sätze richtig, wenn man nur überall die rechte und linke Seite miteinander vertauscht.

§ 4.

Wir gehen jetzt zu den Gruppen über, bei denen keine Seite des Gesetzes der eindeutigen Umkehrbarkeit erfüllt ist, also nur die Postulate I bis III von § 1 gelten. Sei \mathfrak{G} eine solche endliche Gruppe und $P < \mathfrak{G}$; wir wissen schon (§ 1), daß $\mathfrak{G}P$ auch Gruppe ist und $\mathfrak{G}P < \mathfrak{G}$. Nun durchlaufe P alle Elemente von \mathfrak{G} . Unter allen Gruppen $\mathfrak{G}P$ wählen wir die (oder eine von denen), deren Ordnung die kleinste ist. Es sei dies $\mathfrak{G}X \simeq \mathfrak{H}$. Ist $A < \mathfrak{H}$, so ist $\mathfrak{H}A < \mathfrak{H}$; andererseits ist $\mathfrak{H}A = (XA)$;

³⁾ Huntington, Note on the definition of abstract groups and fields by sets of independent postulates. Transact. of the Amer. Math. Soc. 6 (1905).

⁴⁾ Vgl. Huntington loc. cit.; auch E. H. Moore, A definition of abstract groups. Transact. of the Amer. Math. Soc. 3 (1902). Frobenius (Über endliche Gruppen. Sitzungsber. der Berl. Ak. 1895) beweist dasselbe, indem er aber nicht Elemente, sondern Komplexe betrachtet.

also kann die Ordnung von $\mathfrak{A}A$ nicht kleiner sein als die Ordnung von \mathfrak{A} , d. h. $\mathfrak{A}A = \mathfrak{A}$; das zeigt, daß für $\mathfrak{A}IV_7$ gilt (vgl. § 2, Anfang).

Sei jetzt P ein beliebiges Element von \mathfrak{G} ; $\mathfrak{A}P = \mathfrak{G}(XP)$; also ist $\mathfrak{A}P$ auch Gruppe und die Ordnung von $\mathfrak{A}P$ ist gleich der Ordnung von \mathfrak{A} . Sei nun $AP \sim A_1 < \mathfrak{A}$, dann ist $\mathfrak{A}AP = \mathfrak{A}A_1 = \mathfrak{A}P = \mathfrak{A}$. Daraus folgt, daß $\mathfrak{A}P$ entweder mit \mathfrak{A} identisch ist oder *kein* Element mit \mathfrak{A} gemeinsam hat. Im letzten Falle hat $\mathfrak{A}P \sim \mathfrak{A}'$ dieselben Eigenschaften wie \mathfrak{A} ; es gilt also auch für \mathfrak{A}' das Postulat IV_7 .

Sei jetzt $P \leq \mathfrak{A}'$; es ist $\mathfrak{A}'P = \mathfrak{A}'$; auch $\mathfrak{G}P = \mathfrak{A}'$; denn ist $\mathfrak{A}' = \mathfrak{G}Y$, ^(1) nicht sicher, dass das P.) so ist $P = GY$, wo $G < \mathfrak{G}$ ist, also $\mathfrak{G}P = (\mathfrak{G}G)Y < \mathfrak{G}Y$; $\mathfrak{G}P < \mathfrak{A}'$; da aber die Ordnung von \mathfrak{A}' die kleinste ist, so ist $\mathfrak{G}P = \mathfrak{A}'$, also auch $P = \mathfrak{A}'$; das zeigt, daß es in \mathfrak{A} ein Element A gibt, so daß $\underline{AP = P}$ ist.

Wir haben (vgl. § 2, (3)) $\mathfrak{A} = \mathfrak{G}_1 + \mathfrak{G}_2 + \dots$, wo $\mathfrak{G}_1, \mathfrak{G}_2, \dots$ zu einander einstufig isomorphe gewöhnliche Gruppen sind, von denen keine zwei ein Element gemeinsam haben. Sei E_1 die Einheit für \mathfrak{G}_1 , d. h. rechte Einheit für \mathfrak{A} ; sei ferner $A < \mathfrak{G}_1$. Ebenso sei $\mathfrak{A}' = \mathfrak{G}'_1 + \mathfrak{G}'_2 + \dots$, $P < \mathfrak{G}'_1$ und E'_1 die Einheit für \mathfrak{G}'_1 , also rechte Einheit für \mathfrak{A}' . Wir haben

$$AP = P = E'_1 P, \quad \text{also} \quad (E_1 A)P = (E_1 E'_1)P = E_1 P = AP = P;$$

nun hat $\mathfrak{A}P = \mathfrak{A}'$ dieselbe Ordnung wie \mathfrak{A} ; daraus folgt, daß die Gleichung $E_1 P = AP$ nur dann bestehen kann, wenn $\underline{A = E_1}$ ist; es ist also $E_1 P = P$. Ist auch $Q < \mathfrak{G}'_1$, so ist $Q = PR$ (das Element $R < \mathfrak{G}'_1$ kann immer gefunden werden, da \mathfrak{G}'_1 eine gewöhnliche Gruppe ist); also ist auch $E_1 Q = Q$. Es ist also $\underline{E_1}$ die linke Einheit für alle Elemente von \mathfrak{G}'_1 .

Nun ist ja auch $\mathfrak{A}'A = \mathfrak{A}$; wir schließen ebenso, daß es in \mathfrak{A}' eine rechte Einheit, z. B. E'_x geben muß, die als linke Einheit für alle Elemente von \mathfrak{G}_1 erscheint. Sei

$$\mathfrak{G}_1 = E_1 + A_1 + B_1 + C_1 + \dots,$$

dann ist

$$\mathfrak{G}_1 P = P + A_1 P + B_1 P + C_1 P + \dots;$$

$$P = E_1 P; \quad E'_x P = (E'_x E_1)P = E_1 P = P;$$

nun hat ja jedes Element von \mathfrak{A}' nur eine einzige linke Einheit, und diese linke Einheit für P ist E'_1 ; es ist also $E'_x = E'_1$, d. h. $x = 1$. Also haben auch $A_1 P, B_1 P, \dots$ die linke Einheit E'_1 ; das sind also Elemente von \mathfrak{G}'_1 : also $\mathfrak{G}_1 P < \mathfrak{G}'_1$. Wir können ebenso auch umgekehrt finden, daß $\mathfrak{G}'_1 A < \mathfrak{G}_1$ ist. Da nun die Ordnung von $\mathfrak{G}_1 P$ dieselbe ist, wie die von \mathfrak{G}_1 , und die Ordnung von $\mathfrak{G}'_1 A$ dieselbe ist wie die von \mathfrak{G}'_1 , so folgt $\underline{\mathfrak{G}_1 P = \mathfrak{G}'_1}$, $\underline{\mathfrak{G}'_1 A = \mathfrak{G}_1}$; die Gruppen \mathfrak{G}_1 und \mathfrak{G}'_1 haben also dieselbe Ordnung. Also ist die Anzahl der Gruppen $\mathfrak{G}_1, \mathfrak{G}_2, \dots$ dieselbe, wie die Anzahl der Gruppen $\mathfrak{G}'_1, \mathfrak{G}'_2, \dots$.

Nun zeigen wir, daß die Gruppen \mathfrak{G}_1 und \mathfrak{G}'_1 auch einstufig isomorph sind. Sei $AP = Q < \mathfrak{G}'_1$; wir können immer in der gewöhnlichen Gruppe \mathfrak{G}'_1 ein Element R so finden, daß $Q = RP$ wird; es ist also $AP = RP$; man findet leicht, daß diese Gleichung auch für jedes $X < \mathfrak{G}'_1$ gilt:

$$(5) \quad AX = RX.$$

Nun gibt (5) eine eindeutige Zuordnung der Elemente von \mathfrak{G}_1 und \mathfrak{G}'_1 : A entspricht R . Es entspreche nun dem Element $B < \mathfrak{G}_1$ das Element $S < \mathfrak{G}'_1$, also $BP = SP < \mathfrak{G}'_1$; nun setzen wir in (5) $X = BP = SP$; dann kommt

$$A(BP) = R(SP) \quad \text{oder} \quad (AB)P = (RS)P;$$

also auch für jedes $X < \mathfrak{G}'_1$: $(AB)X = (RS)X$ und unsere Behauptung ist bewiesen.

Wir sahen, daß E_1 linke Einheit für \mathfrak{G}'_1 ist; zeigen wir, daß E_1 keine linke Einheit für $\mathfrak{G}'_2, \mathfrak{G}'_3, \dots$ ist. Sei $Q < \mathfrak{G}'_2$ und $E_1 Q = Q$; ist $X < \mathfrak{G}'_2$, so kann man $R < \mathfrak{G}'_2$ so wählen, daß $X = QR$ wird, also $E_1 X = X$. Nun ist $E'_1 E_1 = E_1$; also

$$(E'_1 E_1) X = E_1 X = E'_1 (E_1 X) = E'_1 X = X;$$

also ist E'_1 die linke Einheit für \mathfrak{G}'_2 ; da das nicht richtig ist, so ist unsere Behauptung bewiesen. Ebenso ist E'_1 die linke Einheit für \mathfrak{G}_1 , aber nicht für $\mathfrak{G}_2, \mathfrak{G}_3, \dots$.

Nun ändern wir unsere Bezeichnung. Mit $\mathfrak{U}_1, \mathfrak{U}_2, \dots, \mathfrak{U}_r$ bezeichnen wir alle verschiedene Gruppen $\mathfrak{G} X$ von der kleinsten Ordnung; wir sahen, daß sie Linksgruppen sind, von denen je zwei einander einstufig isomorph sind und kein Element gemeinsam haben. Sei ferner

$$\mathfrak{U}_\kappa = \mathfrak{G}_{\kappa 1} + \mathfrak{G}_{\kappa 2} + \dots + \mathfrak{G}_{\kappa s} \quad (\kappa = 1, 2, \dots, r);$$

dabei sind $\mathfrak{G}_{\kappa \lambda}$ zueinander einstufig isomorphe gewöhnliche Gruppen, von denen je zwei kein Element gemeinsam haben; deren Gesamtanzahl ist $r \cdot s$. Die Elemente von $\mathfrak{G}_{\kappa \lambda}$ bezeichnen wir mit $A_{\kappa \lambda}, B_{\kappa \lambda}, \dots$; $E_{\kappa \lambda}$ sei die Einheit von $\mathfrak{G}_{\kappa \lambda}$; also hat \mathfrak{U}_κ s rechte Einheiten $E_{\kappa 1}, E_{\kappa 2}, \dots, E_{\kappa s}$. Nehmen wir zwei Gruppen \mathfrak{U}_μ und \mathfrak{U}_ν ; wir sahen daß es in \mathfrak{U}_μ eine und nur eine Gruppe $\mathfrak{G}_{\mu \nu}$ gibt, die $E_{\kappa \lambda}$ zur linken Einheit hat; wählen wir die Bezeichnung so, daß diese Gruppe $\mathfrak{G}_{\mu \lambda}$ sei; dann ist auch, wie wir sahen, $E_{\mu \lambda}$ die linke Einheit für die Gruppe $\mathfrak{G}_{\kappa \lambda}$; das soll gelten für $\mu = 1, 2, \dots, r$. Sei nun ϱ auch ein Index aus der Reihe $1, 2, \dots, r$; wir haben

$$E_{\mu \lambda} E_{\varrho \lambda} = E_{\mu \lambda} (E_{\kappa \lambda} E_{\varrho \lambda}) = (E_{\mu \lambda} E_{\kappa \lambda}) E_{\varrho \lambda} = E_{\kappa \lambda} E_{\varrho \lambda} = E_{\varrho \lambda};$$

daraus folgt leicht, daß $E_{\mu \lambda}$ auch linke Einheit für $\mathfrak{G}_{\varrho \lambda}$ (und auch umgekehrt $E_{\varrho \lambda}$ linke Einheit für $\mathfrak{G}_{\mu \lambda}$) ist. Es folgt also bei dieser Bezeichnung, daß alle Gruppen $\mathfrak{G}_{1 \lambda}, \mathfrak{G}_{2 \lambda}, \dots, \mathfrak{G}_{r \lambda}$ (mit einem und demselben

zweiten Index λ) gemeinsame linke Einheiten $E_{1\lambda}, E_{2\lambda}, \dots, E_{r\lambda}$ haben. Es ist auch, wie wir sahen, $A_{\alpha\lambda} B_{\mu\lambda} = C_{\mu\lambda}$. Bezeichnen wir nun

$$\mathfrak{B}_\lambda \simeq \mathfrak{C}_{1\lambda} + \mathfrak{C}_{2\lambda} + \dots + \mathfrak{C}_{r\lambda} \quad (\lambda = 1, 2, \dots, s),$$

so sind die \mathfrak{B}_λ Rechtsgruppen, da alle ihre Hauptelemente $E_{1\lambda}, E_{2\lambda}, \dots, E_{r\lambda}$ zugleich auch linke Einheiten sind (vgl. § 3).

Bezeichnen wir noch: $\mathfrak{K} \simeq \mathfrak{A}_1 + \mathfrak{A}_2 + \dots + \mathfrak{A}_r$, so ist \mathfrak{K} auch eine Gruppe, die wir den Kern der Gruppe \mathfrak{G} oder (allein genommen) die Kerngruppe nennen; das ist dieselbe Gruppe, die ich in der Einleitung angekündigt habe und die den dort geschilderten Bau hat (vgl. Schema I in der Einleitung). Ist n die Ordnung von $\mathfrak{C}_{\alpha\lambda}$, so ist $n \cdot r \cdot s$ die Ordnung von \mathfrak{K} . Es ist $\mathfrak{K}^2 = \mathfrak{K}$.

Keine zwei Gruppen \mathfrak{B}_λ und \mathfrak{B}_μ haben ein Element gemeinsam; der Durchschnitt von \mathfrak{A}_α und \mathfrak{B}_λ ist $\mathfrak{C}_{\alpha\lambda}$.

Wir werden jetzt zeigen, daß die Gruppen \mathfrak{B}_λ die Form $X\mathfrak{G}$ ($X < \mathfrak{G}$) haben; und zwar sind \mathfrak{B}_λ Gruppen dieser Form von möglichst kleiner Ordnung.

Es ist $\mathfrak{G} E_{\alpha\lambda} = \mathfrak{A}_\alpha$; $E_{\alpha\lambda} \mathfrak{A}_\alpha = E_{\alpha\lambda} \mathfrak{G} E_{\alpha\lambda} = \mathfrak{C}_{\alpha\lambda}$; sei $E_{\alpha\lambda} \mathfrak{G} \simeq \mathfrak{B}'_{\alpha\lambda}$; da $\mathfrak{C}_{\alpha\lambda} < \mathfrak{G}$ und $E_{\alpha\lambda} \mathfrak{C}_{\alpha\lambda} = \mathfrak{C}_{\alpha\lambda}$ ist, so ist $\mathfrak{C}_{\alpha\lambda} < \mathfrak{B}'_{\alpha\lambda}$; also auch $E_{\alpha\lambda} < \mathfrak{B}'_{\alpha\lambda}$.

Nun gilt für $\mathfrak{B}'_{\alpha\lambda}$ das Postulat IV_r. Sei nämlich:

$$E_{\alpha\lambda} G \cdot E_{\alpha\lambda} H = E_{\alpha\lambda} G \cdot E_{\alpha\lambda} H_1,$$

wo G, H, H_1 Elemente aus \mathfrak{G} sind. Oder

$$(E_{\alpha\lambda} G E_{\alpha\lambda}) H = (E_{\alpha\lambda} G E_{\alpha\lambda}) H_1;$$

es ist $E_{\alpha\lambda} G E_{\alpha\lambda} \simeq A_{\alpha\lambda} < \mathfrak{C}_{\alpha\lambda}$; also

$$(6) \quad A_{\alpha\lambda} H = A_{\alpha\lambda} H_1;$$

sei $A_{\alpha\lambda}^{-1}$ das inverse zu dem Element $A_{\alpha\lambda}$ in der gewöhnlichen Gruppe $\mathfrak{C}_{\alpha\lambda}$; multiplizieren wir (6) links mit $A_{\alpha\lambda}^{-1}$, so bekommen wir

$$E_{\alpha\lambda} H = E_{\alpha\lambda} H_1, \quad \text{w. z. b. w.}$$

Da $\mathfrak{B}_\lambda < \mathfrak{G}$ und $E_{\alpha\lambda} \mathfrak{B}_\lambda = \mathfrak{B}_\lambda$ ist, so ist $\mathfrak{B}'_{\alpha\lambda} > \mathfrak{B}_\lambda$.

Ist G ein beliebiges Element von \mathfrak{G} , so ist die Ordnung von $(G E_{\alpha\lambda}) \mathfrak{G} \leq$ als die Ordnung von $E_{\alpha\lambda} \mathfrak{G}$; die Ordnung von $(E_{\alpha\lambda} G)(E_{\alpha\lambda} \mathfrak{G}) \leq$ als die Ordnung von $G E_{\alpha\lambda} \mathfrak{G}$; nun ist $(E_{\alpha\lambda} G)(E_{\alpha\lambda} \mathfrak{G}) = E_{\alpha\lambda} \mathfrak{G}$, da $E_{\alpha\lambda} \mathfrak{G} = \mathfrak{B}'_{\alpha\lambda}$ Rechtsgruppe ist. Also ist die Ordnung von $G E_{\alpha\lambda} \mathfrak{G}$ gleich der Ordnung von $E_{\alpha\lambda} \mathfrak{G}$. Andererseits ist die Ordnung von $G E_{\alpha\lambda} \mathfrak{G} \leq$ als die Ordnung von $G \mathfrak{G}$. Sei G so gewählt, daß die Ordnung von $G \mathfrak{G}$ die kleinste ist unter den Ordnungen aller Gruppen von der Form $X\mathfrak{G}$. Dann ist also die Ordnung von $G \mathfrak{G}$ gleich der Ordnung von $E_{\alpha\lambda} \mathfrak{G}$ oder von $\mathfrak{B}'_{\alpha\lambda}$; diese Gruppen sind also Gruppen der kleinsten Ordnung von dieser Art.

Nun ist offenbar $E_{\kappa\lambda}$ die linke Einheit für $\mathfrak{B}_{\kappa\lambda} = E_{\kappa\lambda}\mathfrak{G}$; da aber $E_{\kappa\lambda}$ für kein Element von \mathfrak{B}_μ ($\mu \neq \lambda$) linke Einheit ist, so haben $\mathfrak{B}'_{\kappa\lambda}$ und \mathfrak{B}_μ kein einziges Element gemeinsam. Ebenso wie für die \mathfrak{A}_κ beweisen wir auch hier, daß $\mathfrak{B}'_{\kappa\lambda}$ und $\mathfrak{B}'_{\mu\lambda} \simeq E_{\mu\lambda}\mathfrak{G}$ entweder identisch sind oder kein Element gemeinsam haben; da aber $\mathfrak{B}'_{\mu\lambda}$ ebenso wie $\mathfrak{B}_{\kappa\lambda}$ die Gruppe \mathfrak{B}_λ enthält, so muß $\mathfrak{B}'_{\mu\lambda} \equiv \mathfrak{B}'_{\kappa\lambda}$ sein. Es hängt daher $\mathfrak{B}'_{\kappa\lambda}$ nicht von ihrem ersten Index ab, und man kann einfach bezeichnen: $\mathfrak{B}'_{\kappa\lambda} \simeq \mathfrak{B}'_\lambda$. Ferner ist, wie wir schon sahen: $\mathfrak{B}'_\lambda E_{\kappa\lambda} = \mathfrak{C}_{\kappa\lambda}$. Also sind $\mathfrak{C}_{1\lambda}, \mathfrak{C}_{2\lambda}, \dots, \mathfrak{C}_{r\lambda}$ die gewöhnlichen Gruppen, aus denen \mathfrak{B}'_λ besteht (vgl. § 2, (3)); wir können aber noch nicht behaupten, daß dies *alle* solchen Gruppen sind; vielleicht gibt es deren noch weitere: $\mathfrak{C}_{r+1,\lambda}, \mathfrak{C}_{r+2,\lambda}, \dots, \mathfrak{C}_{r',\lambda}$; $r' \geq r$. Ebenso ist die Anzahl s' der Gruppen $\mathfrak{B}'_\lambda \geq s$. Nun können wir mit den Gruppen $G\mathfrak{G}$, d. h. mit \mathfrak{B}'_λ anfangen und nachher zu den Gruppen \mathfrak{A}_κ übergehen; dann würden wir ebenso finden: $r \geq r', s \geq s'$; daraus folgt, daß $r = r', s = s'$, also auch $\mathfrak{B}'_\lambda \equiv \mathfrak{B}_\lambda$ ist.

Ist $A_{\mu\nu} < \mathfrak{C}_{\mu\nu} < \mathfrak{A}_\mu$, so ist $\mathfrak{A}_\kappa A_{\mu\nu} = \mathfrak{A}_\mu$, also $\mathfrak{C}_{\kappa\lambda} A_{\mu\nu} < \mathfrak{A}_\mu$; also auch $\mathfrak{C}_{\kappa\lambda} \mathfrak{C}_{\mu\nu} < \mathfrak{A}_\mu$. Ebenso beweisen wir, daß $\mathfrak{C}_{\kappa\lambda} \mathfrak{C}_{\mu\nu} < \mathfrak{B}_\lambda$ ist; also gehört $\mathfrak{C}_{\kappa\lambda} \mathfrak{C}_{\mu\nu}$ dem Durchschnitte $\mathfrak{C}_{\mu\lambda}$ von \mathfrak{A}_μ und \mathfrak{B}_λ ; da aber die Ordnungen von $\mathfrak{C}_{\kappa\lambda} \mathfrak{C}_{\mu\nu}$ und von $\mathfrak{C}_{\mu\lambda}$ gleich sind, so ist:

$$(7) \quad \mathfrak{C}_{\kappa\lambda} \mathfrak{C}_{\mu\nu} = \mathfrak{C}_{\mu\lambda};$$

also

$$(7') \quad A_{\kappa\lambda} B_{\mu\nu} = C_{\mu\lambda} < \mathfrak{C}_{\mu\lambda}.$$

Aus (7) und (7') folgt leicht, daß die Gleichung $A_{\kappa\lambda} X = B_{\rho\sigma}$ dann und nur dann auflösbar (in \mathfrak{R}) ist, wenn $\sigma = \lambda$ ist; in diesem Falle hat sie s Lösungen in \mathfrak{R} (je eine aus den Gruppen $\mathfrak{C}_{\rho 1}, \mathfrak{C}_{\rho 2}, \dots, \mathfrak{C}_{\rho s}$); ebenso ist die Gleichung $Y A_{\mu\nu} = B_{\rho\sigma}$ dann und nur dann in \mathfrak{R} auflösbar, wenn $\rho = \mu$ ist; in diesem Falle hat sie r Lösungen in \mathfrak{R} (je eine aus den Gruppen $\mathfrak{C}_{1\sigma}, \mathfrak{C}_{2\sigma}, \dots, \mathfrak{C}_{r\sigma}$).

Die Formel (7) zeigt, daß die Gruppen $\mathfrak{C}_{\kappa\lambda}$, als Elemente betrachtet, auch eine Gruppe bilden, und zwar besteht diese Gruppe aus lauter Hauptelementen.

Erwähnen wir noch einige Formeln, die leicht zu beweisen sind. Es ist

$$(8) \quad \mathfrak{A}_\kappa \mathfrak{B}_\lambda = \mathfrak{R};$$

$$(9) \quad \mathfrak{B}_\lambda \mathfrak{A}_\kappa = \mathfrak{C}_{\kappa\lambda}.$$

Ist $X\mathfrak{G} = \mathfrak{B}_\lambda$, $\mathfrak{G}Y = \mathfrak{A}_\kappa$, so ist

$$(10) \quad X\mathfrak{G}Y = \mathfrak{C}_{\kappa\lambda},$$

$$(11) \quad \mathfrak{G}X\mathfrak{G} = \mathfrak{G}Y\mathfrak{G} = \mathfrak{R}.$$

↳ ebenso $\mathfrak{A}_\kappa \mathfrak{B}_\lambda = \mathfrak{R}$

Also, wenn wir die Resultate dieses Paragraphen zusammenfassen:

1. Jede Gruppe \mathfrak{G} der von uns betrachteten Art hat eine besondere Untergruppe \mathfrak{K} — die „Kerngruppe“ —, die die in der Einleitung erklärte Struktur hat.

2. $\mathfrak{A}_x [\mathfrak{B}_\lambda]$ sind dabei Gruppen der Form $\mathfrak{G}Y [X\mathfrak{G}]$, und zwar von einer möglichst kleinen Ordnung.

3. Es gelten die Formeln (7) bis (11).

4. Umgekehrt, ist $X < \mathfrak{G}$ und ist die Gruppe $\mathfrak{G}X [X\mathfrak{G}]$ Linksgruppe [Rechtsgruppe], so ist $\mathfrak{G}X$ eine der Gruppen $\mathfrak{A}_x [\mathfrak{B}_\lambda]$.

Wir wollen 4. beweisen. Für jedes $G < \mathfrak{G}$ ist $\mathfrak{G}X \cdot GX = \mathfrak{G}X$. Sei G so gewählt, daß $\mathfrak{G}G = \mathfrak{A}_x$ ist; dann ist $\mathfrak{G}XG < \mathfrak{A}_x$; da aber die Ordnung von \mathfrak{A}_x die kleinste ist, so ist $\mathfrak{G}XG = \mathfrak{A}_x = \mathfrak{G}G$; $\mathfrak{G}X \cdot GX = \mathfrak{A}_x X$; also ist die Ordnung von $\mathfrak{G}X \cdot GX$, d. h. von $\mathfrak{G}X$ gleich der Ordnung von \mathfrak{A}_x ; es ist also $\mathfrak{G}X$ eine der Gruppen \mathfrak{A}_x .

§ 5.

Nun werden wir die Struktur der Kerngruppe \mathfrak{K} genauer untersuchen. Wir haben $r \cdot s$ zueinander isomorphe gewöhnliche Gruppen $\mathfrak{C}_{x\lambda}$. Nehmen wir eine dieser Gruppen ganz abstrakt an und bezeichnen wir sie als solche mit

$$\mathfrak{C} = E + A + B + C + \dots;$$

E ist dabei die Einheit; andere Elemente sind irgendwie bezeichnet. Wir bezeichnen ferner

$$\mathfrak{C}_{x\lambda} = E_{x\lambda} + A_{x\lambda} + B_{x\lambda} + C_{x\lambda} + \dots;$$

es möge bei dem Isomorphismus zwischen \mathfrak{C} und $\mathfrak{C}_{x\lambda}$ $A_{x\lambda}$ dem A , $B_{x\lambda}$ dem B usw. entsprechen; $E_{x\lambda}$ entspricht gewiß dem E . Dagegen können $A_{x\lambda}$, $B_{x\lambda}$, ... auf mehrere Arten gewählt werden, da $\mathfrak{C}_{x\lambda}$ immer Automorphismen hat; wir müssen sie daher genauer definieren.

1. Wir fangen mit \mathfrak{C}_{11} an und wählen *auf irgendeine Weise* die Elemente $A_{11}, B_{11}, C_{11}, \dots$ so, daß sie in dem Isomorphismus zwischen \mathfrak{C}_{11} und \mathfrak{C} den Elementen A, B, C, \dots von \mathfrak{C} entsprechen.

2. Ferner *bezeichnen* wir:

$$E_{1\lambda} A_{11} \simeq A_{1\lambda}, \quad E_{1\lambda} B_{11} \simeq B_{1\lambda}, \quad \dots \quad (\lambda = 2, 3, \dots, s)$$

(die Gleichung $E_{1\lambda} E_{11} = E_{1\lambda}$ ist von selbst erfüllt); dann sind (nach § 3) alle die Gruppen $\mathfrak{C}_{11}, \mathfrak{C}_{12}, \dots, \mathfrak{C}_{1s}$ gegeneinander gleichgeordnet.

3. Ferner *bezeichnen* wir:

$$A_{11} E_{x1} \simeq A_{x\lambda}, \quad B_{11} E_{x1} \simeq B_{x\lambda}, \quad \dots \quad (x = 2, 3, \dots, r)$$

(es ist von selbst $E_{11} E_{\kappa 1} = E_{\kappa 1}$); dann sind $\mathfrak{C}_{11}, \mathfrak{C}_{21}, \dots, \mathfrak{C}_{r1}$ gegeneinander gleichgeordnet.

4. Endlich bezeichnen wir:

$$E_{\kappa \lambda} A_{\kappa 1} \simeq A_{\kappa \lambda}, \quad E_{\kappa \lambda} B_{\kappa 1} \simeq B_{\kappa \lambda}, \quad \dots \quad (\kappa = 2, \dots, r; \lambda = 2, \dots, s);$$

dann sind $\mathfrak{C}_{\kappa 1}, \mathfrak{C}_{\kappa 2}, \dots, \mathfrak{C}_{\kappa s}$ gegeneinander gleichgeordnet.

Aus 2., 3., 4. folgt leicht:

$$(12) \quad \underline{A_{\kappa \lambda} B_{\kappa \nu} = A_{\kappa \lambda} B_{\nu \lambda}}; \quad \underline{A_{\kappa 1} B_{\mu 1} = A_{\mu 1} B_{\mu 1}}.$$

Das zeigt, daß man bei der Multiplikation der Elemente aus \mathfrak{A}_{κ} ($\kappa = 1, 2, \dots, r$) oder aus \mathfrak{B}_1 einfach die Cayleysche Tafel der Gruppe \mathfrak{C} benutzen kann.

Jetzt sind alle Elemente aller Gruppen $\mathfrak{C}_{\kappa \lambda}$ bestimmt. Nun entsteht die Frage:

5. Sind die Gruppen $\mathfrak{C}_{1\lambda}, \mathfrak{C}_{2\lambda}, \dots, \mathfrak{C}_{r\lambda}$ auch gegeneinander gleichgeordnet? Oder anders gesagt: Kann man auch bei der Multiplikation der Elemente von \mathfrak{B}_{λ} ($\lambda > 1$) die Tafel von \mathfrak{C} einfach benutzen? Ist auch hier $A_{\kappa \lambda} B_{\mu \lambda} = A_{\mu \lambda} B_{\kappa \lambda}$?

Das ist im allgemeinen nicht der Fall; es ist

$$A_{1\lambda} E_{\kappa \lambda} = A'_{\kappa \lambda} \neq A_{\kappa \lambda}, \quad B_{1\lambda} E_{\kappa \lambda} = B'_{\kappa \lambda} \neq B_{\kappa \lambda}, \quad \dots;$$

ebenso auch

$$A_{\mu \lambda} E_{\kappa \lambda} \neq A_{\kappa \lambda}; \quad \text{nur } E_{\mu \lambda} E_{\kappa \lambda} = E_{\kappa \lambda} \text{ bleibt bestehen.}$$

Ferner

$$A_{1\lambda} B_{\kappa \lambda} = A_{1\lambda} (E_{\kappa \lambda} B_{\kappa \lambda}) = (A_{1\lambda} E_{\kappa \lambda}) B_{\kappa \lambda} = A'_{\kappa \lambda} B_{\kappa \lambda}.$$

Ist $A_{\kappa \lambda} = \bar{A}_{1\lambda} E_{\kappa \lambda}$, $\bar{A}_{1\lambda} E_{\mu \lambda} = A''_{\mu \lambda}$, so ist

$$A_{\kappa \lambda} E_{\mu \lambda} = (\bar{A}_{1\lambda} E_{\kappa \lambda}) E_{\mu \lambda} = \bar{A}_{1\lambda} E_{\mu \lambda} = A''_{\mu \lambda};$$

also

$$(13) \quad A_{\kappa \lambda} B_{\mu \lambda} = A_{\kappa \lambda} E_{\mu \lambda} B_{\mu \lambda} = A''_{\mu \lambda} B_{\mu \lambda}.$$

Es wird also auch in diesem Falle die Tafel von \mathfrak{C} benutzt, nur müssen darin die Zeilen vorher permutiert werden. Das geschieht mittelst der Substitution $A \simeq \begin{pmatrix} E & A & B & \dots \\ E & A'' & B'' & \dots \end{pmatrix}$, welche einen Automorphismus der Gruppe \mathfrak{C} darstellt, der dem durch die Gleichung $\mathfrak{C}_{\kappa \lambda} E_{\mu \lambda} = \mathfrak{C}_{\mu \lambda}$ ausgedrückten Isomorphismus der Gruppen $\mathfrak{C}_{\kappa \lambda}$ und $\mathfrak{C}_{\mu \lambda}$ entspricht. Ebenso sei $\Phi = \begin{pmatrix} E & A & B & \dots \\ E & A' & B' & \dots \end{pmatrix}$ der Automorphismus von \mathfrak{C} , der dem Isomorphismus

$\mathfrak{C}_{1\lambda} E_{\kappa\lambda} = \mathfrak{C}_{\kappa\lambda}$ entspricht, und Ψ der dem Isomorphismus $\mathfrak{C}_{1\lambda} E_{\mu\lambda} = \mathfrak{C}_{\mu\lambda}$ entsprechende Automorphismus von \mathfrak{C} . Dann ist

$$(14) \quad A = \Phi^{-1}\Psi.$$

6. Es bleibt uns noch übrig die Produkte $A_{\kappa\lambda} B_{\mu\nu}$ ($\kappa \neq \mu$, $\lambda \neq \nu$) zu ermitteln. Sei $E_{\kappa\lambda} E_{\mu\nu} = X_{\mu\lambda}$ (nach (7') gehört dieses Produkt der Gruppe $\mathfrak{C}_{\mu\lambda}$ an); dann haben wir:

$$(15) \quad A_{\kappa\lambda} B_{\mu\nu} = (A_{\kappa\lambda} E_{\kappa\lambda})(E_{\mu\nu} B_{\mu\nu}) = A_{\kappa\lambda} X_{\mu\lambda} B_{\mu\nu};$$

wenn man also $X_{\mu\lambda}$ kennt, kann man $A_{\kappa\lambda} B_{\mu\nu}$ auf zweierlei Weise berechnen nach 4. und 5. Nun sind die Elemente $X_{\mu\lambda}$ und die Substitutionen A voneinander abhängig, wie wir jetzt sehen werden.

Sei $E_{11} E_{\kappa\lambda} \simeq X_{\kappa 1}$ ($\kappa \neq 1$, $\lambda \neq 1$), $A_{11} < \mathfrak{C}_{11}$ irgendein Element, $A_{1\lambda} E_{\kappa\lambda} \simeq A'_{\kappa\lambda}$, $A_{11} X_{\kappa 1} \simeq Y_{\kappa 1}$, $X_{\kappa 1} A'_{\kappa\lambda} \simeq Y'_{\kappa 1}$. Wir haben einerseits

$$A_{11} E_{\kappa\lambda} = A_{11} X_{\kappa 1} E_{\kappa\lambda} = (A_{11} X_{\kappa 1}) E_{\kappa\lambda} = Y_{\kappa 1} E_{\kappa\lambda} = Y_{\kappa 1};$$

andererseits

$$\begin{aligned} A_{11} E_{\kappa\lambda} &= (E_{11} A_{1\lambda}) E_{\kappa\lambda} = E_{11} (A_{1\lambda} E_{\kappa\lambda}) \\ &= E_{11} A'_{\kappa\lambda} = E_{11} (X_{\kappa 1} A'_{\kappa\lambda}) = E_{11} Y'_{\kappa 1} = Y'_{\kappa 1}; \end{aligned}$$

also

$$Y_{\kappa 1} = Y'_{\kappa 1} \quad \text{oder} \quad X_{\kappa 1} A'_{\kappa\lambda} = A_{11} X_{\kappa 1};$$

und da beide Produkte nach der Tafel von \mathfrak{C} gefunden werden, so besteht auch zwischen den Elementen von \mathfrak{C} die Relation

$$X A' = A X;$$

$$(16) \quad \underline{A' = X^{-1} A X}; \quad \text{also auch} \quad A'_{\kappa\lambda} = X_{\kappa\lambda}^{-1} A_{\kappa\lambda} X_{\kappa\lambda};$$

also

$$(16') \quad \Phi = \begin{pmatrix} Y \\ X^{-1} Y X \end{pmatrix}.$$

Also stellen Φ und ebenso auch Ψ und A *innere* Automorphismen der Gruppe \mathfrak{C} dar.

Bezeichnen wir noch $E_{11} E_{\mu\nu} \simeq Y_{\mu 1}$, so haben wir

$$(17) \quad \underline{A_{\kappa\lambda} B_{\mu\lambda} = (X_{\mu\lambda}^{-1} Y_{\mu\lambda})^{-1} A_{\mu\lambda} (X_{\mu\lambda}^{-1} Y_{\mu\lambda}) B_{\mu\lambda}}.$$

Wir ermitteln ferner die Abhängigkeiten, die zwischen den Produkten $E_{\kappa\lambda} E_{\mu\nu}$ bestehen. Wir haben

$$(E_{\kappa\lambda} E_{\mu\sigma})(E_{\kappa\sigma} E_{\mu\nu}) = E_{\kappa\lambda}(E_{\mu\sigma} E_{\kappa\sigma}) E_{\mu\nu} = (E_{\kappa\lambda} E_{\kappa\sigma}) E_{\mu\nu} = E_{\kappa\lambda} E_{\mu\nu};$$

$$(E_{\kappa\lambda} E_{\rho\nu})(E_{\rho\lambda} E_{\mu\nu}) = E_{\kappa\lambda}(E_{\rho\nu} E_{\rho\lambda}) E_{\mu\nu} = E_{\kappa\lambda}(E_{\rho\nu} E_{\mu\nu}) = E_{\kappa\lambda} E_{\mu\nu}.$$

Es ist also

$$(18) \quad E_{\kappa\lambda} E_{\mu\nu} = (E_{\kappa\lambda} E_{\mu\sigma})(E_{\kappa\sigma} E_{\mu\nu});$$

$$(19) \quad E_{\kappa\lambda} E_{\mu\nu} = (E_{\kappa\lambda} E_{\varrho\nu})(E_{\varrho\lambda} E_{\mu\nu}).$$

Erläutern wir den Sinn dieser Formeln. Es sei uns die Tafel der Einheitsprodukte gegeben; betrachten wir einen Teil dieser Tafel.

Aus (18) und (19) folgt

$$X_{\mu\lambda} = K_{\mu\lambda} L_{\mu\sigma} = M_{\varrho\lambda} N_{\mu\lambda};$$

diese beiden Produkte können wir nach (12) und (17) finden, wenn $K_{\mu\lambda}$, $L_{\mu\sigma}$, $M_{\varrho\lambda}$, $N_{\mu\lambda}$ bekannt sind. Aus dem Schema II sieht man, daß die Elemente $K_{\mu\lambda}$ und $L_{\mu\sigma}$ (ebenso auch $M_{\varrho\lambda}$ und $N_{\mu\lambda}$) in der Zeile und in der Kolonne zu finden sind, auf deren Durchschnitt sich das gesuchte Element $X_{\mu\lambda}$ befindet; dabei ist der

	$E_{\mu\nu}$	$E_{\mu\sigma}$	$E_{\varrho\nu}$
$E_{\varrho\lambda}$	⋮	$N_{\mu\lambda}$	⋮
⋮	⋮	⋮	⋮
$E_{\kappa\lambda}$	⋮	$X_{\mu\lambda}$	⋮
⋮	⋮	⋮	⋮
$E_{\kappa\sigma}$	⋮	$L_{\mu\sigma}$	⋮

Schema II.

erste Faktor in der Zeile, der zweite in der Kolonne zu suchen. Beim Produkte $K_{\mu\lambda} L_{\mu\sigma}$ sind die ersten Indizes dem ersten Index von $X_{\mu\lambda}$ gleich; bei $M_{\varrho\lambda} N_{\mu\lambda}$ gilt dasselbe für die zweiten Indizes.

Setzen wir in (18) $\nu = \lambda$, in (19) $\mu = \kappa$ und ändern ein wenig die Bezeichnungen ab, so bekommen wir

$$(20) \quad (E_{\kappa\lambda} E_{\mu\nu})(E_{\kappa\nu} E_{\mu\lambda}) = E_{\mu\lambda};$$

$$(21) \quad (E_{\kappa\lambda} E_{\mu\nu})(E_{\mu\lambda} E_{\kappa\nu}) = E_{\kappa\lambda};$$

und daraus folgt leicht

$$(22) \quad (E_{\mu\nu} E_{\kappa\lambda})(E_{\mu\lambda} E_{\kappa\nu}) = E_{\kappa\nu}.$$

Diese Formeln gestatten drei der Produkte $E_{\kappa\lambda} E_{\mu\nu}$ usw. zu berechnen, wenn uns das vierte Produkt bekannt ist.

Seien uns jetzt alle Produkte $E_{11} E_{\kappa\lambda}$ ($\kappa = 2, \dots, r$; $\lambda = 2, \dots, s$) bekannt, d. h. alle die Elemente der ersten Zeile der Tafel der Einheitsprodukte. Wir werden nun zeigen, daß man in diesem Falle auch alle Einheitsprodukte $E_{\kappa\lambda} E_{\mu\nu}$ berechnen kann. Nach (20), (21), (22) können wir alle $E_{\kappa\lambda} E_{11}$, $E_{\kappa 1} E_{1\lambda}$, $E_{1\lambda} E_{\kappa 1}$ berechnen. Setzen wir ferner $\sigma = 1$ in (18) und $\varrho = 1$ in (19), so ist

$$(23) \quad E_{\kappa\lambda} E_{\mu\nu} = (E_{\kappa\lambda} E_{\mu 1})(E_{\kappa 1} E_{\mu\nu});$$

$$(24) \quad E_{\kappa\lambda} E_{\mu\nu} = (E_{\kappa\lambda} E_{1\nu})(E_{1\lambda} E_{\mu\lambda});$$

nun setzen wir in (23) $\kappa = 1$ und in (24) $\lambda = 1$:

$$(25) \quad E_{1\lambda} E_{\mu\nu} = (E_{1\lambda} E_{\mu 1})(E_{11} E_{\mu\nu});$$

$$(26) \quad E_{\kappa 1} E_{\mu\nu} = (E_{\kappa 1} E_{1\nu})(E_{11} E_{\mu\nu}).$$

Nach (25), (26) können wir alle $E_{1\lambda} E_{\mu\nu}$ und $E_{\kappa 1} E_{\mu\nu}$ finden, folglich nach (20) bis (22) auch alle $E_{\mu\nu} E_{1\lambda}$, $E_{\mu\nu} E_{\kappa 1}$, und nachher nach (23) oder (24) auch alle $E_{\kappa\lambda} E_{\mu\nu}$. Sind uns aber alle $E_{\kappa\lambda} E_{\mu\nu}$ bekannt, so können wir nach (16), (16') und (14) auch alle Substitutionen A ermitteln, und darauf alle die Produkte $A_{\kappa\lambda} B_{\mu\nu}$ (nach (12), (17) und (15)); dadurch ist also die ganze Gruppe \mathfrak{S} vollständig bestimmt.

Damit ist also der Satz II der Einleitung vollständig bewiesen.

Bemerkung. In den vorhergehenden Betrachtungen spielt die Gruppe \mathfrak{S}_{11} eine ausgesonderte Rolle; das ist natürlich nicht wesentlich: statt \mathfrak{S}_{11} können wir jede der Gruppen $\mathfrak{S}_{\kappa\lambda}$ als „die erste“ nehmen; es kommt nur auf eine gewisse Vertauschung der Bezeichnungen an.

§ 6.

Wir gehen jetzt zur Behandlung der in der Einleitung, III., aufgestellten Frage über; dazu müssen wir einige Eigenschaften der dort eingeführten verallgemeinerten Substitutionen kennen.

Bei einer solchen Substitution kann es vorkommen, daß mehrere Symbole in ein und dasselbe Symbol übergeführt werden; solche Symbole nennen wir *konjugiert*. Sind in einer Substitution k Symbole untereinander konjugiert, so bezeichnen wir diese Tatsache als $k-1$ *Konjugationen*. Die Anzahl der Konjugationen wächst gleichzeitig mit der Verminderung der Anzahl der verschiedenen Symbole in der unteren Zeile der Substitution.

Ist $AB = C$, so hat die untere Zeile der Substitution C keine anderen Symbole als die der unteren Zeile von B , und ihre Anzahl ist nicht größer als die Anzahl der verschiedenen Symbole der unteren Zeile von A (und auch von B). Die Anzahl der Konjugationen in C kann dagegen nicht kleiner sein als diese Anzahl in A oder in B ; zwei Symbole, die in A konjugiert sind, sind es auch in C (aber nicht umgekehrt).

Die Gleichung $AX = B$ hat dann und nur dann Auflösungen (im allgemeinen mehrere), wenn alle Konjugationen von A auch in B vorkommen (außer diesen kann B auch andere Konjugationen haben). Die Gleichung $YA = B$ hat dann und nur dann Auflösungen (im allgemeinen mehrere), wenn in der unteren Zeile von B keine anderen Symbole vorkommen als die der unteren Zeile von A (doch brauchen nicht alle Symbole von A auch in B vorzukommen). Daraus folgt:

Damit eine Substitutionsgruppe \mathfrak{G} Rechtsgruppe sei, ist notwendig und hinreichend, daß alle Substitutionen dieser Gruppe dieselben Konjugationen haben. Damit eine Substitutionsgruppe \mathfrak{G} Linksgruppe sei, ist notwendig und hinreichend, daß in den unteren Zeilen aller Substitutionen \mathfrak{G} dieselben Symbole stehen. Ist \mathfrak{G} eine gewöhnliche Gruppe, so müssen natürlich beide Bedingungen gleichzeitig erfüllt sein (sie sind auch hinreichend).

Sei jetzt \mathfrak{G} eine beliebige Substitutionsgruppe, \mathfrak{K} ihr Kern. Aus dem Vorhergehenden folgt leicht, daß die Anzahl der Symbole in der unteren Zeile *aller* Substitutionen von \mathfrak{K} dieselbe ist, und zwar ist diese Anzahl *die kleinste*, die überhaupt in den Substitutionen von \mathfrak{G} vorkommt. Ebenso ist die Anzahl der Konjugationen in allen Substitutionen von \mathfrak{K} dieselbe, nämlich die *größte*, die überhaupt in den Substitutionen von \mathfrak{G} vorkommt.

Indem wir jetzt zur Darstellung abstrakter Kerngruppen als Gruppen unserer Substitutionen übergehen, schlagen wir folgenden Weg ein: die abstrakte Gruppe \mathfrak{K} sei uns nicht gegeben; dagegen nehmen wir als gegeben die drei Bestimmungsstücke an, die in der Einleitung, II eingeführt worden sind. Es sei uns also die gewöhnliche Gruppe $\mathfrak{C} = E + A + B + \dots$ gegeben, die wir uns jetzt als Gruppe der gewöhnlichen Substitutionen der m Symbole a_1, a_2, \dots, a_m vorstellen; ferner seien uns die Zahlen r und s und alle Produkte $E_{11} E_{\kappa\lambda}$ (als Substitutionen der Gruppe \mathfrak{C}) gegeben ($\kappa = 2, \dots, r$; $\lambda = 2, \dots, s$). Wir wollen nun zeigen, daß man bei diesen Bedingungen die Substitutionsgruppe \mathfrak{K} immer konstruieren kann.

Zu dem Zwecke nehmen wir r Systeme von je m Symbolen: a'_1, \dots, a'_m ; a''_1, \dots, a''_m ; \dots ; $a^{(r)}_1, \dots, a^{(r)}_m$, entsprechend den r Gruppen $\mathfrak{A}_1, \mathfrak{A}_2, \dots, \mathfrak{A}_r$, so daß in der unteren Zeile jeder Substitution aus \mathfrak{A}_κ nur die Symbole $a^{(\kappa)}_1, \dots, a^{(\kappa)}_m$ vorkommen sollen. Außer diesen nehmen wir noch irgendeine Anzahl k von Symbolen b_1, \dots, b_k , die wir „Trennungssymbole“ nennen und deren Rolle sich im folgenden aufklären wird. Die Substitutionen aus \mathfrak{A}_κ mögen folgenden Bau haben: in bezug auf $a^{(\kappa)}_1, \dots, a^{(\kappa)}_m$ seien sie gewöhnliche Substitutionen, dieselben wie die Substitutionen aus \mathfrak{C} in bezug auf a_1, \dots, a_m sind; solche Substitutionen in \mathfrak{A}_κ und in \mathfrak{C} sollen, als einander entsprechend, mit einem und demselben Buchstaben bezeichnet werden: $A_{\kappa\lambda} \in \mathfrak{A}_\kappa$, $A \in \mathfrak{C}$. Was die Symbole $a^{(\mu)}_1, \dots, a^{(\mu)}_m$ ($\mu \neq \kappa$) anbetrifft, so gehen sie in $A_{\kappa\lambda}$ in *alle* (in einer bestimmten Ordnung genommenen) Symbole $a^{(\kappa)}_1, \dots, a^{(\kappa)}_m$ über. Also ist jedes der Symbole $a^{(\mu)}_1, \dots, a^{(\mu)}_m$ in $A_{\kappa\lambda}$ mit einem und nur einem der Symbole $a^{(\kappa)}_1, \dots, a^{(\kappa)}_m$ konjugiert (und umgekehrt). Jedes der Trennungssymbole b_1, \dots, b_k ist konjugiert mit irgendeinem der Symbole $a^{(\kappa)}_1, \dots, a^{(\kappa)}_m$ (also auch mit einem der Symbole $a^{(\mu)}_1, \dots, a^{(\mu)}_m$ bei jedem μ). Im ganzen hat *jede* Substitution von \mathfrak{K} $m(r-1) + k$ Konjugationen.

Es ist $\mathfrak{A}_x = \mathfrak{C}_{x_1} + \mathfrak{C}_{x_2} + \dots + \mathfrak{C}_{x_s}$; alle Substitutionen von \mathfrak{C}_{x_λ} haben dieselben Konjugationen, die sich von den Konjugationen von \mathfrak{C}_{x_ν} ($\nu \neq \lambda$) unterscheiden. Die Substitutionen A_{x_λ} und A_{x_ν} stimmen in bezug auf die Symbole $a_1^{(\nu)}, \dots, a_m^{(\nu)}$ völlig miteinander überein; sie unterscheiden sich voneinander nur durch ihre Konjugationen. Andererseits haben alle Substitutionen von $\mathfrak{B}_\lambda = \mathfrak{C}_{1\lambda} + \mathfrak{C}_{2\lambda} + \dots + \mathfrak{C}_{r\lambda}$ dieselben Konjugationen. Die Substitutionen A_{x_λ} und $A_{\mu\lambda}$ unterscheiden sich nur durch die Symbole der unteren Zeile; die sind $a_1^{(\nu)}, \dots, a_m^{(\nu)}$ in A_{x_λ} und $a_1^{(\mu)}, \dots, a_m^{(\mu)}$ in $A_{\mu\lambda}$; A_{x_λ} ist dieselbe Substitution in bezug auf $a_1^{(\nu)}, \dots, a_m^{(\nu)}$ wie $A_{\mu\lambda}$ in bezug auf $a_1^{(\mu)}, \dots, a_m^{(\mu)}$, — dieselbe wie $A < \mathfrak{C}$ in bezug auf a_1, \dots, a_m . In einer Einheit E_{x_λ} muß jedes der Symbole $a_1^{(\nu)}, \dots, a_m^{(\nu)}$ in sich selbst übergehen; das ist für die Einheit notwendig und hinreichend. Es bleibt uns noch die Frage zu beantworten: Wie soll man die Konjugationen in $\mathfrak{B}_1, \dots, \mathfrak{B}_s$ bestimmen?

Bei den eingeführten Bezeichnungen stellt sich folgendes heraus:

1. $E_{x_\lambda} E_{x_\nu} = E_{x_\lambda}$; $E_{x_\lambda} E_{\mu\lambda} = E_{\mu\lambda}$, wie leicht zu verifizieren ist; dagegen ist im allgemeinen: $E_{x_\lambda} E_{\mu\nu} \neq E_{\mu\lambda}$; man kann bloß behaupten, daß $E_{x_\lambda} E_{\mu\nu} < \mathfrak{C}_{\mu\lambda}$ ist, da überhaupt $\mathfrak{C}_{x_\lambda} \mathfrak{C}_{\mu\nu} = \mathfrak{C}_{\mu\lambda}$ ist.

2. Ist in \mathfrak{C} $AB = C$, so ist auch in \mathfrak{A}_x $A_{x_\lambda} B_{x_\mu} = C_{x_\lambda}$; d. h. die Gruppen \mathfrak{C}_{x_λ} und \mathfrak{C}_{x_μ} sind gegeneinander gleich geordnet.

3. Dagegen ist im allgemeinen $A_{x_\lambda} B_{\mu\lambda} \neq C_{\mu\lambda}$, wenn $AB = C$ in der Gruppe \mathfrak{C} ist. Die Konjugationen von A_{x_λ} und $B_{\mu\lambda}$ sind dieselben (die der Gruppe \mathfrak{B}_λ); betrachten wir die Konjugationen der Symbole $a_x^{(\nu)}$ und $a_y^{(\mu)}$ in \mathfrak{B}_λ ; sie mögen durch die Substitution $\begin{pmatrix} a_x^{(\nu)} \\ a_y^{(\mu)} \end{pmatrix}$ ausgedrückt werden

oder mittelst der inversen Substitution: $\begin{pmatrix} a_y^{(\mu)} \\ a_x^{(\nu)} \end{pmatrix}$. Setzen wir a_x und a_y statt $a_x^{(\nu)}$ und $a_y^{(\mu)}$, so bekommen wir eine gewöhnliche Substitution der Symbole a_1, \dots, a_m : $A \simeq \begin{pmatrix} a_x \\ a_y \end{pmatrix}$ (wir können vorläufig noch nicht behaupten, daß A der Gruppe \mathfrak{C} angehört). Nun kann die Substitution A_{x_λ} in bezug auf die Symbole $a_1^{(\mu)}, \dots, a_m^{(\mu)}$ folgendermaßen betrachtet werden: $a_y^{(\mu)}$ wird durch $a_x^{(\nu)}$ ersetzt, und $a_x^{(\nu)}$ erleidet dieselbe Substitution, wie a_x in A ; ferner wird in $B_{\mu\lambda}$ $a_x^{(\nu)}$ durch $a_y^{(\mu)}$ ersetzt, und $a_y^{(\mu)}$ erleidet dieselbe Substitution, wie a_y in B ; es kommt also die folgende Operation (in der Gruppe \mathfrak{C}) heraus:

$$A^{-1} A A B;$$

also in der Gruppe $\mathfrak{C}_{\mu\lambda}$:

$$A_{x_\lambda} B_{\mu\lambda} = A_{\mu\lambda}^{-1} A_{\mu\lambda} A_{\mu\lambda} B_{\mu\lambda} = A'_{\mu\lambda} B_{\mu\lambda},$$

wenn man $A^{-1} A A = A'$ setzt. Wir kommen also wieder zur Formel (17) zurück.

4. Nun sei in \mathfrak{B}_1 $a_x^{(\kappa)}$ immer mit $a_x^{(\mu)}$ ($x = 1, 2, \dots, m$) konjugiert (für alle κ und μ). In diesem Falle ist für \mathfrak{B}_1 die Substitution A identisch, und wir bekommen:

$$A_{\kappa 1} B_{\mu 1} = A_{\mu 1} B_{\kappa 1},$$

d. h. es sind $\mathfrak{C}_{11}, \dots, \mathfrak{C}_{r1}$ gegeneinander gleich geordnet.

5. Sei nun nach 3.: $A_{1\lambda} E_{\kappa\lambda} = X_{\kappa\lambda}^{-1} A_{\kappa\lambda} X_{\kappa\lambda}$; sei ferner $E_{11} E_{\kappa\lambda} = X'_{\kappa 1}$. Die Substitution $X_{\kappa\lambda}$ gibt, wie wir in 3. gesehen haben, die Konjugationen der Symbole a'_x mit den Symbolen $a_y^{(\kappa)}$ in \mathfrak{B}_λ . $X'_{\kappa 1}$ hat dieselben Konjugationen wie \mathfrak{C}_{11} ; die untere Zeile von $X'_{\kappa 1}$ enthält die Symbole $a_1^{(\kappa)}, \dots, a_m^{(\kappa)}$. Daher geht in $X'_{\kappa 1} = E_{11} E_{\kappa\lambda}$ das Symbol $a_x^{(\kappa)}$ in a'_x (durch E_{11}) und a'_x in $a_y^{(\kappa)}$ (durch $E_{\kappa\lambda}$) über, schließlich also $a_x^{(\kappa)}$ in $a_y^{(\kappa)}$; es entsprechen also $X_{\kappa\lambda}$ und $X'_{\kappa 1}$ einer und derselben Substitution X der Symbole a_1, \dots, a_m ; es ist $X_{\kappa\lambda} \equiv X'_{\kappa 1}$; $E_{11} E_{\kappa\lambda} = X_{\kappa 1}$. Daraus folgt, daß $X < \mathfrak{C}$ ist, da alle $E_{11} E_{\kappa\lambda}$ uns als Substitutionen der Gruppe \mathfrak{C} gegeben sind. Was die Substitution A anbetrifft, so ist es leicht zu sehen (vgl. § 5, (17)), daß man für A die Substitution $X^{-1} Y$ nehmen kann, wo $X_{\kappa 1} = E_{11} E_{\kappa\lambda}$, $Y_{\mu 1} = E_{11} E_{\mu\lambda}$ ist; es ist also auch $A < \mathfrak{C}$.

Es werden also alle Konjugationen von \mathfrak{B}_λ durch die Produkte: $E_{11} E_{2\lambda}, E_{11} E_{3\lambda}, \dots, E_{11} E_{r\lambda}$ vollständig definiert; dabei können diese Produkte ganz willkürlich (aus \mathfrak{C}) genommen werden. Nehmen wir $\lambda = 2, 3, \dots, s$, so definieren wir die Konjugationen in $\mathfrak{B}_2, \dots, \mathfrak{B}_s$, und so wird die ganze Substitutionsgruppe \mathfrak{R} vollständig konstruiert.

Das eine braucht nur noch aufgeklärt zu werden: Ist es nicht möglich, daß \mathfrak{B}_λ und \mathfrak{B}_ν ($\lambda \neq \nu$) gleiche Konjugationen bekommen? Das wird in der Tat vorkommen, wenn man $E_{11} E_{2\lambda} = E_{11} E_{2\nu}$, $E_{11} E_{3\lambda} = E_{11} E_{3\nu}$, \dots , $E_{11} E_{r\lambda} = E_{11} E_{r\nu}$ zufällig wählt; dann werden \mathfrak{B}_λ und \mathfrak{B}_ν voneinander gar nicht unterschieden werden können. In diesem Falle kommen die „Trennungssymbole“ b_1, \dots, b_k zu Hilfe: man braucht nur bei diesen Symbolen in \mathfrak{B}_λ die einen Konjugationen, in \mathfrak{B}_ν die anderen von den ersten verschiedenen Konjugationen herzustellen; dadurch wird \mathfrak{B}_λ von \mathfrak{B}_ν „getrennt“. Tritt dieser Fall bei keinem Paare der Gruppen $\mathfrak{B}_1, \dots, \mathfrak{B}_s$ ein, so sind die Trennungssymbole gänzlich überflüssig, man braucht sie auch gar nicht einzuführen.

Damit ist die Frage, die in der Einleitung III. gestellt wurde, vollständig und bejahend beantwortet; gleichzeitig ist bewiesen, daß jede beliebige Kerngruppe sich als Gruppe der verallgemeinerten Substitutionen darstellen läßt, und das Verfahren angegeben, wie diese Darstellung wirklich zu vollziehen ist.

Noch einige Bemerkungen sollen hinzugefügt werden:

1. Die einfachste Kerngruppe \mathfrak{R} wird erhalten, wenn man $E_{11} E_{\kappa\lambda} = E_{\kappa 1}$ für jedes κ und λ setzt; dann ist auch allgemein $E_{\kappa\lambda} E_{\mu\nu} = E_{\mu\lambda}$, und jedes

Produkt $A_{\kappa\lambda} B_{\mu\nu} = C_{\mu\lambda}$ wird einfach nach der Tafel der Gruppe \mathfrak{G} gefunden. Doch ist in diesem Falle die Einführung der Trennungssymbole unvermeidlich.

2. Ist \mathfrak{G} eine Abelsche Gruppe, so hat sie bekanntlich keine inneren Isomorphismen; also kann jedes Produkt $A_{\kappa\lambda} B_{\mu\lambda}$ in diesem Falle direkt nach der Tafel der Gruppe \mathfrak{G} gefunden werden.

3. Die Rechts- und Linksgruppen, sowie die gewöhnlichen Gruppen sind ja Spezialfälle der Kerngruppen überhaupt (wenn eine oder beide der Zahlen r, s gleich 1 sind). Es kann also der Fall vorkommen, daß der Kern einer Gruppe \mathfrak{G} eine Rechts- oder Linksgruppe oder gar eine gewöhnliche Gruppe ist. Der letzte Fall tritt gewiß ein, wenn für die Gruppe \mathfrak{G} das kommutative Gesetz erfüllt ist; der Kern einer solchen Gruppe ist eine gewöhnliche Abelsche Gruppe.

Anhang.

Beispiel einer Kerngruppe.

Es sei uns eine Kerngruppe \mathfrak{K} durch folgende Bedingungen gegeben:

1. \mathfrak{G} sei die symmetrische Gruppe sechster Ordnung (dritten Grades), die also als Gruppe folgender Substitutionen dargestellt werden kann:

$$E = \begin{pmatrix} 0 & 1 & 2 \\ 0 & 1 & 2 \end{pmatrix}, \quad A' \simeq \begin{pmatrix} 0 & 1 & 2 \\ 0 & 2 & 1 \end{pmatrix}, \quad A'' \simeq \begin{pmatrix} 0 & 1 & 2 \\ 2 & 1 & 0 \end{pmatrix},$$

$$A''' \simeq \begin{pmatrix} 0 & 1 & 2 \\ 1 & 0 & 2 \end{pmatrix}, \quad B \simeq \begin{pmatrix} 0 & 1 & 2 \\ 1 & 2 & 0 \end{pmatrix}, \quad C \simeq \begin{pmatrix} 0 & 1 & 2 \\ 2 & 0 & 1 \end{pmatrix},$$

deren Cayleysche Tafel danebensteht.

2. Sei $r = s = 3$.

3. Sei $E_{11} E_{22} = E_{11} E_{33} = A'_{21}$;

$$E_{11} E_{32} = E_{11} E_{33} = B_{31}.$$

E	A'	A''	A'''	B	C
A'	E	C	B	A'''	A''
A''	B	E	C	A'	A'''
A'''	C	B	E	A''	A'
B	A''	A'''	A'	C	E
C	A'''	A'	A''	E	B

Schema III.

Hier haben wir den Fall, wo Trennungssymbole unvermeidlich sind, um die Gruppen \mathfrak{B}_2 und \mathfrak{B}_3 voneinander zu trennen; dazu braucht man ein einziges Trennungssymbol. Wir führen folgende Symbole ein:

0, 1, 2 entsprechend der Gruppe \mathfrak{A}_1 ;

3, 4, 5 " " " \mathfrak{A}_2 ;

6, 7, 8 " " " \mathfrak{A}_3 ;

9 als Trennungssymbol.

In \mathfrak{B}_1 seien konjugiert: 0 mit 3 mit 6, 1 mit 4 mit 7, 2 mit 5 mit 8.

Die Konjugationen in \mathfrak{B}_2 werden durch die Gleichungen $E_{11} E_{22} = A'_{21}$, $E_{11} E_{32} = B_{31}$ definiert; die Konjugationen in \mathfrak{B}_3 ebenso durch die Gleichungen $E_{11} E_{23} = A'_{21}$, $E_{11} E_{33} = B_{31}$.

Das Symbol 9 sei in \mathfrak{B}_1 und \mathfrak{B}_2 mit 0, in \mathfrak{B}_3 mit 1 konjugiert.

Dadurch ist die Substitutionsgruppe \mathfrak{K} vollständig bestimmt; ihre Hauptelemente sind:

$$E_{11} = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 0 & 1 & 2 & 0 & 1 & 2 & 0 & 1 & 2 & 0 \end{pmatrix}; E_{21} = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 3 & 4 & 5 & 3 & 4 & 5 & 3 & 4 & 5 & 3 \end{pmatrix}; E_{31} = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 6 & 7 & 8 & 6 & 7 & 8 & 6 & 7 & 8 & 6 \end{pmatrix};$$

$$E_{12} = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 0 & 1 & 2 & 0 & 2 & 1 & 2 & 0 & 1 & 0 \end{pmatrix}; E_{22} = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 3 & 5 & 4 & 3 & 4 & 5 & 4 & 3 & 5 & 3 \end{pmatrix}; E_{32} = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 7 & 8 & 6 & 7 & 6 & 8 & 6 & 7 & 8 & 7 \end{pmatrix};$$

$$E_{13} = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 0 & 1 & 2 & 0 & 2 & 1 & 2 & 0 & 1 & 1 \end{pmatrix}; E_{23} = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 3 & 5 & 4 & 3 & 4 & 5 & 4 & 3 & 5 & 5 \end{pmatrix}; E_{33} = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 7 & 8 & 6 & 7 & 6 & 8 & 6 & 7 & 8 & 8 \end{pmatrix}.$$

Die ganze Gruppe \mathfrak{K} wird vollständig bekannt, wenn alle Einheitsprodukte $E_{\kappa\lambda} E_{\mu\nu}$ bestimmt werden. Wir brauchen also nur die Einheits-tafel zu konstruieren. Es ist vorteilhaft, die Reihenfolge der Einheiten in der Hauptspalte der Tafel ein wenig anders als in der Hauptzeile zu wählen (s. Schema IV). Bei dieser Wahl zerfällt die ganze Tafel in $r \cdot s$ Abteilungen, so daß jede Abteilung einer bestimmten Kombination der Indizes κ, λ entspricht ($\kappa = 1, \dots, r; \lambda = 1, \dots, s$).

	E_{11}	E_{12}	E_{13}	E_{21}	E_{22}	E_{23}	E_{31}	E_{32}	E_{33}
E_{11}	E_{11}	E_{11}	E_{11}	E_{31}	A'_{21}	A'_{21}	E_{31}	B_{31}	B_{31}
E_{21}	E_{11}	A'_{11}	A'_{11}	E_{21}	E_{21}	E_{21}	E_{31}	A'''_{31}	A'''_{31}
E_{31}	E_{11}	C_{11}	C_{11}	E_{21}	A'''_{21}	A'''_{21}	E_{31}	E_{31}	E_{31}
E_{12}	E_{12}	E_{12}	E_{12}	A'_{22}	E_{22}	E_{22}	C_{32}	E_{32}	E_{22}
E_{22}	A'_{12}	E_{12}	E_{12}	E_{22}	E_{22}	E_{22}	A'''_{32}	E_{32}	E_{32}
E_{32}	B_{12}	E_{12}	E_{12}	A'''_{22}	E_{22}	E_{22}	E_{32}	E_{32}	E_{32}
E_{13}	E_{13}	E_{13}	E_{13}	A'_{23}	E_{23}	E_{23}	C_{33}	E_{33}	E_{33}
E_{23}	A'_{13}	E_{13}	E_{13}	E_{23}	E_{23}	E_{23}	A'''_{33}	E_{33}	E_{33}
E_{33}	B_{13}	E_{13}	E_{13}	A'''_{23}	E_{23}	E_{23}	E_{33}	E_{33}	E_{33}

Schema IV.

Hier sind nur die durch Fettdruck bezeichneten Elemente zu bestimmen; sie können entweder direkt durch die Multiplikation der entsprechenden Einheitssubstitutionen oder abstrakt mit Hilfe der Formeln (17) bis (22) berechnet werden. Beide Methoden führen auf dasselbe hinaus.

Woronesch, den 28. Dezember 1926.

Zur Theorie der allgemeinen Zahlringe.

Von

Wolfgang Krull in Freiburg i. Br.

Verzeichnis einiger inhaltlich verwandter, häufig zitierten Arbeiten:

1. Grell, H.: Zur Theorie der Ordnungen in algebraischen Zahl- und Funktionenkörpern. Math. Annalen 97 (1927), S. 524—558; zitiert mit „G.“.
2. Krull, W.: Algebraische Erweiterungen kommutativer hyperkomplexer Systeme. Math. Annalen 97 (1927), S. 473—489; zitiert mit „K.I.“.
3. Krull, W.: Über verallgemeinerte endliche Abelsche Gruppen. Math. Zeitschrift 23 (1925), S. 161—186; zitiert mit „K.II.“.
4. Krull, W.: Theorie und Anwendung der verallgemeinerten Abelschen Gruppen. Sitzungsber. der Heidelberger Akademie d. Wissenschaften 1926, 1. Abhandlung; zitiert mit „K.III.“.
5. Noether, E.: Abstrakter Aufbau der Idealtheorie in algebraischen Zahl- und Funktionenkörpern. Math. Annalen 96 (1926), S. 26—61; zitiert mit „N.“.

In der ersten Hälfte unserer Note beschäftigen wir uns ausschließlich mit idealtheoretischen Problemen; in der zweiten steht das Studium gewisser verallgemeinerter Abelscher Gruppen (im Sinne von K.II und K.III) im Vordergrund. Der Zusammenhang zwischen beiden Teilen wird dadurch hergestellt, daß wir die Ideale eines Ringes nach dem Vorbild von E. Noether als verallgemeinerte Abelsche Gruppen auffassen.

Den Ausgangspunkt bilden die Hauptergebnisse von N., die sich kurz so formulieren lassen:

„Ein nullteilerfreier Ring¹⁾ besitzt dann und nur dann die charakteristische Eigenschaft des Ringes der ganzen Größen eines endlichen algebraischen Zahlkörpers, daß sich jedes Ideal eindeutig als Produkt von Primidealepotenzen darstellen läßt, wenn die folgenden Axiome erfüllt sind²⁾;

¹⁾ Unter „Ring“ schlechtweg verstehen wir stets einen kommutativen Ring mit Einheitslement im Gegensatz zu N., wo (insbesondere in § 6 und § 7) auch Ringe ohne Einheitslement betrachtet werden.

²⁾ Die Axiome 1_a, 1_b und 2 entsprechen den Noetherschen Axiomen I, II und V. Die Noetherschen Axiome III und IV beziehen sich auf die Existenz des Einheitslements und den Ausschluß der Nullteiler.

1_a. Teilerkettensatz. *Jede Kette von Idealen, bei der jedes Ideal ein echter Teiler des vorangehenden ist, bricht im Endlichen ab.*

1_b. Vielfachenkettensatz modulo jedes vom Nullideal verschiedenen Ideals. *Jede Kette von Idealen — die sämtlich Teiler eines festen, vom Nullideal verschiedenen Ideals sind —, bei der jedes Ideal ein echtes Vielfaches des vorangehenden ist, bricht im Endlichen ab.*

2. Ganze Abgeschlossenheit im Quotientenkörper. *Jedes Element des Quotientenkörpers, das ganz in bezug auf den Ring ist, gehört dem Ring an.* ²⁾

Gleichwertig mit Axiom 1_a und 1_b ist die gruppentheoretische Forderung:

1. *Das Restklassensystem nach jedem vom Nullideal verschiedenen Ideal besitzt als verallgemeinerte Abelsche Gruppe eine Jordansche Kompositionsreihe.*

Ein allgemeiner Ring, der den Axiomen 1_a und 1_b genügt, besitzt die folgenden beiden Eigenschaften:

α) *Jedes vom Nullideal verschiedene Ideal ist das kleinste gemeinschaftliche Vielfache von endlich viel teilerfremden (starken³⁾) Primäridealen (mit endlichem Exponenten³⁾).*

β) *Jedes Ideal ist als kleinstes gemeinschaftliches Vielfaches von endlich viel irreduziblen Idealen darstellbar.*

In unserer Note werden die Noetherschen Ergebnisse in verschiedener Hinsicht verschärft. Was zunächst den an letzter Stelle angeführten Satz angeht, so können wir zeigen, daß er sich umkehren läßt, daß also nicht nur die Gültigkeit von α) und β) aus der von 1_a und 1_b, sondern auch umgekehrt die Gültigkeit von 1_a und 1_b aus der von α) und β) abgeleitet werden kann. Damit ist insbesondere für alle endlichen „Ordnungen“ in Zahlkörpern dasselbe geleistet wie bei N. für die Hauptordnung.

Wir erhalten dieses Ergebnis am Ende von § 5 im Anschluß an die axiomatischen Untersuchungen, die in § 1 und § 5 über allgemeine „Zahlringe“ angestellt werden, d. h. über nullteilerfreie Ringe, denen wohl die Eigenschaft α), aber nicht notwendig die Eigenschaft β) zukommt. Das Hauptresultat über Zahlringe selbst lautet: ²⁾

³⁾ Zum Unterschied zwischen „starkem“ und „schwachem“ Primärideal, sowie über die Bedeutung des endlichen Exponenten vgl. N. § 5. Ein starkes Primärideal q mit endlichem Exponenten ist dadurch ausgezeichnet, daß eine Potenz des zugehörigen Primideals p durch q teilbar ist, das kleinste r , für das $p^r \geq q$ ausfällt, ist der Exponent von q . Im folgenden verstehen wir unter „Primärideal“ schlechtweg stets ein starkes Primärideal mit endlichem Exponenten.

¹⁾ folgt aus β): $\#a \geq 2$

²⁾ was folgt aus der Gültigkeit von 2; $\#a \geq 2$ für $\forall a \in \mathfrak{a}$?

Notwendig und hinreichend dafür, daß der nullteilerfreie Ring \mathfrak{R} einen Zahlring darstellt, sind die folgenden beiden Bedingungen:

1_a. Quotientenkettensatz. Jede Kette a_1, a_2, a_3, \dots , bei der a_{i+1} stets die Gestalt $a_i \cdot b_i$ besitzt und echter Teiler von a_i ist, bricht im Endlichen ab. (\rightarrow für d. Stumpfheit \leftarrow teilerfremd irreduzibel)

1_b. Produktkettensatz modulo jedes vom Nullideal verschiedenen Ideals. Jede Kette a_1, a_2, a_3, \dots , bei der stets a_{i+1} die Gestalt $a_i \cdot b_i$ besitzt, echtes Vielfaches von a_i und Teiler eines festen vom Nullideal verschiedenen Ideals ist, bricht im Endlichen ab.)

1_a bzw. 1_b stellt offenbar eine nie stärkere und i. a. schwächere Voraussetzung als 1_a bzw. 1_b dar. Ein ganz entsprechendes Ergebnis erhalten wir für allgemeine „kommutative hyperkomplexe Systeme“, d. h. für Ringe, in denen zwar Nullteiler zugelassen sind, aber die Bedingung α) für jedes Ideal ausnahmslos erfüllt ist. Es zeigt sich, daß die kommutativen hyperkomplexen Systeme durch die Gültigkeit des Quotientenkettensatzes, sowie des Produktkettensatzes modulo jedes beliebigen Ideals eindeutig umkehrbar charakterisiert sind.

Ebenso wie die Axiome 1_a und 1_b lassen auch die Axiome 1_a' und 1_b' eine gruppentheoretische Deutung zu. Man kann nämlich 1_a' und 1_b' ersetzen durch die Forderung, daß das Restklassensystem nach jedem vom Nullideal verschiedenen Ideal eine endliche „Loewysche Kompositionsreihe“ von bestimmtem, leicht angebbarem Bau besitzen soll. (Den Begriff der Loewyschen Kompositionsreihe habe ich in K. III eingeführt; der wesentliche Unterschied zwischen Jordanscher und Loewyscher Kompositionsreihe liegt darin, daß die Glieder der Jordanschen Kompositionsreihe irreduzibel, die der Loewyschen vollständig reduzibel sind.)

Die Brauchbarkeit der Loewyschen Kompositionsreihe zeigt sich auch darin, daß mit ihrer Hilfe der scheinbar rein idealtheoretische Begriff des „Exponenten“ eines Primärideals sich gruppentheoretisch deuten läßt, nämlich als Länge der Loewyschen Kompositionsreihe des Restklassensystems. Allgemeiner habe ich in K. III § 4 gezeigt, wie man mit Hilfe des Begriffes der gruppentheoretischen Irreduzibilität und vollständigen Reduzibilität — auf dem ja die Jordansche und die Loewysche Kompositionsreihe beruhen — sämtliche wesentlichen Eigenschaften eines Primärideals aus dem gruppentheoretischen Bau des Restklassensystems ableiten kann. — Es sei hervorgehoben, daß es für unsere Zwecke notwendig ist, die vollständige Reduzibilität einer Gruppe etwas allgemeiner zu definieren als üblich; wir müssen nämlich auch solche Gruppen vollständig reduzibel nennen, die eine direkte Summendarstellung durch unendlich viel irreduzible Komponenten besitzen. In § 4 wird gezeigt, daß

1) Ziel für zahlringe Sätze über Primidealketten, in über das Abbrechen der

diese Verallgemeinerung an den bekannten Fundamentalsätzen über vollständig reduzierbare Gruppen nichts ändert.

Außer den bisher skizzierten Entwicklungen enthält die Note noch die Behandlung zweier idealtheoretischer Probleme, die sich ausschließlich auf Zahlringe, nicht auf hyperkomplexe Systeme beziehen. Zunächst können wir zeigen (§ 3):

In einem Zahlring \mathfrak{R} sind dann und nur dann die Primidealepotenzen die einzigen Primär Ideale, wenn er hinsichtlich des Quotientenkörpers \mathfrak{K} ganz abgeschlossen ist.

Unser Satz stellt eine Verschärfung des oben an erster Stelle angeführten Noetherschen Theorems dar, weil ja ein beliebiger Zahlring nicht den Axiomen 1_a und 1_b , sondern nur den Axiomen $1'_a$ und $1'_b$ zu genügen braucht; der Beweis ist zugleich wesentlich einfacher.

a) Wir führen nach dem Vorbild der Dedekindschen Supplemente zur vierten Auflage von Dirichlets Vorlesungen über Zahlentheorie gebrochene Ideale (als endliche \mathfrak{R} -Moduln in \mathfrak{K}) ein, und können dann in wenigen Zeilen unter Vermeidung von allem Rechnen mit Elementen zeigen, daß in einem ganz abgeschlossenen Zahlring jedes Primärideal Primidealepotenz ist.

b) Die Richtigkeit der Umkehrung ergibt sich leicht durch Heranziehen gewisser „regulärer Quotientenringe“.

Der skizzierte Beweis scheint mir wegen seiner Kürze und Durchsichtigkeit für eine gleichzeitig elementare und begriffliche Darstellung der Idealtheorie in algebraischen Zahlkörpern brauchbar zu sein⁴⁾.

Die soeben erwähnten „regulären Quotientenringe“ eines Zahlrings werden in § 2 eingeführt und untersucht. Jeder Primidealschar S aus \mathfrak{R} entspricht eindeutig umkehrbar ein regulärer Quotientenring \mathfrak{R}_S , und es besteht Isomorphie hinsichtlich der Teilbarkeitsverhältnisse und der Produktbildung zwischen dem Bereich aller Ideale von \mathfrak{R}_S und dem Bereich derjenigen Ideale von \mathfrak{R} , deren Primidealteiler sämtlich der Schar S angehören.

Auf die Bedeutung der regulären Quotientenringe habe ich in meinem Vortrag auf der Deutschen Mathematikertagung in Danzig 1925 hin-

⁴⁾ Einer mündlichen Mitteilung zufolge hat E. Artin meinen Beweis auf Maximalbereiche endlicher, nichtkommutativer hyperkomplexer Systeme übertragen, und dadurch die Ableitung der Ergebnisse der Arbeit „Allgemeine Zahlentheorie“ von A. Speiser (Vierteljahrsschrift der Naturforschenden Gesellschaft in Zürich 71 (1926)) wesentlich vereinfacht.

Zusatz b. d. Korrektur: In der Zwischenzeit hat Herr Artin seine Ergebnisse veröffentlicht: Zur Arithmetik hyperkomplexer Zahlen, Abhandlungen aus dem Mathematischen Seminar der Hamburgischen Universität 5 (1927), S. 261—289.

gewiesen. Seither hat sich H. Grell eingehend mit ihnen beschäftigt (G. § 2), doch behandelt er nur den wichtigsten, für seine Zwecke ausreichenden Spezialfall⁵⁾. Wir brauchen die allgemeinsten regulären Quotientenringe, um in § 3 den folgenden Satz zu beweisen:

Um sämtliche Unterringe eines endlichen algebraischen Zahlkörpers \mathfrak{K} zu beherrschen, braucht man nur den Aufbau der ganzzahligen zu kennen; denn es ist jeder beliebige Unterring regulärer Quotientenring eines ganzzahligen. (siehe „Ordnung“)

§ 1.

Hyperkomplexe Systeme und Zahlringe.

Im folgenden werden die Grundbegriffe der allgemeinen Idealtheorie als bekannt vorausgesetzt. Für die Teilbarkeitsbeziehungen, den größten gemeinschaftlichen Teiler, das kleinste gemeinschaftliche Vielfache, den Quotienten verwenden wir die bzw. Schreibweisen $a > b$, $b < a$; $a \div b$; $a \cap b$; $a : b$. Gebrochene Ideale bezeichnen wir mit $\frac{a}{b}$ usw., das Restklassensystem von \mathfrak{R} nach a mit \mathfrak{R}/a . Unter einem Ring schlechtweg verstehen wir stets einen kommutativen Ring mit Einheitselement ε . o bedeutet das Einheits-, n das Nullideal.

Besonders hervorgehoben werden möge die folgende Begriffsbildung:

Es sei \mathfrak{R} Unterring von \mathfrak{S} , a_r bzw. a_s ein beliebiges Ideal aus \mathfrak{R} bzw. \mathfrak{S} . Dann verstehen wir unter dem zu a_r (hinsichtlich \mathfrak{S}) gehörigen „Erweiterungsideal“ das Ideal $a_r \cdot \mathfrak{S}$, also das kleinste gemeinschaftliche Vielfache aller der b_s , die sämtliche Elemente von a_r enthalten, unter dem zu a_s (hinsichtlich \mathfrak{R}) gehörigen „Verengungsideal“ das Ideal $a_s \cap \mathfrak{R}$, also den größten gemeinschaftlichen Teiler aller der b_r , für die $b_r \cdot \mathfrak{S}$ Vielfaches von a_s ist⁶⁾.

Es ist $(a_r \cdot \mathfrak{S}) \cap \mathfrak{R} \leq a_r$; $(a_s \cap \mathfrak{R}) \cdot \mathfrak{S} \geq a_s$. Gelten stets die „reziproken Gleichungen“ $(a_r \cdot \mathfrak{S}) \cap \mathfrak{R} = a_r$, $(a_s \cap \mathfrak{R}) \cdot \mathfrak{S} = a_s$, so ordne man jedem a_s sein Verengungsideal, jedem a_r sein Erweiterungsideal zu; man erhält auf diese Weise zwischen den Idealen von \mathfrak{S} und \mathfrak{R} eine eindeutig umkehrbare Beziehung, bei der aus $a_s \geq b_s$ bzw. $a_s \cdot b_s = c_s$ stets $a_r \geq b_r$ bzw. $a_r \cdot b_r = c_r$ folgt und umgekehrt, wenn a_s und a_r , b_s und b_r , c_s und c_r jeweils entsprechende Ideale aus \mathfrak{S} und \mathfrak{R} bedeuten. Wir wollen in diesem Fall die zu \mathfrak{R} und \mathfrak{S} gehörigen Idealbereiche *äquivalent* nennen.

⁵⁾ Vgl. Anm. 7).

⁶⁾ Die Bezeichnungen „Erweiterungs-“ und „Verengungsideal“ stammen von H. Grell: Beziehungen zwischen den Idealen verschiedener Ringe, *Math. Annalen* 97 (1927), § 3. Es ist dort auch die im folgenden betrachtete Zuordnung genauer ausgeführt.

Von Äquivalenz soll auch dann gesprochen werden, wenn sich die Gültigkeit der reziproken Gleichungen und mithin die Zuordnung nicht auf die vollen Idealbereiche, sondern nur auf gewisse Scharen von Idealen aus \mathfrak{K} und \mathfrak{S} (Idealkörper im Sinne von H. Grell) bezieht.

Definition 1. \mathfrak{K} heißt (kommutatives) System hyperkomplexer Größen, wenn 1. kein Primideal aus \mathfrak{K} einen von \mathfrak{o} verschiedenen echten Teiler besitzt und 2. jedes Ideal aus \mathfrak{K} eindeutig als Produkt endlich vieler teilerfremder (starker⁸⁾) Primärideale (von endlichen Exponenten⁹⁾) dargestellt werden kann. \mathfrak{K} heißt Zahlring, wenn \mathfrak{n} Primideal ist, und die Bedingungen 1 und 2 für jedes Ideal $\neq \mathfrak{n}$ gelten.

Wir ziehen zunächst einige Folgerungen aus unserer Definition und verstehen dabei unter \mathfrak{n} ein beliebiges Ideal aus dem hyperkomplexen System \mathfrak{K} bzw. ein von \mathfrak{n} verschiedenes Ideal aus dem Zahlring \mathfrak{K} . Die nach Voraussetzung vorhandene, eindeutige Primärfaktorzerlegung $\alpha = \alpha_1 \cdot \alpha_2 \cdot \dots \cdot \alpha_n$ bezeichnen wir als die „Normalzerlegung von α “. Offenbar sind die zu den α_i gehörigen Primideale \mathfrak{p}_i die einzigen in α überhaupt aufgehenden Primideale, und es wird (für genügend großes r) $(\mathfrak{p}_1 \cdot \mathfrak{p}_2 \cdot \dots \cdot \mathfrak{p}_n)^r \supseteq \alpha$. Ist \mathfrak{p} ein beliebiges Primideal, so verstehen wir unter der „ \mathfrak{p} -Komponente von α “ den zugehörigen Primärfaktor der Normalzerlegung oder \mathfrak{o} , je nachdem ob \mathfrak{p} Teiler von α ist oder nicht.

Zwei Ideale ohne gemeinsamen Primidealteiler sind teilerfremd. \mathfrak{K}/α stellt (auch wenn \mathfrak{K} Zahlring) stets ein hyperkomplexes System dar, dessen Primideale eindeutig umkehrbar den in α aufgehenden Primidealen aus \mathfrak{K} entsprechen.

Unter einer Quotienten- bzw. Produktkette verstehen wir eine unendliche Folge $\alpha_1, \alpha_2, \alpha_3, \dots$ von Idealen, bei der sich α_{i+1} stets auf die Form $\alpha_{i+1} = \alpha_i : \mathfrak{b}_i$ bzw. $\alpha_{i+1} = \alpha_i \cdot \mathfrak{b}_i$ bringen läßt. Wird von einem gewissen n an: $\alpha_n = \alpha_{n+1} = \alpha_{n+2} = \dots$, so sagen wir, die Kette „breche im Endlichen ab“.

Satz 1. \mathfrak{K} ist dann und nur dann hyperkomplexes System, wenn jede Quotientenkette aus \mathfrak{K} und jede Produktkette aus \mathfrak{K}/α für beliebiges α im Endlichen abbricht.

\mathfrak{K} ist dann und nur dann Zahlring, wenn \mathfrak{n} Primideal ist, und wenn jede Quotientenkette aus \mathfrak{K} und jede Produktkette aus \mathfrak{K}/α für beliebiges $\alpha \neq \mathfrak{n}$ im Endlichen abbricht.

Wir führen den Beweis nur für das hyperkomplexe System. Die Änderungen, die sich im Falle des Zahlrings durch die Sonderstellung von \mathfrak{n} ergeben, sind trivial.

1_a. Es seien $\mathfrak{p}_1, \mathfrak{p}_2, \dots, \mathfrak{p}_n$ die sämtlichen in α aufgehenden Primideale, $\mathfrak{p}_1^{r_1} \cdot \mathfrak{p}_2^{r_2} \cdot \dots \cdot \mathfrak{p}_n^{r_n} \supseteq \alpha$, $\mathfrak{b} = \alpha : c$. Dann ist entweder $\mathfrak{b} = \alpha$, wenn

nämlich c durch kein p_i teilbar ist, oder es gilt eine Gleichung $p_1^{s_1} \cdot p_2^{s_2} \cdot \dots \cdot p_n^{s_n} \geq b$, bei der kein Exponent s_i größer als der entsprechende Exponent r_i , aber mindestens einer kleiner ist. Daraus folgt leicht, daß in einem hyperkomplexen System die Quotientenkettenvoraussetzung erfüllt ist.

1_b. Um zu zeigen, daß in einem hyperkomplexen System \mathfrak{R} die Produktkettenvoraussetzung gilt, genügt es nachzuweisen, daß in \mathfrak{R} selbst jede Produktkette im Endlichen abbricht, weil ja mit \mathfrak{R} stets auch \mathfrak{R}/α ein hyperkomplexes System darstellt. — Ist $n = q_1 \cdot q_2 \cdot \dots \cdot q_n$ die Normalzerlegung des Nullideals, so sind die zu den q_i gehörenden Primideale p_i die einzigen Primideale in \mathfrak{R} , und es gelten n Gleichungen $q_i = p_i^{r_i} = p_i^{r_i+1} = p_i^{r_i+2} = \dots$. Daraus folgt sofort das Abbrechen der Produktketten. (Man beachte, daß $a \cdot b \geq p^{r+s}$, wenn $a \geq p^r$, $b \geq p^s$!)

2_a. Gilt in \mathfrak{R} die Produktkettenvoraussetzung, so besitzt in \mathfrak{R} kein Primideal einen echten Teiler $\neq 0$.

Es sei p ein Primideal, $\bar{a} = (\alpha)$ sei ein beliebiges, vom Nullideal verschiedenes Hauptideal aus \mathfrak{R}/p . Dann gilt nach Voraussetzung eine Gleichung $\bar{a}^r = \bar{a}^{r+1}$, und durch Übergang zu den Elementen erhalten wir: $\alpha^{r+1} \cdot \beta = \alpha^r$, $\alpha^r \cdot (\alpha \cdot \beta - \varepsilon) = 0$. Da in \mathfrak{R}/p keine Nullteiler auftreten, muß $\alpha \cdot \beta = \varepsilon$ sein, d. h. jedes von 0 verschiedene Element aus \mathfrak{R}/p ist Einheit, \mathfrak{R}/p ist ein Körper und 0 daher der einzige echte Teiler von p .

2_b. In K.I § 1 habe ich bewiesen, daß ein Ring notwendig ein hyperkomplexes System darstellt, wenn in ihm die Quotientenkettenvoraussetzung gilt, und außerdem kein Primideal einen von 0 verschiedenen echten Teiler besitzt.

Satz 2. \mathfrak{R} ist dann und nur dann Zahlring, wenn n Primideal ist und für beliebiges $\alpha \neq n$ der Ring \mathfrak{R}/α stets ein hyperkomplexes System darstellt.

Daß die angegebene Bedingung notwendig ist, wurde beiläufig oben hervorgehoben. Daß sie hinreicht, ergibt sich aus Satz 1 und der Tatsache, daß eine mit n beginnende Quotientenkette höchstens zwei verschiedene Glieder (nämlich n und 0) enthält.

§ 2.

Die regulären Quotientenringe eines Zahlringes.

In den folgenden beiden Paragraphen beschäftigen wir uns ausschließlich mit Zahlringen.

Da ein Zahlring \mathfrak{R} keine Nullteiler enthält, kann er zu einem eindeutig bestimmten Körper \mathfrak{K} ergänzt werden, den wir als „Quotienten-

Körper von \mathfrak{K} “ bezeichnen, weil sich jedes Element aus \mathfrak{K} als Quotient von Elementen aus \mathfrak{K} darstellen läßt. Der Aufbau von \mathfrak{K} über \mathfrak{K} geschieht genau nach demselben Schema, nach dem wir im Sonderfall vom Zahlring der ganzen natürlichen Zahlen zum Körper der rationalen Zahlen aufsteigen.

(*endl. oder unendl. viele*)
 Es sei S eine beliebige Schar von Primidealen aus \mathfrak{K} . Dann ordnen wir ihr einen eindeutig bestimmten, zwischen \mathfrak{K} und \mathfrak{K} liegenden Ring \mathfrak{K}_S durch folgende Festsetzung zu:

Ein Element α aus \mathfrak{K} soll dann und nur dann zu \mathfrak{K}_S gehören, wenn in \mathfrak{K} ein durch kein Primideal von S teilbares Ideal a existiert, das die Eigenschaft hat, daß das Produkt eines beliebigen seiner Elemente mit α stets zu \mathfrak{K} gehört.

Wir können, wie leicht einzusehen, \mathfrak{K}_S auch folgendermaßen definieren:

\mathfrak{K}_S besteht aus der Gesamtheit der Elemente, die eine Quotientendarstellung besitzen, bei der für jedes p aus S die p -Komponente des (als Hauptideal in \mathfrak{K} aufgefaßten) Zählers durch die p -Komponente des Nenners teilbar ist.

Die Ringe \mathfrak{K}_S bezeichnen wir als die „regulären Quotientenringe von \mathfrak{K} “, während wir unter „Quotientenring“ schlechtweg einen beliebigen Zwischenring zwischen \mathfrak{K} und \mathfrak{K} verstehen⁷⁾.

Hilfssatz 1. *$(a \cdot \mathfrak{K}_S) \cap \mathfrak{K}$ ist nur durch solche Primideale teilbar, die zu S gehören.*

Jedes von n verschiedene, durch a teilbare Hauptideal (a) besitzt eine Produktzerlegung $(a) = q_1 \cdot q_2$, bei der q_1 durch kein Primideal aus S , q_2 nur durch Primideale aus S teilbar ist. Bestimmen wir, was stets möglich ist, das Element b aus \mathfrak{K} so, daß die Beziehungen $(b) + q_1 = 0$, $(b) \geq q_2$ gelten, so gehört $\frac{b}{a}$ zu \mathfrak{K}_S und mithin $b = \frac{b}{a} \cdot a$ zu $(a \cdot \mathfrak{K}_S) \cap \mathfrak{K}$, d. h. es ist $(a, b) \geq (a \cdot \mathfrak{K}_S) \cap \mathfrak{K}$. Daraus folgt aber die Behauptung des Hilfssatzes, weil offenbar (a, b) nur durch Primideale aus S teilbar ist. — Wenden wir den Hilfssatz insbesondere auf ein Ideal a an, das zu sämtlichen Primidealen aus S teilerfremd ist, so erhalten wir:

Ist a durch kein Primideal aus S teilbar, so wird $a \cdot \mathfrak{K}_S = 0 \cdot \mathfrak{K}_S$; $(a \cdot \mathfrak{K}_S) \cap \mathfrak{K} = 0$.

⁷⁾ In G. § 2 wird Satz 3 für diejenigen regulären Quotientenringe bewiesen, bei denen die zugehörige Primidealschar nur aus endlich viel Individuen besteht. Das Wort „Quotientenring“ wird in G. und auch in der in Anm. ⁶⁾ zitierten Arbeit in wesentlich speziellerem Sinne gebraucht als bei uns, es braucht nicht einmal jeder „reguläre Quotientenring“ in unserem Sinne ein „Quotientenring“ im Grelleschen Sinne zu sein. (Dieser Fall tritt z. B. ein, wenn S aus allen Primidealen bis auf p besteht, und wenn kein zu p gehöriges Primärideal in \mathfrak{K} Hauptideal ist.)

Satz 3. *Der zu \mathfrak{R}_S gehörige Idealbereich ist äquivalent zum Bereich aller der Ideale aus \mathfrak{R} , die zu sämtlichen nicht in S vorkommenden Primidealen teilerfremd sind.*

Nach der in § 1 gegebenen Definition der Äquivalenz haben wir zum Beweise unseres Satzes im wesentlichen zu zeigen, daß jedes Ideal aus \mathfrak{R}_S Erweiterungsideal seines Verengungsideals, jedes zu allen nicht in S vorkommenden Primidealen teilerfremde Ideal aus \mathfrak{R} Verengungsideal seines Erweiterungsideals ist.

a) Aus dem Zusatz zu Hilfssatz 1 sowie aus der Definition von \mathfrak{R}_S folgt leicht, daß jedes Hauptideal und mithin jedes beliebige Ideal aus \mathfrak{R}_S eine (natürlich nicht notwendig endliche) Basis aus \mathfrak{R} besitzt, daß also für jedes Ideal α_s aus \mathfrak{R}_S die Gleichung $(\alpha_s \cap \mathfrak{R}) \cdot \mathfrak{R}_S = \alpha_s$ gilt.

b) Daß bei beliebigem α aus \mathfrak{R} das Ideal $(\alpha \cdot \mathfrak{R}_S) \cap \mathfrak{R}$ zu jedem nicht in S vorkommenden Primideal teilerfremd ist, lehrt Hilfssatz 1. Es sei jetzt umgekehrt α nur durch Primideale aus S teilbar, und es bedeute α ein beliebiges Element aus $(\alpha \cdot \mathfrak{R}_S) \cap \mathfrak{R}$. Dann gilt eine Gleichung $\alpha = \sum_{i=1}^m \alpha_i \alpha_i$, bei der die α_i Elemente aus α , die α_i Elemente aus \mathfrak{R}_S sind.

Bestimmen wir, was nach Definition von \mathfrak{R}_S möglich, ein zu α teilerfremdes Ideal \mathfrak{b} aus \mathfrak{R} so, daß das Produkt eines α_i mit einem Elemente aus \mathfrak{b} stets zu \mathfrak{R} gehört, so ist $\mathfrak{b} \cdot (\alpha)$ ein Vielfaches von α , und daraus folgt wegen $\alpha + \mathfrak{b} = \mathfrak{o}$ weiter $(\alpha) \supseteq \alpha$. Es gehört also ein beliebiges Element aus $(\alpha \cdot \mathfrak{R}_S) \cap \mathfrak{R}$ stets zu α , d. h. es gilt die Gleichung $(\alpha \cdot \mathfrak{R}_S) \cap \mathfrak{R} = \alpha$.

Durch Satz 3 ist die Frage nach dem Bau der regulären Quotientenringe vollkommen geklärt. Eine besonders wichtige Rolle spielen diejenigen regulären Quotientenringe, bei denen die Schar S durch ein einziges Primideal \mathfrak{p} gebildet wird. Wir wollen einen solchen Ring auch mit $\mathfrak{R}_{\mathfrak{p}}$ bezeichnen.

$\mathfrak{R}_{\mathfrak{p}}$ besteht aus der Gesamtheit derjenigen Elemente aus \mathfrak{R} , die eine Quotientendarstellung mit zu \mathfrak{p} teilerfremdem Nenner besitzen.

In der Tat, jedes derartige Element gehört zu $\mathfrak{R}_{\mathfrak{p}}$. Ist andererseits α ein Element aus $\mathfrak{R}_{\mathfrak{p}}$ mit der Quotientendarstellung $\frac{a}{b}$, so ist die \mathfrak{p} -Komponente von (a) durch diejenige von (b) teilbar. Wir können daher ein zu \mathfrak{p} teilerfremdes Element c so wählen, daß $(a) \cdot (c) \supseteq (b)$, d. h. $\frac{a \cdot c}{b} = d$ ein Element aus \mathfrak{R} wird; dann aber haben wir in $\alpha = \frac{d}{c}$ eine Quotientendarstellung von α mit zu \mathfrak{p} teilerfremdem Nenner.

Satz 4. \mathfrak{R}_S ist gleich dem kleinsten gemeinschaftlichen Vielfachen (also dem mengentheoretischen Durchschnitt) aller der $\mathfrak{R}_{\mathfrak{p}}$, deren zugehöriges \mathfrak{p} in S vorkommt.

Es ist nur zu zeigen, daß α in \mathfrak{R}_p auftreten muß, wenn es in jedem der genannten \mathfrak{R}_p vorkommt. Das ergibt sich aber aus der Tatsache, daß zu einem solchen α für beliebiges p aus S ein zu p teilerfremdes Ideal \mathfrak{a} aus \mathfrak{R} existiert, derart, daß das Produkt von α mit einem beliebigen Elemente aus \mathfrak{a} in \mathfrak{R} , also um so mehr in \mathfrak{R}_p enthalten ist.

Satz 5. *In einem Zahlring sind (abgesehen natürlich von 0 und 1) die Potenzen eines Ideals sämtlich verschieden.*

Ist $a^2 = a$, so muß für jede Primärkomponente q von a die Gleichung $q^2 = q$ gelten. Es genügt daher, wenn wir den Satz für Primär Ideale beweisen, und wir dürfen uns dabei nach Satz 3 auf den Fall eines \mathfrak{R}_p mit dem einzigen Primideal p beschränken. Es sei $q^2 = q$, $p^r \nsubseteq q$. Dann ist $q^{r+1} = q \supseteq p^{r+1}$, also $q = p^r = p^{r+1} = \dots$, d. h. p^r stellt in \mathfrak{R}_p das kleinste gemeinschaftliche Vielfache aller von π verschiedenen Ideale dar.) Daraus folgt, daß für $a \neq 0$ aus p^r die Gleichung $p^r = (a) = (a)^2$ gelten muß, und durch Übergang von den Idealen zu den Elementen erhalten wir $a = a^2 \cdot b$, $a(\varepsilon - a \cdot b) = 0$. Die letztere Gleichung ist aber unmöglich, weil a weder Nullteiler noch Einheit ist, die Ausgangsannahme $q^2 = q$ war also falsch.

X

§ 3.

Ganz abgeschlossene Ringe.

Definition 2. *Ein Element α aus dem Quotientenkörper \mathfrak{K} heißt „von \mathfrak{R} ganz abhängig“, wenn es in \mathfrak{R} ein Element $a \neq 0$ gibt, dessen Produkt mit einer beliebigen Potenz von α stets zu \mathfrak{R} gehört⁸⁾.*

Die Gesamtheit aller von \mathfrak{R} ganz abhängigen Elemente bildet ersichtlich einen zwischen \mathfrak{R} und \mathfrak{K} liegenden Ring \mathfrak{R}^* . Ist $\mathfrak{R}^* = \mathfrak{R}$, so heißt \mathfrak{R} „ganz abgeschlossen“.

Hilfssatz 2. *α hängt dann und nur dann ganz von \mathfrak{R} ab, wenn es von sämtlichen zu \mathfrak{R} gehörigen \mathfrak{R}_p ganz abhängt. \mathfrak{R} ist dann und nur dann ganz abgeschlossen, wenn es alle \mathfrak{R}_p sind.*

Bei der ersten Behauptung bedarf allein das „dann“, bei der zweiten allein das „nur dann“ eines Beweises. Wir verstehen unter $\alpha = \frac{a}{b}$ ein

⁸⁾ Es ist wesentlich, daß in unserem Falle diese grundsätzlich auf Dedekind zurückgehende Definition des algebraisch ganz abhängigen Elementes zugrunde gelegt wird; denn man kann bei einem allgemeinen Zahlring \mathfrak{R} unmittelbar nur zeigen, daß ein Element α aus dem Quotientenkörper \mathfrak{K} , das einer Gleichung $\alpha^n + a_1 \alpha^{n-1} + \dots + a_n = 0$ mit Koeffizienten aus \mathfrak{R} genügt, von \mathfrak{R} im Sinne der Definition 2 ganz abhängt, während zum Beweis der Umkehrung der Satz 1a von der endlichen Teilerkette unentbehrlich zu sein scheint. (Vgl. N. S. 31 f., insbesondere die „Dedekindsche Folgerung“ auf S. 33.)

1) Da jedes a in \mathfrak{R}_p ein y^2 enthält

von sämtlichen \mathfrak{R}_p ganz abhängiges Element, unter p_1, p_2, \dots, p_n die in (b) aufgehenden Primideale, unter b_i ein (stets existierendes) von 0 verschiedenes Element aus \mathfrak{R} , dessen Produkt mit einer beliebigen Potenz von α in \mathfrak{R}_p auftritt. Dann sind die Elemente $b_1 \cdot b_2 \cdot \dots \cdot b_n \cdot \alpha^\kappa$ ($\kappa = 1, 2, \dots$) in sämtlichen \mathfrak{R}_p , also nach Satz 4 auch in \mathfrak{R} enthalten, d. h. α hängt ganz von \mathfrak{R} ab. ~~Es sei ferner α ein von \mathfrak{R}_p ganz abhängiges, aber nicht in \mathfrak{R}_p auftretendes Element, und es sei, was immer möglich, ein zu p teilerfremdes Element a so bestimmt, daß $a \cdot \alpha$ in sämtlichen von \mathfrak{R}_p verschiedenen \mathfrak{R}_p vorkommt. Dann hängt $a \cdot \alpha$ nach dem bereits bewiesenen Teil des Hilfssatzes von \mathfrak{R} ganz ab, gehört aber nicht zu \mathfrak{R}_p , also auch nicht zu \mathfrak{R} .~~ *sh. = \mathfrak{R} ganz abg., so jede \mathfrak{R}_p .*

Satz 6. *\mathfrak{R} ist dann und nur dann ganz abgeschlossen, wenn seine Primär Ideale stets Primidealpotenzen sind.*

Nach Hilfssatz 2 und Satz 3 genügt es, den Beweis für einen nur ein einziges Primideal p enthaltenden \mathfrak{R}_p zu führen.

a) Zu \mathfrak{R}_p existieren dann und nur dann außer den Potenzen von p keine weiteren Primär Ideale, wenn $p = (p)$ Hauptideal ist. Diese Bedingung ist gleichzeitig für die ganze Abgeschlossenheit von \mathfrak{R}_p hinreichend.

Ist $p = (p)$, so läßt sich jedes Element aus \mathfrak{R} in der Form $\alpha = e \cdot p^r$ schreiben, wobei r eine eindeutig bestimmte ganze Zahl und e eine Einheit aus \mathfrak{R}_p bedeutet; α gehört dann und nur dann zu \mathfrak{R}_p , wenn $r \geq 0$. Aus dieser Darstellung der Elemente von \mathfrak{R} ergibt sich mühelos die ganze Abgeschlossenheit von \mathfrak{R}_p sowie die Tatsache, daß p, p^2, p^3, \dots die einzigen Primär Ideale in \mathfrak{R}_p sind. — Gibt es umgekehrt in \mathfrak{R}_p außer den Potenzen von p keine Primär Ideale, und bezeichnet p ein (nach Satz 5 sicher vorhandenes) durch p , aber nicht durch p^2 teilbares Element, so ist $p = (p, p^2)$, da (p, p^2) primär, also $p = (p)$.

b) Es bleibt noch übrig zu zeigen, daß \mathfrak{R}_p sicher nicht ganz abgeschlossen ist, wenn es in \mathfrak{R}_p ein Primär Ideal gibt, das keine Primidealpotenz darstellt.

Wir bezeichnen mit $\frac{0}{p^r}$ das gebrochene Ideal, das aus der Gesamtheit derjenigen Elemente aus \mathfrak{R} besteht, deren Produkt mit einem beliebigen Elemente aus p^r in \mathfrak{R}_p enthalten ist. $\frac{0}{p}$ ist sicher von \mathfrak{R}_p verschieden, denn sonst wäre für beliebiges r auch $\frac{0}{p^r} = 0$ im Widerspruch mit der Tatsache, daß jedes Hauptideal (außer 0) Teiler einer genügend hohen Potenz von p ist. Für das Produkt $p \cdot \frac{0}{p}$ gibt es nur die beiden Möglichkeiten $p \cdot \frac{0}{p} = 0$ und $p \cdot \frac{0}{p} = p$, da in \mathfrak{R}_p das primäre Ideal teilbar

i) oder da \mathfrak{R} Zahlring in p selbst primär.

Ist $p \cdot \frac{o}{p} = o$, so gilt für beliebiges $q \geq p$ die Gleichung $q = p \cdot \left(\frac{o}{p} \cdot q\right)$, und hier stellt $\frac{o}{p} \cdot q = q_1$ ein ganzes Ideal aus \mathfrak{R}_p dar. Es ist also jedes Vielfache von p im Produktsinne durch p teilbar, und daraus folgt leicht, daß die Potenzen von p die einzigen Primärideale in \mathfrak{R}_p sind⁹⁾.

Ist $p \cdot \frac{o}{p} = p$, so ist auch $p \cdot \left(\frac{o}{p}\right)^2 = \left(p \cdot \frac{o}{p}\right) \cdot \frac{o}{p} = p \cdot \frac{o}{p} = p$, $p \cdot \left(\frac{o}{p}\right)^3 = p$ usw. $\frac{o}{p}$ besteht daher aus lauter von \mathfrak{R}_p ganz abhängigen Elementen, \mathfrak{R}_p ist nicht ganz abgeschlossen.

Um weiter zu gelangen, verallgemeinern wir die Definition der ganzen Abgeschlossenheit folgendermaßen: \mathfrak{R} heißt *hinsichtlich des Quotientenringes \mathfrak{S} ganz abgeschlossen*, wenn jedes von \mathfrak{R} algebraisch ganz abhängige Element aus \mathfrak{S} bereits in \mathfrak{R} vorkommt.

Ist \mathfrak{S} irgendein Quotientenring, p ein Primideal von \mathfrak{R} , so verstehen wir unter \mathfrak{S}_p den (nach dem Schema des größten gemeinschaftlichen Teilers gebildeten) Vereinigungsring von \mathfrak{R}_p und \mathfrak{S} . Mit Hilfe der beim Beweise von Satz 4 und Hilfssatz 2 verwandten Schlußweisen zeigt man:

Hilfssatz 3. \mathfrak{S} ist das kleinste gemeinschaftliche Vielfache sämtlicher \mathfrak{S}_p . \mathfrak{R} ist dann und nur dann hinsichtlich \mathfrak{S} ganz abgeschlossen, wenn jedes \mathfrak{R}_p hinsichtlich \mathfrak{S}_p ganz abgeschlossen ist.

Satz 7. Ist \mathfrak{R} hinsichtlich \mathfrak{S} ganz abgeschlossen, so ist \mathfrak{S} regulärer Quotientenring von \mathfrak{R} .

Ansichts von Satz 4 und Hilfssatz 3 brauchen wir unsere Behauptung nur für $\mathfrak{R} = \mathfrak{R}_p$ zu beweisen und haben dann zu zeigen, daß unter den Voraussetzungen unseres Satzes \mathfrak{S} entweder gleich \mathfrak{R}_p oder gleich \mathfrak{R} sein muß. Ist \mathfrak{R}_p absolut (d. h. hinsichtlich \mathfrak{R}) ganz abgeschlossen, so ergibt sich die Richtigkeit dieser Behauptung aus der beim Beweise von Satz 6 unter a) auseinandergesetzten Struktur von \mathfrak{R}_p . Sollte aber \mathfrak{R}_p nicht absolut ganz abgeschlossen sein, so besteht nach dem beim Beweise von Satz 6 unter b) Bemerkten $\frac{o}{p}$ aus lauter ganz von \mathfrak{R}_p abhängigen Elementen; daraus folgt aber, daß $\mathfrak{S} = \mathfrak{R}_p$ sein muß, denn jeder echte (in \mathfrak{R} enthaltene) Erweiterungsring von \mathfrak{R}_p enthält sicher ein in $\frac{o}{p}$, aber nicht in \mathfrak{R}_p vorkommendes Element, weil jedes Element aus \mathfrak{S} in einem der Ideale $\frac{o}{p^r}$ ($r = 1, 2, \dots$) auftritt.

⁹⁾ Man beachte Satz 6 und die Quotientenkettenvoraussetzung, oder auch die Tatsache, daß der Exponent des Primärideals $q \cdot \frac{o}{p}$ mindestens um 1 niedriger ist als der von q .

Satz 7 ist wichtig für die Zahlentheorie, denn man kann mit seiner Hilfe schließen:

Ist \mathfrak{K} ein beliebiger Unterring eines endlichen algebraischen Zahlkörpers, \mathfrak{K}^* derjenige Unterring von \mathfrak{K} , der aus allen in \mathfrak{K} vorkommenden ganzen algebraischen Zahlen besteht, so ist \mathfrak{K} regulärer Quotientenring von \mathfrak{K}^* .

Wir beherrschen also sämtliche Unterringe eines endlichen algebraischen Zahlkörpers, wenn wir den Aufbau der ganzzahligen kennen. Das analoge Ergebnis für algebraische Funktionkörper einer Variablen lautet in der Form ein klein wenig umständlicher, nämlich folgendermaßen:

Es sei \mathfrak{K} ein beliebiger, den Konstantenkörper \mathfrak{K}_0 enthaltender Unterring eines endlichen algebraischen Funktionkörpers einer Variablen, φ sei eine nichtkonstante Funktion aus \mathfrak{K} . Bedeutet dann \mathfrak{K}^* denjenigen Unterring von \mathfrak{K} , der aus allen denjenigen Funktionen aus \mathfrak{K} besteht, die vom Ringe der ganzen rationalen Funktionen in φ mit Koeffizienten aus \mathfrak{K}_0 ganz abhängen, so ist \mathfrak{K} regulärer Quotientenring von \mathfrak{K}^* .

~~Wann diese ganze Aussage~~

§ 4.

wegen der
für den § 4
Satz

Vollständig reduzible verallgemeinerte Abelsche Gruppen¹⁰⁾.

Definition 3. Ein Elementesystem A heißt hinsichtlich eines Operatorbereiches Ω „verallgemeinerte Abelsche Gruppe“ (in Zukunft kurz „v. A. G.“ oder „Gruppe“), wenn folgende beiden Bedingungen erfüllt sind:

1. Die Elemente von A bilden hinsichtlich einer kommutativen, als „Addition“ bezeichneten Operation eine Gruppe.

2. Aus einem beliebigen Elemente a von A kann durch Anwendung eines beliebigen Operators ξ von Ω ein eindeutig bestimmtes Element $b = \xi \cdot a$ von A abgeleitet werden, und es gilt das „distributive Gesetz“ $\xi \cdot (a + b) = \xi \cdot a + \xi \cdot b$.

Das nach 1 vorhandene Einheitselement der Addition bezeichnen wir wie üblich mit 0. Für die gleich aufzuzählenden Verknüpfungen kommen nur solche Gruppen in Betracht, die ein und denselben festen Operatorbereich Ω besitzen und das Einheitselement, also die aus 0 allein bestehende „Nullgruppe N “ gemein haben.

Die Begriffe Untergruppe, Summe ($A_1 + A_2 + \dots$), Durchschnitt ($A_1 \cap A_2 \cap \dots$), elementfremd, direkte Summe ($A_1 \dot{+} A_2 \dot{+} \dots$), Restklassengruppe (A/B) werden ganz analog eingeführt wie in der gewöhnlichen (multiplikativ geschriebenen) Gruppentheorie die Begriffe Untergruppe,

¹⁰⁾ Zu den in diesem und dem folgenden Paragraphen benutzten gruppentheoretischen Begriffsbildungen vgl. K. II § 1 und K. III § 1.

Produkt, Durchschnitt, elementefremd, direktes Produkt, Quotientengruppe. Die Restklassengruppe besitzt den gleichen Operatorenbereich (aber natürlich nicht das gleiche Einheitslement) wie die ursprüngliche Gruppe. Hinsichtlich der Summenbildung sei bemerkt, daß wir auch Summen einer wohlgeordneten¹¹⁾ unendlichen Menge von Summanden zulassen; doch muß jedes Element der Summe als Summe von endlich viel Elementen aus den verschiedenen Summanden darstellbar sein.

Ist A die Summe der endlich oder abzählbar unendlich vielen Gruppen $A^{(1)}, A^{(2)}, A^{(3)}, \dots$ und stellt $A^{(i+1)}$ stets eine echte Obergruppe von $A^{(i)}$ dar, so sprechen wir von einer „Kompositionsreihe“ (Kr.) von A mit den „Gliedern“ $A_1 = A^{(1)}, A_2 = A^{(2)}/A^{(1)}, A_3 = A^{(3)}/A^{(2)}, \dots$ ¹²⁾.

Die Isomorphie zweier Gruppen mit gleichem Operatorenbereich wird in üblicher Weise definiert; nur ist zu beachten, daß auch Isomorphie hinsichtlich der Anwendung der Operatoren gefordert werden muß.

Definition 4. Eine von N verschiedene v. A. G. heißt „irreduzibel“, wenn sie außer N keine echte Untergruppe besitzt, sie heißt „vollständig reduzibel“ (v. red.), wenn sie die direkte Summe von endlich oder unendlich viel irreduziblen Gruppen darstellt.

Gewöhnlich fordert man von einer v. red. Gruppe Darstellbarkeit durch endlich viel irreduzible Summanden. Wir zeigen, daß auch bei unserer allgemeineren Definition die bekannten Hauptsätze in Gültigkeit bleiben.

Hilfssatz 4. Die Summe A einer beliebigen Menge irreduzibler Gruppen ist stets v. red.

Nach Voraussetzung gilt eine Gleichung $A = A_1 + A_2 + \dots + A_r + \dots$ ¹³⁾ mit irreduziblen A_r . Wir setzen $A^{(\tau)} = A_1 + A_2 + \dots + A_\sigma + \dots$ ($\sigma \leq \tau$) und bezeichnen mit τ_1, τ_2, \dots diejenigen Ordnungszahlen, für die A_{τ_σ} nicht in $A^{(\tau_\sigma-1)}$ enthalten ist. Dann wird $A = A_{\tau_1} + A_{\tau_2} + \dots + A_{\tau_\sigma} + \dots$.

Definition 5. V heißt „größte v. red. Untergruppe von A “, wenn V selbst v. red. ist, und wenn es in A keine echte Obergruppe von V gibt, die gleichfalls v. red. wäre.

¹¹⁾ Im folgenden wird die Gültigkeit des Wohlordnungssatzes für alle in Betracht kommenden Mengen vorausgesetzt. Die direkte Summe ist auch bei unendlich viel Komponenten gegenüber der gewöhnlichen Summe dadurch ausgezeichnet, daß jede Komponente zur Summe aller übrigen elementefremd ist.

¹²⁾ Man könnte auf genau dieselbe Weise auch Kompositionsreihen mit einer eventuell nicht abgezählten, wohlgeordnet unendlichen Menge von Gliedern einführen, doch soll hiervon abgesehen werden.

¹³⁾ Wenn wir die Schreibweise $A = A_1 + A_2 + \dots + A_r + \dots$ bzw. $A = A_1 + A_2 + \dots + A_r + \dots$ benutzen, so wollen wir uns immer die Festsetzung getroffen denken, daß einer Limeszahl r keine Komponente entspricht.

Aus Hilfssatz 4 folgt, daß jede Gruppe, die eine irreduzible Untergruppe besitzt, auch eine eindeutig bestimmte „größte v. red. Untergruppe“ enthält.

Satz 8. Jede von N verschiedene Untergruppe U einer v. red. G . A ist selbst v. red., und es gilt eine Gleichung $A = U \dot{+} \bar{U}$.

Es sei

$$A = A_1 \dot{+} A_2 \dot{+} \dots \dot{+} A_\tau \dot{+} \dots \text{ mit irreduziblen } A_\tau,$$

$$A^{(\tau)} = A_1 \dot{+} A_2 \dot{+} \dots \dot{+} A_\sigma \dot{+} \dots \quad (\sigma \leq \tau),$$

$$\bar{A}^{(\tau)} = A_{\tau+1} \dot{+} A_{\tau+2} \dot{+} \dots \dot{+} A_{\tau+\sigma} \dot{+} \dots$$

Satz 8 wird bewiesen sein, wenn wir zeigen können:

Zu jedem τ gibt es ein der Gleichung $A = B^{(\tau)} \dot{+} \bar{A}^{(\tau)}$ genügendes $B^{(\tau)}$ mit folgenden Eigenschaften: a) $B^{(\tau)}$ besitzt eine Darstellung $B^{(\tau)} = B_1 \dot{+} B_2 \dot{+} \dots \dot{+} B_\sigma \dot{+} \dots$ ($\sigma \leq \tau$), bei der B_σ gleich A_σ oder gleich einer zu A_σ isomorphen Untergruppe von U ist. b) Bedeuten $B_{\sigma_1}, B_{\sigma_2}, \dots, B_{\sigma_\rho}, \dots$ die in U enthaltenen Gruppen unter den B_σ , so ist $U \cap B^{(\tau)} = B_{\sigma_1} \dot{+} B_{\sigma_2} \dot{+} \dots \dot{+} B_{\sigma_\rho} \dot{+} \dots$, und es wird $U = (U \cap B^{(\tau)}) \dot{+} (U \cap \bar{A}^{(\tau)})$.

Aus der Definition der $B^{(\tau)}$ folgt, daß für eine Limeszahl τ die Gruppe $B^{(\tau)}$ gleich der Summe aller $B^{(\sigma)}$ ($\sigma < \tau$) ist. Um die Existenz von $B^{(\tau)}$ für beliebiges τ nachzuweisen, braucht man daher nur zu zeigen, daß sich zu gegebenem $B^{(\tau-1)}$ stets ein $B^{(\tau)} = B^{(\tau-1)} \dot{+} B_\tau$ in der gewünschten Weise bestimmen läßt.

Ist $\bar{A}^{(\tau-1)} \cap U = \bar{A}^{(\tau)} \cap U$, so ist die Konstruktion von $B^{(\tau)}$ trivial; man hat einfach $B_\tau = A_\tau$, $B^{(\tau)} = B^{(\tau-1)} \dot{+} A_\tau$ zu setzen. Ist andererseits $\bar{A}^{(\tau-1)} \cap U$ echte Obergruppe von $\bar{A}^{(\tau)} \cap U$, so kann man endlich viel Ordnungszahlen $\sigma_i > \tau$ ($i = 1, 2, \dots, n$) so bestimmen, daß $(A_\tau \dot{+} A_{\sigma_1} \dot{+} \dots \dot{+} A_{\sigma_n}) \cap U = D$ von N verschieden ist, während alle Gruppen, die aus $A_\tau \dot{+} A_{\sigma_1} \dot{+} \dots \dot{+} A_{\sigma_n}$ durch Weglassen einer Komponente entstehen, zu U elementfremd sind. Dann lehren die Ergebnisse von K. II, § 3, daß D zu A_τ isomorph ist, und die Gleichung $D \dot{+} A_{\sigma_1} \dot{+} \dots \dot{+} A_{\sigma_n} = A_\tau \dot{+} A_{\sigma_1} \dot{+} \dots \dot{+} A_{\sigma_n}$ gilt; daraus folgt aber, daß $\bar{A}^{(\tau-1)} = D \dot{+} A_{\tau+1} \dot{+} A_{\tau+2} \dot{+} \dots \dot{+} A_{\tau+\sigma} \dot{+} \dots$ wird, und $B_\tau = D$, $B^{(\tau)} = B^{(\tau-1)} \dot{+} D$ gesetzt werden kann.

Satz 9. Die Restklassengruppe einer v. red. G . nach einer echten Untergruppe ist stets v. red.

Satz 10. Sieht man isomorphe Gruppen als nicht verschieden an, so ist die direkte Summendarstellung einer v. red. G . durch irreduzible Summanden bis auf die Wohlordnung eindeutig bestimmt.

Satz 9 ist ein selbstverständlicher Zusatz zu Satz 8. Satz 10 ergibt sich ohne wesentliche Schwierigkeit aus der folgenden Tatsache, die beim Beweis von Satz 8 durch die Konstruktion der $B^{(\tau)}$ mitbewiesen wurde:

Ist $A = A_1 \dot{+} A_2 \dot{+} \dots \dot{+} A_r \dot{+} \dots$ eine Darstellung von A durch irreduzible Summanden, so besitzt eine von N verschiedene Untergruppe U von A eine Darstellung $U = B_{\tau_1} \dot{+} B_{\tau_2} \dot{+} \dots \dot{+} B_{\tau_r} \dot{+} \dots$, bei der B_{τ_σ} zu A_{τ_σ} isomorph ist.

§ 5.

Die Loewyschen Gruppen und ihre idealtheoretische Bedeutung.

Definition 6. A heißt „Loewysche Gruppe“, wenn A eine ^{kompositionelle} Kompositionsreihe mit lauter v. red. Gliedern besitzt. \vee $\#$

Hilfssatz 5. Ist U eine Untergruppe der Loewyschen Gruppe A , so enthält für $U \neq N$ bzw. $U \neq A$ stets U bzw. A/U eine irreduzible Untergruppe.

Es seien die v. red. G. $A_1 = A^{(1)}$, $A_2 = A^{(2)}/A^{(1)}$, $A_3 = A^{(3)}/A^{(2)}$, ... die Glieder einer Kompositionsreihe von A ; $A_0 = N$. Wählen wir dann für $U \neq N$ den Index $n \geq 0$ so, daß $A^{(n)} \cap U = N$, $A^{(n+1)} \cap U \neq N$ ausfällt, so muß $A^{(n+1)} \cap U$ zu einer Untergruppe von A_{n+1} isomorph und daher v. red. sein; bestimmen wir andererseits im Falle $U \neq A$ die Zahl n derart, daß $A^{(n)}$, aber nicht $A^{(n+1)}$ Untergruppe von U ist, so zeigt sich, daß A/U eine Untergruppe enthält, die v. red. sein muß, weil sie zu einer von der Nullgruppe verschiedenen Restklassengruppe von A_{n+1} isomorph ist.

Gestützt auf Hilfssatz 5 und den Zusatz zu Hilfssatz 4 können wir zu einer Loewyschen Gruppe A folgendermaßen eine ausgezeichnete Kompositionsreihe konstruieren: Wir verstehen unter $L_1 = L^{(1)}$ bzw. unter L_2 die größte v. red. Untergruppe von A bzw. $A/L^{(1)}$, unter $L^{(2)}$ diejenige Untergruppe von A , die der Gleichung $(A/L^{(1)})/L_2 = A/L^{(2)}$ genügt, d. h. die Gesamtheit aller in den Restklassen von L_2 auftretender Elemente aus A ; allgemein definieren wir L_i als größte v. red. Untergruppe von $A/L^{(i-1)}$, $L^{(i)}$ durch die Gleichung $(A/L^{(i-1)})/L_i = A/L^{(i)}$. Ist nun irgendeine Kompositionsreihe von A vorgelegt, deren Glieder $A_1 = A^{(1)}$, $A_2 = A^{(2)}/A^{(1)}$, $A_3 = A^{(3)}/A^{(2)}$, ... sämtlich v. red. sind, so stellt ersichtlich $L^{(i)}$ eine Obergruppe von $A^{(i)}$ dar. Die v. red. G. L_1, L_2, L_3, \dots bilden daher die Glieder einer ausgezeichneten Kompositionsreihe von A , die wir als die „Loewysche Kompositionsreihe“¹⁴⁾ von A bezeichnen wollen. Eine Gruppe, deren Loewysche Kompositionsreihe nur endlich viel Glieder besitzt, soll „endliche Loewysche Gruppe“ heißen.²⁾

¹⁴⁾ Die Loewyschen Kompositionsreihen habe ich in K. III, § 3 eingeführt. Man kommt zu ihnen dadurch, daß man gewisse matrizentheoretische Definitionen von A. Loewy in die Sprache der Gruppentheorie übersetzt. (Vgl. A. Loewy, Über Matrizen- und Differentialkomplexe I—III, Math. Annalen 78 (1917), S. 1—51 bzw. S. 343—368.)

¹⁾ Analogie für nichtkommut. Gr.

²⁾ wann fällt Loewysche v. Jordansche Disk. zus. (für gewöhnl. Ab. Gr. = mit. null.)

Satz 11. Die Summe zweier Loewyschen Gruppen A und B ist selbst eine Loewysche Gruppe. Sind $L_1 = L^{(1)}$, $L_i = L^{(i)} / L^{(i-1)}$ bzw. $M_1 = M^{(1)}$, $M_i = M^{(i)} / M^{(i-1)}$ die Glieder der Loewyschen Kompositionsreihe von A bzw. B , so sind die Glieder der Loewyschen Kompositionsreihe von $A + B$ durch $N_1 = L_1 + M_1$, $N_i = (L^{(i)} + M^{(i)}) / (L^{(i-1)} + M^{(i-1)})$ gegeben.

Der Beweis ergibt sich unmittelbar aus der leicht einzusehenden Tatsache¹⁵⁾, daß $A^* + B^*$ die größte v. red. Untergruppe von $A + B$ darstellt, falls A^* die von A , B^* die von B bedeutet. — Der folgende Satz zeigt die Rolle, die die Loewyschen Kompositionsreihen in der Theorie der gewöhnlichen Abelschen Gruppen spielen:

Satz 12. Eine gewöhnliche Abelsche Gruppe A ist dann und nur dann Loewysch, wenn ihre Elemente sämtlich endliche Ordnung¹⁶⁾ besitzen. Ist diese Bedingung erfüllt, so besteht die zur Loewyschen Kompositionsreihe von A gehörige Gruppe $L^{(i)}$ (abgesehen vom Einheitselement) aus allen und nur den Elementen, in deren Ordnung kein Primfaktor in höherer als i -ter Potenz aufgeht.

Zum Beweis beachte man: α) Eine zyklische Gruppe, deren erzeugendes Element unendlich hohe Ordnung besitzt, enthält überhaupt keine irreduzible Untergruppe. β) Eine zyklische Gruppe, deren erzeugendes Element als Ordnung die r -te Potenz einer Primzahl besitzt, hat eine Loewysche Kompositionsreihe von r Gliedern. γ) Besitzen die Elemente von G sämtlich endliche Ordnung, so gilt eine direkte Summenzerlegung $G = G_1 + G_2 + G_3 + \dots$, bei der die Elemente von G_i als Ordnungen Potenzen einer festen Primzahl p_i haben. Existiert für die Ordnungen aller Elemente von G_i eine obere Schranke, so läßt sich G_i als direkte Summe zyklischer Gruppen darstellen.

Wir wenden uns jetzt unserm Hauptziele zu, nämlich der Untersuchung der idealtheoretischen Bedeutung der endlichen Loewyschen Gruppen. Zu diesem Zweck wählen wir als Gruppen die Ideale eines Ringes \mathfrak{R} , als Operatorenbereich den Ring \mathfrak{R} selber.¹⁷⁾ Die Restklassengruppen nach einem Ideale α aus \mathfrak{R} werden dann durch die Ideale von \mathfrak{R}/α dargestellt; es sei besonders hervorgehoben, daß wir gemäß den Festsetzungen von § 4 auch für die Restklassengruppen nach α als Operatorenbereich den Ring \mathfrak{R} zugrunde legen¹⁷⁾, und nicht, was auf dasselbe herauskäme, den Ring \mathfrak{R}/α .

¹⁵⁾ Vgl. K. III, § 3, Satz 8.

¹⁶⁾ Einem Element, dessen sämtliche Vielfachen verschieden sind, schreiben wir als Ordnung das Symbol ∞ zu. Das ist zweckmäßig, weil wir hier unter der Ordnung eine Zahl, und nicht, wie in Definition 7, ein Ideal verstehen.

¹⁷⁾ Unter dem Produkt des Elementes α mit der Restklasse $\bar{\beta}$ verstehen wir das Produkt der durch α bestimmten Restklasse mit $\bar{\beta}$.

1) eine Id. ist als Gruppe irreduzibel, wenn es keine echten ^{Multipl.} ~~Teile~~ enthält

Definition 7. Unter der „Ordnung“¹⁸⁾ eines als Gruppe aufgefaßten Ideals α verstehen wir das Ideal $\alpha \cdot \mathfrak{R}$, das aus der Gesamtheit derjenigen Operatoren besteht, die mit einem beliebigen Element aus α komponiert das Nullelement liefern.

Hilfssatz 6. Zwei Hauptideale sind als Gruppen dann und nur dann isomorph, wenn sie gleiche Ordnung besitzen.

Die angegebene Bedingung ist offenbar notwendig. Sie ist aber auch hinreichend. Denn es besteht ein Hauptideal (α) mit der Ordnung m aus der Gesamtheit aller Elemente von der Form $\xi \cdot \alpha$, und es ist $\xi_1 \cdot \alpha = \xi_2 \cdot \alpha$ dann und nur dann, wenn $\xi_1 \equiv \xi_2 (m)$; daraus folgt aber sofort, daß zwei Hauptideale (α) und (β) mit gleicher Ordnung durch Zuordnung von α und β gruppenisomorph aufeinander abgebildet werden.

Hilfssatz 7. α ist als Gruppe dann und nur dann irreduzibel, wenn α Hauptideal ist, und die Ordnung von α ein Primideal \mathfrak{p} ohne von \mathfrak{o} verschiedenen echten Teiler darstellt. Zwei irreduzible Gruppen mit gleicher Ordnung sind isomorph.

Ist α irreduzibel, und bedeutet $\alpha \neq 0$ ein Element aus α , so ist (α) eine von n verschiedene Untergruppe von α , und es gilt daher die Gleichung $\alpha = (\alpha)$. Stellt ferner \mathfrak{q} einen echten Teiler der Ordnung \mathfrak{p} von α dar, so ist wegen $\mathfrak{q} \cdot (\alpha) \neq n$ sicher $\mathfrak{q} \cdot (\alpha) = (\alpha)$, es gilt also eine Gleichung $\beta \cdot \alpha = \alpha$; $(\beta) \geq \mathfrak{q}$, und daraus folgt $(\beta - \varepsilon) \cdot \alpha = 0$; $(\beta - \varepsilon) \geq \mathfrak{p} > \mathfrak{q}$; $\mathfrak{q} + \mathfrak{p} = \mathfrak{q} = \mathfrak{o}$. Daß ein Hauptideal, dessen Ordnung außer \mathfrak{o} keinen echten Teiler besitzt, als Gruppe irreduzibel sein muß, ist klar. Die auf die Isomorphie bezügliche Bemerkung von Hilfssatz 7 folgt aus Hilfssatz 6.

Satz 13. Der Ring \mathfrak{R} ist dann und nur dann ein hyperkomplexes System bzw. ein Zahlring, wenn \mathfrak{R} bzw. \mathfrak{R}/α für beliebiges $\alpha \neq n$ eine endliche Loewysche Gruppe ist, die die Eigenschaft hat, daß unter den irreduziblen Komponenten der Glieder ihrer Loewyschen Kompositionsreihe nur endlich viele nichtisomorphe Typen auftreten.

Der Beweis braucht nur für hyperkomplexe Systeme geführt zu werden.

a) Es sei \mathfrak{R} ein Ring, der den gruppentheoretischen Voraussetzungen von Satz 13 genügt, n sei die Gliederzahl der Loewyschen Kompositionsreihe von \mathfrak{R} . Nach Hilfssatz 7 besitzen die irreduziblen Komponenten

¹⁸⁾ An Stelle von „Ordnung“ braucht H. Grell in G. das Wort „Periode“. Ich möchte lieber die Bezeichnung „Ordnung“ beibehalten, obwohl wir uns dadurch in einen gewissen Gegensatz zu der in der Theorie der gewöhnlichen endlichen Abelschen Gruppen üblichen Ausdruckweise setzen. (Dort versteht man bekanntlich unter der Ordnung einer Gruppe die Anzahl ihrer Elemente, während man nach unserer Festsetzung unter der Ordnung das kleinste gemeinschaftliche Vielfache der Ordnungen der Elemente verstehen mußte.) Vgl. auch K. III, S. 23, Anm. 1.

*Man folgt aus bloßer Sicht d. endl. Loewyschen Kompos. d. „...“
 „...“
 „...“
 „...“*

der Glieder dieser Kompositionsreihen nur endlich viel verschiedene Ordnungen p_1, p_2, \dots, p_m , und aus der Definition der Kompositionsreihe und der Ordnung folgt $p_1^{n_1} \cdot p_2^{n_2} \cdot \dots \cdot p_m^{n_m} = n$. Aus dieser letzteren Gleichung schließt man leicht, daß \mathfrak{R} ein hyperkomplexes System ist.

b) Jedes hyperkomplexe System läßt sich als direkte Summe von endlich vielen, nur ein einziges Primideal enthaltenden „primären“ hyperkomplexen Systemen darstellen¹⁹⁾. Nach Satz 11 ist daher die gruppentheoretische Voraussetzung von Satz 13 bei einem hyperkomplexen System stets erfüllt, wenn sie nur für jedes primäre hyperkomplexe System zutrifft. Das ist aber in der Tat der Fall. Ist nämlich \mathfrak{R} ein primäres hyperkomplexes System, dessen einziges Primideal \mathfrak{p} der Gleichung $\mathfrak{p}^n = n$, $\mathfrak{p}^{n-1} \neq n$ genügt, so besitzt \mathfrak{R} eine endliche Loewysche Kompositionsreihe mit den Gliedern $L_i = (n : \mathfrak{p}^i) / (n : \mathfrak{p}^{i-1})$ ($i = 1, 2, \dots, n$), und die irreduziblen Komponenten sämtlicher L_i sind isomorph, weil sie alle die gleiche Ordnung \mathfrak{p} haben.

Definition 8. *Eine Kompositionsreihe heißt Jordansch, wenn sie aus endlich vielen Gliedern besteht, die sämtlich irreduzibel sind.*

Aus der bekannten Invarianz der Gliederzahl einer Jordanschen Kompositionsreihe ergibt sich:

Hilfssatz 8. *Eine Gruppe mit Jordanscher Kompositionsreihe ist stets eine endliche Loewysche Gruppe. — Eine endliche Loewysche Gruppe besitzt dann und nur dann eine Jordansche Kompositionsreihe, wenn jedes Glied ihrer Loewyschen Kompositionsreihe die direkte Summe von endlich viel irreduziblen Komponenten darstellt.*

Um die idealtheoretische Bedeutung der Jordanschen Kompositionsreihen formulieren zu können, müssen wir — analog wie in § 1 Quotienten- und Produktketten — Teiler- und Vielfachenketten einführen. Offenbar zieht das endliche Abbrechen sämtlicher Teiler- bzw. Vielfachenketten das endliche Abbrechen sämtlicher Quotienten- bzw. Produktketten nach sich. Aus den Ergebnissen von N. § 10 und K. III, § 2, können wir den folgenden Satz übernehmen:

Hilfssatz 9. *\mathfrak{R} besitzt dann und nur dann als Gruppe eine Jordansche Kompositionsreihe, wenn in \mathfrak{R} sämtliche Teiler- und Vielfachenketten im Endlichen abbrechen.*

Ein Ideal, das sich nicht als kleinstes gemeinschaftliches Vielfaches von echten Teilern darstellen läßt, heißt bei E. Noether „irreduzibel“. Wir wollen es „ideal-irreduzibel“ nennen, um jedes Mißverständnis mit dem bei Gruppen eingeführten Irreduzibilitätsbegriff zu vermeiden.

¹⁹⁾ Vgl. K. I, § 1, Satz 2, sowie die dort in Anm. *) angegebene Literatur.

Hilfssatz 10. *In einem Ring \mathfrak{R} , der eine Loewysche Kompositionsreihe mit der größten v. red. Untergruppe \mathfrak{v} besitzt, ist \mathfrak{n} dann und nur dann ideal-irreduzibel, wenn \mathfrak{v} als Gruppe irreduzibel ist.*

Einer etwa vorhandenen direkten Summendarstellung von \mathfrak{v} durch endlich viel irreduzible Untergruppen entspricht eine kleinste-gemeinschaftliche-Vielfachendarstellung von \mathfrak{n} durch ebensoviel ideal-irreduzible Teiler.

Ist \mathfrak{v} nicht irreduzibel, so gilt eine Gleichung $\mathfrak{v} = \mathfrak{v}_1 + \mathfrak{v}_2$, und wir haben in $\mathfrak{n} = \mathfrak{v}_1 \cap \mathfrak{v}_2$ eine kleinste-gemeinschaftliche-Vielfachendarstellung von \mathfrak{n} durch echte Teiler. Ist aber \mathfrak{v} irreduzibel, so muß \mathfrak{v} nach Hilfssatz 5 durch alle von \mathfrak{n} verschiedenen Ideale teilbar sein, und es kann daher keine Gleichung $\mathfrak{n} = \mathfrak{t}_1 \cap \mathfrak{t}_2$, $\mathfrak{t}_1 \neq \mathfrak{n}$, $\mathfrak{t}_2 \neq \mathfrak{n}$ bestehen.

Zum Beweise des zweiten Teiles von Hilfssatz 10 vgl. K. III, Satz 9b und Satz 11, sowie die dort auf S. 15 zitierte Arbeit von E. Noether.

Satz 14. *Ein hyperkomplexes System bzw. ein Zahlring besitzt dann und nur dann die Eigenschaft, daß in \mathfrak{R} selbst jede Teilerkette und in \mathfrak{R}/α für beliebiges α bzw. für beliebiges $\alpha \neq \mathfrak{n}$ jede Vielfachenkette im Endlichen abbricht, wenn jedes Ideal aus \mathfrak{R} als kleinstes gemeinschaftliches Vielfaches von endlich vielen, ideal-irreduziblen Teilern dargestellt werden kann.*

Der Beweis folgt leicht aus den Hilfssätzen 8, 9 und 10. Satz 14 ist vor allem deshalb wichtig, weil in jedem Unterring eines endlichen algebraischen Zahlkörpers alle Teilerketten und modulo $\alpha \neq \mathfrak{n}$ alle Vielfachenketten im Endlichen abbrechen.

(Eingegangen am 6. 2. 1927.)

Über die simultane Approximation von Irrationalzahlen.

Zweite Mitteilung¹⁾.

Von

Ph. Furtwängler in Wien.

Der Hauptzweck der folgenden Mitteilung ist die Aufstellung eines Algorithmus zur simultanen Approximation von zwei Irrationalzahlen, der als eine Verallgemeinerung der gewöhnlichen Kettenbruchentwicklung für eine einzelne Irrationalität aufgefaßt werden kann. Um ihn zu entwickeln, ist zunächst in der Ebene die Reduktion von Parallelogittern kurz zu erörtern, wenn zur Definition der Strahldistanz als Eichfigur ein Quadrat mit dem Anfangspunkt als Mittelpunkt und mit Seiten, die den Achsen parallel laufen, benutzt wird. Auf Grund dieser Erörterungen gelingt es dann im Anschluß an die geometrischen Ideen Minkowskis den gewünschten Algorithmus zu erhalten, wenn als Eichkörper im Raume ein Prisma mit quadratischem Querschnitt benutzt wird.

Vorweggeschickt ist eine Ergänzung zum § 2 der ersten Mitteilung, um eine dort vorhandene Lücke auszufüllen, auf die mich Herr Perron freundlichst aufmerksam gemacht hat²⁾. Die Lücke ist erst bei der endgültigen Redaktion des Aufsatzes entstanden, bei der ich den Beweis vereinfachen zu können glaubte. Ich gebe daher hier im § 1 meinen ursprünglichen Beweis wieder; an den Resultaten der ersten Mitteilung ist nichts zu ändern.

¹⁾ Die erste Mitteilung ist in *Math. Annalen* 96 (1926), S. 169–175 erschienen.

²⁾ Auch Herr A. Scholz in Berlin hat die Lücke bemerkt und einen Weg zu ihrer Beseitigung gefunden, der von dem im folgenden entwickelten verschieden ist.

§ 1.

Wir stützen uns auf folgenden Approximationssatz.

Satz 1. *Die Beziehungen*

$$(1) \quad \begin{cases} a_{11} x_1 + \dots + a_{1n} x_n + l_1 = O(1) \\ \dots \\ a_{n-1,1} x_1 + \dots + a_{n-1,n} x_n + l_{n-1} = O(1) \\ a_{n,1} x_1 + \dots + a_{n,n} x_n + l_n = O(g) \end{cases}$$

sind stets durch ganzzahlige x_i zu erfüllen, wenn die a_{ik} reelle Konstanten mit nicht verschwindender Determinante sind und die l_i von der Größenordnung 1 sind; g bedeutet eine beliebig große positive Zahl.

Beweis. Man kann offenbar (1) ersetzen durch

$$(2) \quad \begin{cases} y_1 = a_{11} x_1 + \dots + a_{1n} x_n = O(1) \\ \dots \\ y_{n-1} = a_{n-1,1} x_1 + \dots + a_{n-1,n} x_n = O(1) \\ y_n = a_{n,1} x_1 + \dots + a_{n,n} x_n = O(g). \end{cases}$$

Nach bekannten Approximationssätzen besitzen die ersten $n - 1$ Gleichungen ganzzahlige Lösungen mit beliebig großen x_i . Würde für alle diese Lösungssysteme $|y_n|$ geschränkt sein, so würde die Geschränktheit der $|x_i|$ folgen, was zu einem Widerspruch führt.

Wir behalten im folgenden durchweg die Bezeichnungen der ersten Mitteilung bei. Es sei der Körper k reell und die konjugierten $k'' \dots k^{(n)}$ seien so angeordnet, daß zwei konjugiert komplexe Körper aufeinander folgen. Die Aufgabe ist, die Basiszahlen $\omega_1, \omega_2, \dots, \omega_n$ so zu normieren, daß der Quotient Q höchstens beliebig wenig unter 1 liegt. Dies wollen wir dadurch erreichen, daß wir ein Glied im Zähler und Nenner, und zwar in beiden Fällen dasselbe Glied so groß machen, daß es alle übrigen Glieder beliebig stark überwiegt. Wir nehmen nun an, daß die ersten $i - 1$ Basiszahlen, die wir normiert mit $\tau_1, \tau_2, \dots, \tau_{i-1}$ bezeichnen, bereits normiert seien, und zwar derart, daß für die noch nicht normierten Zahlen $\omega_i \dots \omega_n$ gilt:

$$(3) \quad |\omega_j^{(i)}| = O(1) \quad \left(\begin{array}{l} j = i, i+1, \dots, n \\ l = 2, 3, \dots, n \end{array} \right),$$

und zeigen, wie die Normierung fortzusetzen ist.

1. $k^{(i+1)}$ sei ein reeller Körper.

Wir ersetzen nur ω_i durch

$$(4) \quad \tau_i = \omega_i + \tau_1 u_1 + \dots + \tau_{i-1} u_{i-1} + \omega_{i+1} u_i + \dots + \omega_n u_{n-1}$$

und bestimmen die u_i ganzzahlig so, daß

$$(5) \quad |\tau_i^{(l)}| = O(1) \quad (l = 2, 3, \dots, i, i+2, \dots, n), \quad |\tau_i^{(i+1)}| = O(g),$$

was nach Satz 1 möglich ist. Da alle Basiszahlen außer ω_i ungeändert bleiben, ist dieser Fall erledigt.

2. $k^{(i+1)}$ und $k^{(i+2)}$ seien konjugiert komplex.

Dieser Fall ist etwas komplizierter zu behandeln, da gleichzeitig zwei Basiszahlen geändert werden. Es handelt sich hier darum, bei einem Ausdruck von der Bauart

$$\alpha = \frac{2|\varrho_1\sigma_2 - \varrho_2\sigma_1|}{(\varepsilon_1\varrho_1 + \varepsilon_2\varrho_2)^2 + (\varepsilon_1\sigma_1 + \varepsilon_2\sigma_2)^2}$$

durch geeignete Verfügung über die ϱ, σ zu erreichen, daß er höchstens beliebig wenig unter 1 sinkt, wobei $|\varepsilon_1| \leq 1, |\varepsilon_2| \leq 1$ ist. Fordert man zunächst:

$$\varrho_1 - \sigma_2 = O(1), \quad \varrho_2 + \sigma_1 = O(1),$$

so wird

$$\begin{aligned} \alpha &\geq \frac{2[\varrho_1^2 + \sigma_1^2 + \varrho_1 O(1) + \sigma_1 O(1)]}{[\varepsilon_1\varrho_1 - \varepsilon_2\sigma_1 + O(1)]^2 + [\varepsilon_1\sigma_1 + \varepsilon_2\varrho_1 + O(1)]^2} \\ &= \frac{2[\varrho_1^2 + \sigma_1^2 + \varrho_1 O(1) + \sigma_1 O(1)]}{(\varepsilon_1^2 + \varepsilon_2^2)(\varrho_1^2 + \sigma_1^2) + \varrho_1 O(1) + \sigma_1 O(1)}. \end{aligned}$$

Es genügt also $\varrho_1^2 + \sigma_1^2$, d. h. ϱ_1 oder σ_1 von der Ordnung g zu machen, um das gewünschte Ziel zu erreichen.

Wir ersetzen jetzt ω_i und ω_{i+1} durch

$$\begin{aligned} \tau_i &= \omega_i + u_1\tau_1 + \dots + u_{i-1}\tau_{i-1} + u_i\omega_{i+1} + u_{i+1}\omega_{i+2} + \dots + u_{n-1}\omega_n \\ &= \varrho_i + i\sigma_i, \\ \tau_{i+1} &= \omega_{i+1} + v_1\tau_1 + \dots + v_{i-1}\tau_{i-1} + v_i\omega_{i+2} + \dots + v_{n-2}\omega_n \\ &= \varrho_{i+1} + i\sigma_{i+1}, \end{aligned}$$

indem wir die übrigen Basiszahlen vorläufig ungeändert lassen, und fordern als Normierungsbedingungen:

$$(6) \quad \begin{cases} |\tau_i^{(l)}| = O(1), & |\tau_{i+1}^{(l)}| = O(1) \quad (l = 2, 3, \dots, i, i+3, \dots, n), \\ |\varrho_i^{(i+1)} - \sigma_{i+1}^{(i+1)}| = O(1), & |\sigma_i^{(i+1)} + \varrho_{i+1}^{(i+1)}| = O(1), \end{cases}$$

$$(7_1) \quad |\varrho_i^{(i+1)}| = O(g)$$

oder

$$(7_2) \quad |\sigma_{i+1}^{(i+1)}| = O(g).$$

Die Bedingungssysteme (6), (7₁) und (6), (7₂) haben nach Satz 1 stets Lösungen in ganzzahligen u, v , wenn die in jedem System auftretenden

^{2a}) Für $i = 1$ läuft l von 3 bis n .

$2n - 3$ linearen homogenen Polynome linear unabhängig sind. Die Determinanten dieser beiden Polynomsysteme zerfallen beide in das Produkt einer $(n - 1)$ - und einer $(n - 2)$ -zeiligen Determinante. Von der ersten steht ohne weiteres fest, daß sie von Null verschieden ist. Es ist daher nur zu prüfen, ob die beiden auftretenden $(n - 2)$ -zeiligen Determinanten gleichzeitig verschwinden können. Es muß dann

$$(8) \quad \begin{vmatrix} \tau_1^{(2)} & \dots & \tau_{i-1}^{(2)} & , & \omega_{i+2}^{(2)} & \dots & \omega_n^{(2)} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \tau_1^{(i)} & \dots & \tau_{i-1}^{(i)} & , & \omega_{i+2}^{(i)} & \dots & \omega_n^{(i)} \\ \tau_1^{(i+2)} & \dots & \tau_{i-1}^{(i+2)} & , & \omega_{i+2}^{(i+2)} & \dots & \omega_n^{(i+2)} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \tau_1^{(n)} & \dots & \tau_{i-1}^{(n)} & , & \omega_{i+2}^{(n)} & \dots & \omega_n^{(n)} \end{vmatrix} = 0$$

sein. Man kann nun zu jeder Basiszahl ω_i oder ω_{i+1} addieren; dadurch bleiben die bisher normierten Basiszahlen normiert und die noch nicht normierten behalten samt ihren konjugierten die Größenordnung 1, da die konjugierten von ω_i und ω_{i+1} von der Größenordnung 1 sind. Wären nun alle entsprechenden neuen Determinanten Null, so müßte die Matrix

$$\begin{pmatrix} \tau_1^{(2)} & \dots & \tau_{i-1}^{(2)} & , & \omega_i^{(2)} & , & \omega_{i+1}^{(2)} & \dots & \omega_n^{(2)} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \tau_1^{(i)} & \dots & \tau_{i-1}^{(i)} & , & \omega_i^{(i)} & , & \omega_{i+1}^{(i)} & \dots & \omega_n^{(i)} \\ \tau_1^{(i+2)} & \dots & \tau_{i-1}^{(i+2)} & , & \omega_i^{(i+2)} & , & \omega_{i+1}^{(i+2)} & \dots & \omega_n^{(i+2)} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \tau_1^{(n)} & \dots & \tau_{i-1}^{(n)} & , & \omega_i^{(n)} & , & \omega_{i+1}^{(n)} & \dots & \omega_n^{(n)} \end{pmatrix}$$

den Rang $n - 3$ haben, was nicht möglich ist. Damit ist die Ergänzung zu § 2 der ersten Mitteilung gegeben.

§ 2.

Um die Entwicklung eines Algorithmus zur simultanen Approximation von zwei Irrationalzahlen vorzubereiten, führen wir zunächst eine Betrachtung über Parallelengitter in der Ebene durch. Es seien

$$(9) \quad \xi = \xi_1 x + \xi_2 y, \quad \eta = \eta_1 x + \eta_2 y$$

zwei reelle Linearformen mit nicht verschwindender Determinante. Wir interpretieren ξ, η als rechtwinklige Koordinaten in der Ebene und zeichnen das Gitter der Punkte mit ganzzahligen x, y . Wir definieren dann als Entfernung eines Punktes $P(x, y)$ vom Anfangspunkt:

$$(10) \quad \mu(P) = \mu(x, y) = \text{Max}(|\xi|, |\eta|).$$

Punkte mit gleicher Entfernung vom Anfangspunkt O liegen auf Quadraten

mit O als Mittelpunkt und mit Seiten parallel den Achsen. Wir nennen nun das Formenpaar (9) oder die „definite Form“ $\mu(x, y)$ reduziert, wenn unter den Gitterpunkten die Punkte $(1, 0)$ und $(0, 1)$ die ersten beiden Entfernungsminima von O liefern. Es zeigt sich dann, daß die Reduktionsbedingungen ebenso wie für den gewöhnlichen Entfernungsbegriff $\sqrt{\xi^2 + \eta^2}$ lauten:

$$(11) \quad \mu(1, 0) \leq \mu(0, 1) \leq \mu(1, \pm 1).$$

Daß diese Bedingungen notwendig sind, ist evident. Um zu zeigen, daß sie hinreichend sind, kann man sich auf folgenden Satz stützen, dessen Beweis der Hauptzweck unserer Betrachtungen ist.

Satz 2. *Ist das Formenpaar (9) reduziert, d. h. sind die Bedingungen (11) erfüllt, so folgt aus der Ungleichung $\mu(x, y) < \mu(0, 1)$, daß $|y| < 2$ ist und daß $|x| < 1$ ist, wenn $|y| \geq 1$ ist. Dabei brauchen x, y nicht ganzzahlig zu sein.*

Beweis. Sind nicht alle Koeffizienten ξ_i, η_i von Null verschieden, so liegt die Richtigkeit des Satzes auf der Hand; es sei also $\xi_1 \xi_2 \eta_1 \eta_2 \neq 0$. Es folgt dann sofort aus (11):

$$(12) \quad \operatorname{sgn} \xi_1 \xi_2 \neq \operatorname{sgn} \eta_1 \eta_2.$$

Es möge nun $\mu(0, 1) = |\xi_2|$ sein, was durch eventuelle Vertauschung von ξ und η stets erreicht werden kann. Damit $\mu(x, y) < \mu(0, 1)$ sei, muß dann gelten:

$$(13) \quad \text{a) } |\xi_1 x + \xi_2 y| < |\xi_2|, \quad \text{b) } |\eta_1 x + \eta_2 y| < |\xi_2|.$$

Ferner folgt aus der Reduktionsbedingung $\mu(1, \pm 1) \geq \mu(0, 1)$, daß

$$(14) \quad |\eta_1| + |\eta_2| \geq |\xi_2|$$

sein muß. Wäre nun $|y| \geq 2$, so würde aus (13a) $\operatorname{sgn} xy \neq \operatorname{sgn} \xi_1 \xi_2$ und daher $\operatorname{sgn} xy = \operatorname{sgn} \eta_1 \eta_2$ folgen. Ferner müßte $|x| > 1$ sein, da $|\xi_2| \geq |\xi_1|$ ist. Dann würde aber aus (13b) und (14) folgen

$$|\eta_1 x + \eta_2 y| = |\eta_1 x| + |\eta_2 y| > |\eta_1| + |\eta_2| \geq |\xi_2|.$$

Es muß also $|y| < 2$ sein. Weiter sieht man, daß auch die Ungleichungen $|y| \geq 1, |x| \geq 1$ nicht gleichzeitig gelten können. Denn aus (13a) würde $\operatorname{sgn} xy = -\operatorname{sgn} \xi_1 \xi_2 = \operatorname{sgn} \eta_1 \eta_2$ folgen und daher wäre $|\eta_1 x + \eta_2 y| \geq |\xi_2|$.

Damit ist unser Satz bewiesen, aus dem sofort folgt, daß die Reduktionsbedingungen (11) auch hinreichend sind. Die Möglichkeit der Reduktion liegt auf der Hand.

§ 3.

Wir brauchen noch den folgenden Satz über Raumpgitter.

Satz 3. Sind P, Q, R drei Gitterpunkte, die mit dem Anfangspunkt O nicht in einer Ebene liegen, und enthält das durch P, Q, R und ihre Gegenpunkte bestimmte Oktaeder außer O und den Eckpunkten keine Gitterpunkte im Innern und auf der Begrenzung, so hat das durch die drei Kanten OP, OQ, OR bestimmte Parallelepipèd denselben oder den doppelten Inhalt wie das Fundamentalparallelepiped des Gitters.

Beweis. Nach Voraussetzung ist jeder Gitterpunkt in der Gestalt $lP + mQ + nR$ mit rationalen l, m, n darstellbar. Ich behaupte, daß dabei nur der Nenner 2 auftreten kann. Das Oktaeder ist durch $|l| + |m| + |n| \leq 1$ gegeben. Angenommen, es wäre

$$\frac{lP + mQ + nR}{p}$$

ein Gitterpunkt, wo p eine ungerade Primzahl und wenigstens eine der ganzen Zahlen l, m, n , etwa l , nicht durch p teilbar ist. Da mit P auch qP Gitterpunkt ist, wo q eine beliebige ganze Zahl bedeutet, muß dann auch ein Gitterpunkt $\frac{P + \bar{m}Q + \bar{n}R}{p}$ existieren, wo $|\bar{m}|$ und $|\bar{n}| < \frac{p}{2}$

sind. Da $\frac{|1 + \bar{m} + \bar{n}|}{p} \leq 1$ ist, müßte dieser Gitterpunkt in das Oktaeder oder auf die Begrenzung fallen, was zu einem Widerspruch führt. Wir

betrachten zweitens den Nenner 4 und nehmen an, daß $\frac{lP + mQ + nR}{4}$ ein Gitterpunkt wäre, wo etwa l ungerade sei. Es müßte dann auch ein Gitterpunkt $\frac{P + \bar{m}Q + \bar{n}R}{4}$ existieren, wo $|\bar{m}|$ und $|\bar{n}| \leq 2$ ist. Da

$|\bar{m}| = |\bar{n}| = 2$ offenbar ausgeschlossen ist, würde dieser Gitterpunkt wieder in das Oktaeder oder auf seine Begrenzung fallen. Es können also nur Gitterpunkte von der Gestalt $\frac{lP + mQ + nR}{2}$ auftreten. Da die Punkte

$\frac{P+Q}{2}, \frac{P+R}{2}, \frac{Q+R}{2}$ keine Gitterpunkte sind, bleibt nur der Fall $l \equiv m \equiv n$

(2). Ist daher $\frac{P+Q+R}{2}$ kein Gitterpunkt, so bestimmen OP, OQ, OR ein Fundamentalparallelepiped; ist dagegen $\frac{P+Q+R}{2}$ ein Gitterpunkt, so bestimmen die Verbindungsstrecken von O mit $P, Q, \frac{P+Q+R}{2}$ ein solches. Damit ist unser Satz bewiesen.

§ 4.

Wir erörtern den Algorithmus zur simultanen Approximation zweier Irrationalitäten zunächst in geometrischer Gestalt. Es seien

$$(15) \quad \xi = \alpha_1 x + \beta_1 y + \gamma_1 z, \quad \eta = \alpha_2 x + \beta_2 y + \gamma_2 z, \quad \zeta = \alpha_3 x + \beta_3 y + \gamma_3 z$$

drei reelle Linearformen mit der Determinante ± 1 ³⁾). Später werden wir speziell

$$(16) \quad \xi = x - \alpha z, \quad \eta = y - \beta z, \quad \zeta = z$$

setzen, wo α, β die beiden zu approximierenden Irrationalitäten sind, zwischen denen keine Relation $g_1 \alpha + g_2 \beta + g_3 = 0$ mit ganzzahligen g_i bestehen soll.

Wir interpretieren ξ, η, ζ als rechtwinklige Raumkoordinaten und denken uns das Raumgitter aller Punkte mit ganzzahligen x, y, z hergestellt. Es soll nun die ζ -Achse durch Gitterpunkte approximiert werden, wobei wir annehmen, daß auf der ζ -Achse außer O keine Gitterpunkte liegen. Hat ein Punkt P die Koordinaten ξ, η, ζ , so soll $\mu(P) = \text{Max}(|\xi|, |\eta|)$ seine Achsendistanz und $h(P) = |\zeta|$ seine Höhe heißen. Wir nehmen zur Vereinfachung der Diskussion noch weiter an, daß außerhalb der Ebene $\zeta = 0$ weder Gitterpunkte mit verschwindenden ξ oder η noch solche mit gleicher Achsendistanz vorhanden seien. Bei den Gittern (16) ist dies stets erfüllt.

Zur Approximation bedienen wir uns vierseitiger Prismen mit dem Mittelpunkt in O und der ζ -Achse als Längsachse; ihr Querschnitt soll ein Quadrat sein, dessen Seiten der ξ - und η -Achse parallel laufen. Ein derartiges Prisma ist durch die Ungleichungen

$$|\xi| \leq \varrho, \quad |\eta| \leq \varrho, \quad |\zeta| \leq \tau$$

bestimmt, wo ϱ, τ positive Konstanten bedeuten. Wenn im folgenden von Prismen die Rede ist, sind immer Prismen der angegebenen Art gemeint.

Ist eine beliebige Anzahl von Gitterpunkten gegeben, so bezeichnen wir als zugehöriges Prisma das kleinste Prisma, das alle Gitterpunkte im Innern und auf der Begrenzung enthält. Wenn das zu einem Gitterpunkt P gehörige Prisma im Innern und auf der Begrenzung keine anderen Gitterpunkte als O, P und seinen Gegenpunkt $-P$ enthält, soll P ein *Näherungspunkt* heißen; mit P ist also stets auch $-P$ Näherungspunkt. Das Problem ist nun, die Reihe der Näherungspunkte, die man sich nach wachsender Höhe geordnet denken kann und bei denen man sich auf diejenigen mit positiven ζ beschränken kann, zu ermitteln; bei unseren Annahmen bricht diese Reihe nicht ab.

Um den Algorithmus möglichst einfach darstellen zu können, brauchen wir noch gewisse Hilfspunkte, die wir so definieren. Es sei P ein Näherungspunkt und Q ein Punkt mit größerer Achsendistanz außerhalb der Geraden OP von solcher Art, daß das zu Q gehörige Prisma außer $\pm Q$

³⁾ Es ist nur wesentlich, daß die Determinante nicht Null ist.

höchstens Gitterpunkte der Geraden OP enthält; dies Prisma werde als *Näherungsprisma* bezeichnet. Es sei endlich R ein solcher Punkt, daß OP, OQ, OR ein Fundamentalparallelepiped definieren⁴⁾ und $h(R)$ nicht größer als die größere der Höhen $h(P), h(Q)$ ist; drei Punkte von der Art P, Q, R werden dann als Näherungstriplel bezeichnet. Die Durchführung des Algorithmus beruht nun auf folgendem Umstand. Es sei durch P, Q ein Näherungsprisma bestimmt, das wir in der Höhe wachsen lassen, bis es zum ersten Male an einen Gitterpunkt S anstößt. Dieser liegt dann entweder in der Ebene OPQ oder P, Q, S bilden ein Fundamentaltripel. Die Punkte P, S oder S, P bestimmen dann wieder ein Näherungsprisma; im letzten Falle, d. h. wenn $\mu(S) < \mu(P)$ ist, ist S der auf P folgende Näherungspunkt. Anderenfalls wird das Verfahren fortgesetzt, das schließlich zu dem nächsten Näherungspunkt führen muß. Gleichzeitig ist klar, daß man dabei keinen Näherungspunkt übergehen kann.

Um unsere Behauptung bezüglich des Punktes S zu beweisen, nehmen wir an, daß S nicht in der Ebene OPQ liegt; es ist dann zu zeigen, daß P, Q, S ein Fundamentaltripel bilden. Das durch die Punkte P, Q, S und ihre Gegenpunkte bestimmte Oktaeder enthält außer E und den Eckpunkten keine Gitterpunkte. Nach Satz 3 bilden dann P, Q, S ein Fundamentaltripel, wenn nicht $\frac{\pm P \pm Q \pm S}{2}$ Gitterpunkte sind, was tatsächlich nicht der Fall ist. Wir betrachten die drei Punkte $\frac{P+Q-R}{2}, \frac{P-Q+R}{2}, \frac{-P+Q+R}{2}$, deren Höhe $< h(S)$ ist. Sind die Koordinaten von P, Q, S resp. ξ_i, η_i, ζ_i ($i=1, 2, 3$) und sollen die genannten drei Punkte nicht in das Prisma (Q, S) fallen, muß

$$(17) \quad \mu\left(\frac{P+Q-R}{2}\right) > \mu(Q), \quad \mu\left(\frac{P-Q+R}{2}\right) > \mu(Q), \\ \mu\left(\frac{-P+Q+R}{2}\right) > \mu(Q)$$

sein. Diese Ungleichungen sind nur erfüllt, wenn

$$\begin{aligned} \text{entweder } \operatorname{sgn} \xi_1 = \operatorname{sgn} \xi_2 = -\operatorname{sgn} \xi_3 \text{ oder } \operatorname{sgn} \eta_1 = \operatorname{sgn} \eta_2 = -\operatorname{sgn} \eta_3 \text{ und} \\ \text{,, } \operatorname{sgn} \xi_1 = -\operatorname{sgn} \xi_2 = \operatorname{sgn} \xi_3 \text{ ,, } \operatorname{sgn} \eta_1 = -\operatorname{sgn} \eta_2 = \operatorname{sgn} \eta_3 \text{ und} \\ \text{,, } \operatorname{sgn} \xi_1 = \operatorname{sgn} \xi_2 = \operatorname{sgn} \xi_3 \text{ ,, } -\operatorname{sgn} \eta_1 = \operatorname{sgn} \eta_2 = \operatorname{sgn} \eta_3 \text{ ist.} \end{aligned}$$

Da von den drei untereinander stehenden Bedingungen immer nur eine erfüllt sein kann, muß einer der drei Punkte in das Prisma (Q, S) fallen,

⁴⁾ Drei Punkte P, Q, R von solcher Art, daß OP, OQ, OR ein Fundamentalparallelepiped definieren, sollen ein Fundamentaltripel heißen.

kann also kein Gitterpunkt sein. Damit ist gezeigt, daß P, Q, S ein Fundamentaltripel bilden.

§ 5.

Wir entwickeln jetzt die analytische Gestalt des Algorithmus. Es sei P, Q, R ein Näherungstriple mit den Koordinaten ξ_i, η_i, ζ_i ($i = 1, 2, 3$), wobei $\zeta_i > 0$ sei. Es ist dann der im vorigen Paragraphen mit S bezeichnete Punkt zu bestimmen. Wir unterscheiden zwei Fälle.

I. Das Formenpaar $(\xi_1 x + \xi_2 y, \eta_1 x + \eta_2 y)$ ist reduziert (§ 2), $\xi_1 \eta_2 - \xi_2 \eta_1 \neq 0$. Da in der Ebene OPQ kein Gitterpunkt mit kleinerer Achsendistanz als Q liegt, weil sonst das genannte Formenpaar nicht reduziert wäre, muß S die Gestalt $aP + bQ + R$ haben und die Bedingung $\mu(S) < \mu(Q) = \text{Max}(|\xi_2|, |\eta_2|) = \mu_2$ erfüllen. Um a, b zu ermitteln, sei

$$(18) \quad \left. \begin{aligned} \xi_1 a + \xi_2 b + \xi_3 &= \xi_1 a' + \xi_2 b' = \varepsilon_1 \mu_2 \\ \eta_1 a + \eta_2 b + \eta_3 &= \eta_1 a' + \eta_2 b' = \varepsilon_2 \mu_2 \end{aligned} \right\} \quad |\varepsilon_1|, |\varepsilon_2| < 1.$$

Hieraus folgt

$$(19) \quad b - b' = \frac{\eta_1 \xi_3 - \eta_3 \xi_1}{\eta_2 \xi_1 - \eta_1 \xi_2}$$

Soll nun $\mu(S) < \mu_2$ sein, so ist dazu nach Satz 2 notwendig $|b'| < 2$. Setzt man daher ⁵⁾:

$$(20) \quad \frac{\eta_1 \xi_3 - \eta_3 \xi_1}{\eta_2 \xi_1 - \eta_1 \xi_2} = \left[\frac{\eta_1 \xi_3 - \eta_3 \xi_1}{\eta_2 \xi_1 - \eta_1 \xi_2} \right]_n + \delta = c + \delta,$$

so kann nach (19) b nur einen der Werte

$$(21) \quad \begin{matrix} \nearrow \\ \searrow \end{matrix} c, \quad c + 1, \quad c - 1, \quad c + 2 \operatorname{sgn} \delta$$

haben, die wir mit b_i ($i = 1, 2, 3, 4$) bezeichnen. Damit dann $\mu(S) < \mu(Q)$ sei, ist notwendig und hinreichend, daß ein zu einem b_i zugehöriges a_i den Intervallen

$$(22) \quad \left[\frac{\mu_2 - b_i \xi_2 - \xi_3}{\xi_1}, \frac{-\mu_2 - b_i \xi_2 - \xi_3}{\xi_1} \right] \quad \text{und} \quad \left[\frac{\mu_2 - b_i \eta_2 - \eta_3}{\eta_1}, \frac{-\mu_2 - b_i \eta_2 - \eta_3}{\eta_1} \right]$$

mit Ausschluß der Grenzen angehöre. Es sei nun:

$$(23) \quad \begin{cases} \alpha_i = \text{Min} \left[\frac{\mu_2}{|\xi_1|} - \frac{b_i \xi_2 + \xi_3}{\xi_1}, \frac{\mu_2}{|\eta_1|} - \frac{b_i \eta_2 + \eta_3}{\eta_1} \right] \\ \beta_i = \text{Max} \left[\frac{-\mu_2}{|\xi_1|} - \frac{b_i \xi_2 + \xi_3}{\xi_1}, \frac{-\mu_2}{|\eta_1|} - \frac{b_i \eta_2 + \eta_3}{\eta_1} \right]. \end{cases}$$

⁵⁾ $\{\alpha\}_n, \{\alpha\}_g, \{\alpha\}_k =$ nächste, nächstgrößere, nächstkleinere Zahl an α . Ist α ganz, sei $\{\alpha\}_g = \alpha + 1, \{\alpha\}_k = \alpha - 1$.

Ist dann $\alpha_i \leq \beta_i$, so existiert kein zu b_i gehöriger Wert a_i ; dasselbe gilt, wenn im Innern des Intervalls (β_i, α_i) keine ganze Zahl liegt. Enthält das Intervall dagegen eine solche, was aus $[\alpha_i]_k > \beta_i$ folgt, so setze man

$$(24) \quad a_i = [\beta_i]_g \text{ resp. } [\alpha_i]_k,$$

je nachdem $[\beta_i]_g \zeta_1 + b_i \zeta_2 + \zeta_3 > 0$ oder < 0 ist. Denn läßt man in $S' = \lambda P + b_i Q + R$ den Parameter λ das Intervall (β_i, α_i) durchlaufen, so geht dabei $h(S')$ nicht durch Null; denn anderenfalls würde im Intervall (β_i, α_i) eine ganze Zahl \bar{a} existieren, so daß der Gitterpunkt $\bar{S} = \bar{a}P + b_i Q + R$ die Bedingungen $h(\bar{S}) < h(P)$, $\mu(\bar{S}) < \mu(Q)$ erfüllte, was nicht möglich ist. Es wechselt daher $\zeta(S')$ im Intervall (β_i, α_i) nicht das Zeichen und daher ist für a_i eine der Zahlen zu nehmen, die möglichst nahe den Intervallgrenzen liegen, d. h. $[\beta_i]_g$ oder $[\alpha_i]_k$, und zwar diejenige, die den Gitterpunkt mit kleinerer Höhe ergibt. Ist $[\beta_i]_g \zeta_1 + b_i \zeta_2 + \zeta_3 > 0$, so ist offenbar $a_i = [\beta_i]_g$ zu setzen; ist der Ausdruck aber < 0 , so ist auch $[\alpha_i]_k \zeta_1 + b_i \zeta_2 + \zeta_3 < 0$ und $a_i = (\alpha_i)_k$ zu nehmen. Man erhält auf diese Weise höchstens vier und mindestens einen Punkt $S_i = (a_i, b_i, 1)$, unter denen der Punkt mit kleinster Höhe der gesuchte Punkt S ist. Das nächste Näherungstripel, mit dem der Algorithmus fortzusetzen ist, ist dann $P, \pm S, Q$ oder $\pm S, P, Q$, je nachdem $\mu(P) < 0$ oder $> \mu(S)$ ist.

II. Das Formenpaar $(\xi_1 x + \xi_2 y, \eta_1 x + \eta_2 y)$ ist nicht reduziert.

Es ist jetzt $\mu(P + Q) < \mu(Q)$ und der Punkt $P + Q$ ist der Punkt mit kleinster Höhe in der Ebene OPQ , der diese Bedingung erfüllt. Denn zunächst ist $\mu(P - Q) > \mu(Q)$, da die Höhe von $P - Q$ kleiner als die von P oder Q ist. Ist nun $\xi_1 \eta_2 - \xi_2 \eta_1 \neq 0$, so folgt $\mu(P + Q) < \mu(Q)$, weil wir anderenfalls Fall (I) hätten. Ist dann $\mu(Q) = (\eta_2)$, was man eventuell durch Vertauschung von ξ und η erreichen kann, so gilt $\text{sgn } \eta_1 \eta_2 = -1$, und soll daher $\mu(aP + bQ) < \mu(Q)$ sein, muß $\text{sgn } a b = 1$ sein, wenn $a b \neq 0$ ist, daher $h(aP + bQ) > h(P + Q)$ für $a b > 1$. Ist $\xi_1 \eta_2 - \xi_2 \eta_1 = 0$, so folgt aus $\mu(P - Q) > \mu(Q)$, daß entweder $\text{sgn } \eta_1 \eta_2$ oder $\text{sgn } \xi_1 \xi_2 = -1$ ist; es müssen daher beide Ausdrücke -1 sein und es folgt wieder $\mu(P + Q) < \mu(Q)$, und weiter ergeben sich die gleichen Schlüsse wie oben.

Wir untersuchen jetzt, ob es einen Punkt $S' = aP + bQ + R$ von solcher Art gibt, daß $\mu(S') < \mu(Q)$, $h(S') < h(P + Q)$ wird. Für einen solchen muß gelten

$$(25) \quad \begin{aligned} \xi_1 a + \xi_2 b + \xi_3 &= \varepsilon_1 \mu_2, \\ \eta_1 a + \eta_2 b + \eta_3 &= \varepsilon_2 \mu_2 & |\varepsilon_i| < 1, \\ \zeta_1 a + \zeta_2 b + \zeta_3 &= \varepsilon_3 (\zeta_1 + \zeta_2). \end{aligned}$$

Aus der zweiten und dritten Gleichung folgt

$$(26) \quad b = \frac{\zeta_1 \eta_3 - \zeta_3 \eta_1}{\zeta_2 \eta_1 - \zeta_1 \eta_2} + \varepsilon_3 \cdot \frac{\eta_1 (\zeta_1 + \zeta_2)}{\eta_1 \zeta_2 - \eta_2 \zeta_1} - \varepsilon_2 \cdot \frac{\zeta_1 \mu_3}{\eta_1 \zeta_3 - \eta_2 \zeta_1}.$$

Da die Summe der beiden letzten Glieder absolut den Betrag 2 nicht übersteigen kann, kann b , wenn man

$$(27) \quad \frac{\zeta_1 \eta_3 - \zeta_3 \eta_1}{\zeta_2 \eta_1 - \zeta_1 \eta_2} = c + \varepsilon = \left[\frac{\zeta_1 \eta_3 - \zeta_3 \eta_1}{\zeta_2 \eta_1 - \zeta_1 \eta_2} \right]_n + \varepsilon$$

setzt, nur einen der Werte $c, c + 1, c - 1, c + 2 \operatorname{sgn} \varepsilon$ haben, die wir mit b_i ($i = 1, 2, 3, 4$) bezeichnen. Wir bestimmen dann zugehörige Werte a_i nach den Formeln (23), (24) unter I. Von den höchstens fünf Punkten $P + Q$ und $S_i = a_i P + b_i Q + R$ wähle man den mit kleinster Höhe und nenne ihn S . Das nächste Näherungstriplet ist dann

$$\begin{aligned} P, \pm S, Q \quad \text{oder} \quad P, \pm S, R, & \quad \text{wenn} \quad \mu(S) > \mu(P), \\ \pm S, P, Q \quad \text{oder} \quad \pm S, P, R, & \quad \text{wenn} \quad \mu(S) < \mu(P). \end{aligned}$$

Das erste oder zweite Triplet ist in jeder Zeile zu nehmen, je nachdem $S \neq$ oder $= P + Q$ ist.

Ist $\mu_2 = \mu(Q) = |\xi_2|$, gelten die analogen Formeln mit vertauschten ξ und η .

§ 6.

Wir stellen in diesem Paragraphen zunächst die Formeln für den Algorithmus zusammen, bestimmen dann ein Ausgangstriplet und geben schließlich zwei Beispiele.

Es sei P, Q, R ein gegebenes Näherungstriplet mit den Koordinaten

$$\left. \begin{array}{l} P \\ Q \\ R \end{array} \right\} \begin{array}{l} \xi \quad \eta \quad \zeta \\ \xi_1 \quad \eta_1 \quad \zeta_1 \\ \xi_2 \quad \eta_2 \quad \zeta_2 \\ \xi_3 \quad \eta_3 \quad \zeta_3 \end{array} \quad \left. \begin{array}{l} \text{Hat Punkt } S \text{ die Koordinaten } \xi, \eta, \zeta, \text{ sei} \\ \mu(S) = \text{Max} (|\xi|, |\eta|) \text{ (Achsenabstand),} \\ h(S) = |\zeta| \text{ (Höhe).} \end{array} \right\}$$

$$(1) \quad \text{I. } \mu(P) \leq \mu(Q) \leq \mu(P \pm Q), \quad \xi_1 \eta_2 - \xi_2 \eta_1 \neq 0.$$

$$(2) \quad \left\{ \begin{array}{l} \frac{\xi_3 \eta_1 - \xi_1 \eta_3}{\xi_1 \eta_2 - \xi_2 \eta_1} = \left[\frac{\xi_3 \eta_1 - \xi_1 \eta_3}{\xi_1 \eta_2 - \xi_2 \eta_1} \right]_n + \delta = c + \delta, \\ b_i = c, c + 1, c - 1, c + 2 \operatorname{sgn} \delta \quad (i = 1, 2, 3, 4). \end{array} \right. \\ \mu_2 = \mu(Q).$$

$$(3) \quad \left\{ \begin{array}{l} a_i = \text{Min} \left[\frac{\mu_2}{|\xi_1|} - \frac{b_i \xi_2 + \xi_3}{\xi_1}, \frac{\mu_2}{|\eta_1|} - \frac{b_i \eta_2 + \eta_3}{\eta_1} \right] \\ \beta_i = \text{Max} \left[\frac{-\mu_2}{|\xi_1|} - \frac{b_i \xi_2 + \xi_3}{\xi_1}, \frac{-\mu_2}{|\eta_1|} - \frac{b_i \eta_2 + \eta_3}{\eta_1} \right] \end{array} \right. \quad (i = 1, 2, 3, 4).$$

$$(4) \quad \begin{cases} [\alpha_i]_k \leq \beta_i & \text{kein } a_i \\ [\alpha_i]_k > \beta_i \begin{cases} [\beta_i]_g + b_i \zeta_2 + \zeta_3 > 0, & a_i = [\beta_i]_g \\ [\beta_i]_g + b_i \zeta_2 + \zeta_3 < 0, & a_i = [\alpha_i]_k \end{cases} \end{cases} \quad (i = 1, 2, 3, 4).$$

Man wähle unter den Punkten

$$(5) \quad S_i = a_i P + b_i Q + R$$

denjenigen mit kleinster Höhe und nenne ihn S . Das nächste Näherungstriple ist dann $P, \pm S, Q$ oder $\pm S, P, Q$, je nachdem $\mu(S) >$ oder $< \mu(P)$ ist⁶⁾.

II. Die Bedingungen (1) sind nicht erfüllt.

A. $\mu(Q) = |\eta_2|.$

$$(6) \quad \begin{cases} \frac{\zeta_1 \eta_3 - \zeta_3 \eta_1}{\zeta_2 \eta_1 - \zeta_1 \eta_2} = \left[\frac{\zeta_1 \eta_3 - \zeta_3 \eta_1}{\zeta_2 \eta_1 - \zeta_1 \eta_2} \right]_n + \varepsilon = d + \varepsilon, \\ b_i = d, d + 1, d - 1, d + 2 \operatorname{sgn} \varepsilon \end{cases} \quad (i = 1, 2, 3, 4).$$

Man bestimme die Punkte S_i , soweit sie existieren, nach den Formeln (3), (4), (5) unter I. und wähle dann unter den Punkten $P + Q, S_i$ den Punkt mit kleinster Höhe und nenne ihn S . Das nächste Näherungstriple ist dann:

$$\begin{array}{ll} P, \pm S, Q \quad \text{oder} \quad P, \pm S, R, & \text{wenn } \mu(P) < \mu(S), \\ \pm S, P, Q \quad \text{oder} \quad \pm S, P, R, & \text{wenn } \mu(P) > \mu(S). \end{array}$$

Das erste oder zweite Triple in jeder Zeile ist zu nehmen, je nachdem $S \neq$ oder $= P + Q$ ist.

B. $\mu(Q) = |\xi_2|.$

Es gelten die Formeln unter A mit vertauschten ξ und η .

Der auf P folgende Näherungspunkt ist der erste bei dem Algorithmus auftretende Punkt, dessen Achsendistanz kleiner als die von P ist.

Wir brauchen noch eine Ausgangstriple, mit dem wir den Algorithmus beginnen können. Es sei jetzt speziell

$$\xi = x - \alpha z, \quad \eta = y - \beta z, \quad \zeta = z,$$

wo α, β zwei reelle Zahlen bedeuten, zwischen denen keine Relation $g_1 \alpha + g_2 \beta + g_3 = 0$ mit ganzzahligen g_i besteht. Wir setzen dann:

$$\alpha = [\alpha]_n + \delta = a + \delta, \quad \beta = [\beta]_n + \varepsilon = b + \varepsilon$$

und wählen als Ausgangstriple:

$$\begin{array}{l} P = (a, b, 1), \quad Q = (a + \operatorname{sgn} \delta, b, 1), \quad R = (a, b + \operatorname{sgn} \varepsilon, 1), \quad \text{wenn } |\delta| > |\varepsilon|, \\ P = (a, b, 1), \quad Q = (a, b + \operatorname{sgn} \varepsilon, 1), \quad R = (a + \operatorname{sgn} \delta, b, 1), \quad \text{wenn } |\delta| < |\varepsilon|. \end{array}$$

⁶⁾ Das Vorzeichen bei $\pm S$ ist so zu bestimmen, daß $\zeta(\pm S) > 0$ wird.

Wir führen noch zwei Beispiele an, in denen die Näherungspunkte und Hilfspunkte angegeben sind. Die ersten sind dadurch kenntlich, daß bei ihnen der Näherungskoeffizient k hinzugefügt ist.

Erstes Beispiel. $\alpha = \sqrt[3]{2} = 1,2599$; $\beta = \sqrt[3]{4} = 1,5874$.

x	1	1	2	3	4	5	9	15	24	34	49	58	73	165	223
y	2	1	2	3	5	6	11	19	30	43	62	73	92	208	281
z	1	1	1	2	3	4	7	12	19	27	39	46	58	131	177
k	0,41				0,41		0,48	0,41				0,32			0,40

Zweites Beispiel. $\alpha = 1,3248$ (Wurzel von $x^3 - x - 1 = 0$),
 $\beta = \alpha^2 = 1,7551$.

x	1	2	1	3	4	5	9	12	16
y	2	2	1	4	5	7	12	16	21
z	1	1	1	2	3	4	7	9	12
k	0,33				0,46			0,61	0,36

(Eingegangen am 6. 2. 1927.)

Newton'sche Polygone in der Theorie der algebraischen Körper.

Von

Öystein Ore in Oslo.

Einleitung.

Bei den arithmetischen Untersuchungen über algebraische Zahlen und algebraische Körper bildet die allgemeine Dedekindsche Theorie der Ideale das Fundament, worauf sich alle weiteren Resultate aufbauen. Der vorgelegte algebraische Körper ist aber allgemein durch eine bestimmte irreduzible Gleichung definiert und es entsteht daher natürlich das folgende Problem, das besonders für die Anwendung der Idealtheorie wichtig ist: Welche Verbindung besteht zwischen der allgemeinen Dedekindschen Idealtheorie im Körper und den Eigenschaften einer beliebigen Gleichung, die den entsprechenden Körper festlegt?

Diese Frage habe ich in zwei Arbeiten: „Über den Zusammenhang zwischen den definierenden Gleichungen und der Idealtheorie in algebraischen Körpern.“ Erste und zweite Mitteilung¹⁾ behandelt. Man erhielt daraus bei einer beliebigen irreduziblen Gleichung sofort die vollständige Form der Primidealzerlegung einer beliebigen Primzahl im entsprechenden Körper, weiter eine vollständige und einfache Verzweigungstheorie sowie eine Reihe von anderen Resultaten.

Die vorliegende Arbeit bildet eine direkte Fortsetzung der beiden erwähnten Arbeiten, indem sie auch die Verbindung zwischen der Gleichung und der Idealtheorie behandelt. Bekanntlich spielen in der Theorie der algebraischen Funktionen die Newton(-Puiseux)schen Polygone eine fundamentale Rolle, und es handelt sich hier um die Übertragung dieser Methode auf die Theorie der algebraischen Körper. Dabei muß aber der Methode eine neue und allgemeinere Form gegeben werden; die Notwen-

¹⁾ *Math. Annalen* 96 (1926), S. 313–352 und 97 (1927), S. 569–598. Die beiden Arbeiten werden im folgenden kurz bzw. mit I und II bezeichnet.

digkeit dieser Umformung folgt u. a. daraus, daß bei den algebraischen Zahlen Primideale von allen Graden auftreten, während bei den algebraischen Funktionen alle Primideale vom ersten Grade sind. Von dieser allgemeineren Theorie kann man auch umgekehrt interessante Anwendungen auf algebraische Funktionen machen, worauf ich aber hier nicht eingehe.

Durch Anwendung der Polygone kann man, analog wie man bei den algebraischen Funktionen die Reihenentwicklungen bestimmt, die vollständige Primidealzerlegung einer beliebigen Primzahl bestimmen; diese Aufgabe ist schon allgemein durch die beiden Hauptsätze (I, Kap. 2, Satz 3 und II, Kap. 1, Satz 9) gelöst, aber die Methode der Polygone wird bei einer vorgelegten Gleichung mit numerischen Koeffizienten viel rascher zum Ziele führen. Weiter erhält man durch die Polygone auch eine Darstellung in der Normalform $p = (p, \varphi(\theta))$ für jedes Primideal, woraus man sofort ein Fundamentalsystem für \mathfrak{p} und p aufstellen kann, und weiter auch numerisch alle Größen der Verzweigungstheorie berechnen kann. Diese Probleme habe ich schon in einigen früheren Arbeiten³⁾ behandelt, es war mir aber nicht gelungen, eine einfache Form der Theorie zu erhalten und auch nicht für *alle* Fälle die Untersuchung durchzuführen.

In Kap. 1 wird die allgemeine Theorie der Polygone behandelt und zunächst ist in § 2 der Multiplikationssatz für Polygone bewiesen. Wie ich schon früher bewiesen habe, enthält dieser Satz einen allgemeinen Irreduzibilitätssatz, worin z. B. die Sätze von Schönemann, Eisenstein, Königsberger, Perron, Bauer und Dumas enthalten sind. Weiter kann man daraus unter Anwendung des ersten Hauptsatzes allgemeine Primitivitätskriterien für Gleichungen aufstellen, wovon Sätze von Schur, Bauer und Furtwängler Spezialfälle sind. Weiter werden verschiedene andere Eigenschaften der Polygone behandelt, um den allgemeineren Produktsatz 7 zu beweisen.

In Kap. 2 wird der Zusammenhang zwischen Ideal- und Polygoneigenschaften studiert, und unter Anwendung eines wichtigen Hilfssatzes über das Polygon eines $(\text{mod } p^a)$ irreduziblen Polynoms folgt leicht der Beweis des allgemeinen Satzes 5 über Polygone. Für die sogenannten *regulären* Gleichungen liefert dieser Satz die vollständige Primidealzerlegung einer Primzahl; durch den Existenzbeweis für reguläre Gleichungen erhält man eine neue Normalform der Gleichungen des Körpers, die eine Reihe von interessanten Eigenschaften besitzt. Zuletzt wird im allgemeinen Falle die Primidealzerlegung bestimmt.

Es soll nur noch beiläufig bemerkt werden, worauf ich hier aber

³⁾ Ö. Ore, „Zur Theorie der algebraischen Körper“, Acta math. 44 (1923), S. 219 bis 314. „Weitere Untersuchungen zur Theorie der algebraischen Körper“, Acta math. 45 (1924), S. 145–160.

nicht eingehe, daß man unter Anwendung der Newtonschen Polygone die noch fehlende allgemeine Theorie der höheren Kongruenzen für Primzahlpotenzmoduln $(\text{mod } p^\alpha)$ schaffen kann.

Kapitel 1.

Allgemeine Eigenschaften der Polynome.

§ 1.

Einführung der Newtonschen Polygone.

Zuerst soll hier die Zerlegung der Polynome in Faktoren $(\text{mod } p^\alpha)$ studiert werden. Es sei

$$f(x) = x^n + a_1 x^{n-1} + \dots + a_n$$

ein beliebiges Polynom, wobei die Diskriminante genau durch p^δ teilbar ist. Für die Zerlegung von $f(x)$ in irreduzible Faktoren $(\text{mod } p^\alpha)$ hat man dann nach I, Kap. 1, Satz 4 und Kap. 2, Satz 5 den fundamentalen Satz:

Satz 1. *Besteht für $f(x)$ die Zerlegung in irreduzible Faktoren*

$$(1) \quad f(x) \equiv f_1(x) \dots f_r(x) \pmod{p^{\delta+1}},$$

so besteht auch für alle $\alpha \geq \delta + 1$ eine entsprechende Zerlegung in irreduzible Faktoren

$$(2) \quad f(x) \equiv f_1^{(\alpha)}(x) \dots f_r^{(\alpha)}(x) \pmod{p^\alpha},$$

und diese Zerlegung ist $(\text{mod } p^{\alpha-\delta})$ eindeutig.

Dieser Satz ist zunächst in I nur unter der Voraussetzung bewiesen, daß $f(x)$ rational irreduzibel ist; wie man aber zeigen kann, bleibt der Satz auch dann richtig, wenn $f(x)$ reduzibel ist, wenn nur vorausgesetzt wird, daß die Diskriminante von Null verschieden ist. Im folgenden wird der Satz aber nur in dem Falle benutzt, wo $f(x)$ irreduzibel ist.

Aus der Eindeutigkeit der Zerlegung folgt, daß wenn $g(x)$ und $h(x)$ zwei unzerlegbare Teiler von $f(x) \pmod{p^\alpha}$ sind, so ist entweder

$$g(x) \equiv h(x) \pmod{p^{\alpha-\delta}} \quad \text{oder} \quad g(x) \not\equiv h(x) \pmod{p^{\delta+1}}.$$

Wenn im folgenden von *verschiedenen* Teilern von $f(x) \pmod{p^\alpha}$ die Rede ist, so wird dabei gemeint, daß der letzte Fall vorliegt.

Es sei nun vorläufig $f(x)$ ein beliebiges Polynom und

$$(3) \quad f(x) \equiv \varphi_1(x)^{a_1} \dots \varphi_s(x)^{a_s} \pmod{p}$$

die Primfunktionzerlegung von $f(x) \pmod{p}$, wobei allgemein m_i der Grad einer Primfunktion $\varphi_i(x) \pmod{p}$ ist. Nach dem Schönemannschen Satz (I, Kap. 1, § 3) folgt aus (3), daß für alle α eine Zerlegung

$$(4) \quad f(x) \equiv \Phi_1(x) \dots \Phi_s(x) \pmod{p^\alpha}$$

besteht, wobei

$$\Phi_i(x) \equiv \varphi_i(x)^{a_i} \pmod{p} \quad (i = 1, 2, \dots, s)$$

und daher der Grad eines Faktors $\Phi_i(x)$ gleich $a_i m_i$ ist. Diese Faktoren sind natürlich im allgemeinen nicht $\pmod{p^\alpha}$ unzerlegbar und es wird im folgenden eine wichtige Aufgabe, die weitere Zerlegung der Faktoren $\Phi_i(x)$ zu bestimmen. Eben für diese Aufgabe sind die Newtonschen Polygone ein vorzügliches Hilfsmittel.

Es sei $\varphi(x)$ eine beliebige unter den Primfunktionen $\varphi_i(x)$ in (3), welche $f(x) \pmod{p}$ teilt; der Grad von $\varphi(x)$ soll gleich m sein. Man kann dann das gegebene Polynom $f(x)$ in der Form

$$(5) \quad f(x) = \sum_{i=0}^t Q_i(x) p^{\alpha_i} \varphi(x)^{t-i}$$

schreiben, wobei die Polynome $Q_i(x)$ vom Grade kleiner als m sind, und wo weiter der Exponent α_i so gewählt wird, daß $Q_i(x)$ nicht durch p teilbar ist, wenn $Q_i(x) \neq 0$. Die Zahl t ist durch

$$t = \left[\frac{n}{m} \right]$$

bestimmt.

Eine Darstellung von $f(x)$ in der Form (5) soll eine *Entwicklung* $(p, \varphi(x))$ von $f(x)$ heißen. Wenn $\varphi(x)$ gegeben ist, so ist die Entwicklung $(p, \varphi(x))$ von $f(x)$ eindeutig bestimmt und kann durch sukzessive Divisionen mit Potenzen von $\varphi(x)$ ermittelt werden. Da in $f(x)$ der höchste Koeffizient gleich eins ist, so wird in (5) immer $\alpha_0 = 0$. Bei der Definition der Entwicklung $(p, \varphi(x))$ war vorausgesetzt, daß $\varphi(x) \pmod{p}$ ein Teiler von $f(x)$ sei. Wie man aber sieht, kann man auch die Entwicklung für eine beliebige Primfunktion $\varphi(x) \pmod{p}$ bilden, doch wird im folgenden nur der Fall von Interesse, daß $\varphi(x)$ das Polynom $f(x) \pmod{p}$ wirklich teilt. Der Begriff der Entwicklung $(p, \varphi(x))$ hat sich schon in I, Kap. 3, § 3 bei der Bestimmung der Körperdifferente als nützlich erwiesen.

Aus der Entwicklung (5) von $f(x)$ leitet man die Zahlpaare

$$(6) \quad (0, 0), (1, \alpha_1), \dots, (t, \alpha_t)$$

ab, und diese können als Gitterpunkte in einem rechtwinkligen Koordinatensystem abgebildet werden. Zu den Punkten (6) konstruiert man in bekannter Weise das entsprechende Newtonsche Polygon, das gegen die X -Achse konvex wird; das so entstandene Polygon heißt *das Polygon* $(p, \varphi(x))$ von $f(x)$.

Das Polygon $(p, \varphi(x))$ wird im Punkte $(0, 0)$ seinen Anfangspunkt und im Punkte (t, α_t) seinen Endpunkt haben; alle Punkte (6), welche nicht auf dem Polygone liegen, müssen oberhalb des Polygons fallen. Man kann im folgenden immer voraussetzen, daß das Polygon im Endlichen liegt; man könnte sich nämlich auch den Fall denken, daß die letzte Seite des Polygons ins Unendliche gehe, indem in (6) $\alpha_t = \infty$ wäre. Dieser Fall tritt dann ein, wenn $\varphi(x)$ nicht nur $(\text{mod } p)$, sondern auch absolut ein Teiler von $f(x)$ wäre, indem dann in (5) das letzte Glied verschwinden würde und man daher $\alpha_t = \infty$ setzen müßte. Im folgenden ist dieser Fall ohne Bedeutung und soll ausgeschlossen werden, indem es sich um die Zerlegung von $f(x)$ in irreduzible Faktoren $(\text{mod } p^a)$ handelt, und $\varphi(x)$ kann daher als solcher ausgeschieden werden. Setzt man, wie später in Kap. 2, $f(x)$ irreduzibel voraus, tritt dieser Fall nie ein.

Da $\varphi(x) \pmod{p}$ ein Teiler von $f(x)$ ist, wird sicher $\alpha_t \geq 1$ und das Polygon $(p, \varphi(x))$ von $f(x)$ kann daher nicht vollständig mit der X -Achse zusammenfallen. Derjenige Teil des Polygons, der nicht mit der X -Achse zusammenfällt, soll das *Hauptpolygon* $(p, \varphi(x))$ von $f(x)$ heißen. Wenn $\varphi(x) \pmod{p}$ kein Teiler von $f(x)$ wäre, so würde $\alpha_t = 0$ sein, und das Polygon $(p, \varphi(x))$ fiel dann mit der X -Achse zusammen. Für eine gegebene Primzahl p existieren also genau so viele Hauptpolygone $(p, \varphi(x))$ wie es verschiedene Primfunktionenteiler $\varphi(x)$ von $f(x) \pmod{p}$ gibt, also mit der Bezeichnung in (3) genau s .

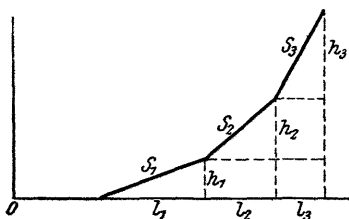
Für das Hauptpolygon $(p, \varphi(x))$ von $f(x)$ sollen immer die folgenden Bezeichnungen angewandt werden: Die Seiten des Polygons sind

$$S_1, S_2, \dots, S_k,$$

und diese Seiten haben auf die X -Achse bzw. Y -Achse die Projektionen

$$l_1, l_2, \dots, l_k$$

$$h_1, h_2, \dots, h_k.$$



Die Zahlen l_i und h_i sind ganz rational, und wenn ε_i der größte gemeinsame Faktor von l_i und h_i ist, kann man also

$$(7) \quad l_i = \varepsilon_i \lambda_i, \quad h_i = \varepsilon_i \kappa_i \quad (i = 1, 2, \dots, k)$$

setzen, wobei λ_i zu κ_i relativ prim ist. Die Zahlen λ_i und κ_i haben auch, wie man später bemerkt, eine einfache geometrische Bedeutung. Bildet die Seite S_i mit der X -Achse den Winkel φ_i , so ist

$$\text{tg } \varphi_i = \frac{h_i}{l_i} = \frac{\kappa_i}{\lambda_i};$$

die Zahl $\frac{\alpha_i}{\lambda_i}$ soll die *Neigung* der Seite heißen. Offenbar gelten dann die Ungleichungen

$$(8) \quad \frac{\alpha_1}{\lambda_1} < \frac{\alpha_2}{\lambda_2} < \dots < \frac{\alpha_k}{\lambda_k}.$$

Wenn $f(x)$ genau durch $\varphi(x)^a \pmod{p}$ teilbar ist, so müssen in (5) die a letzten Exponenten $\alpha_i > 0$ sein, während $\alpha_{i-a} = 0$ ist. Das Hauptpolygon von $f(x) \pmod{p, \varphi(x)}$ hat daher eine Projektion auf die X -Achse mit der Länge a , folglich ist

$$(9) \quad l_1 + l_2 + \dots + l_k = a.$$

Es sei nun $\Phi(x)$ derjenige Teiler von $f(x) \pmod{p^a}$ in (4), wofür $\Phi(x) \equiv \varphi(x)^a \pmod{p}$ ist; dann wird zunächst das Polygon von $\Phi(x) \pmod{p, \varphi(x)}$ ein Hauptpolygon, indem in der Entwicklung $(p, \varphi(x))$ alle Glieder außer dem ersten $\varphi(x)^a$ durch p teilbar sind. Man sieht nun allgemein ein, daß das Hauptpolygon von einem Polynome $g(x) \pmod{p, \varphi(x)}$ ungeändert wird, wenn man $g(x)$ mit einem Polynome $k(x)$ multipliziert, das \pmod{p} nicht durch $\varphi(x)$ teilbar ist. In (4) sind aber die übrigen Faktoren $\Phi_i(x) \pmod{p}$ nicht durch $\varphi(x)$ teilbar, und daher folgt leicht:

Wenn man in (4) $\alpha > \alpha_i$ wählt, so ist das Polygon $(p, \varphi(x))$ eines Faktors $\Phi(x)$ gleich dem Hauptpolygone von $f(x) \pmod{p, \varphi(x)}$.

Man zeigt auch leicht, daß das Hauptpolygon von $f(x) \pmod{p, \varphi(x)}$ ungeändert bleibt, wenn man zur Bestimmung des Polygons nicht die reduzierte Darstellung (5) anwendet, sondern annimmt, daß bei den Polynomen $Q_i(x)$ auch Gradzahlen größer als $m - 1$ gestattet sind, wenn nur immer $Q_i(x) \equiv 0 \pmod{p, \varphi(x)}$ ist.

§ 2.

Der Multiplikationssatz für Polygone.

Es soll jetzt der folgende wichtige Satz über die Eigenschaften der Polygone $(p, \varphi(x))$ bewiesen werden:

Satz 2. *Sind*

$$(10) \quad \begin{cases} g(x) = \sum_{i=0}^{i'} Q_i'(x) p^{\alpha_i'} \varphi(x)^{i'-i}, \\ h(x) = \sum_{j=0}^{j''} Q_j''(x) p^{\alpha_j''} \varphi(x)^{j''-j} \end{cases}$$

zwei Polynome mit den Hauptpolygonen S' und S'' , so hat das Produkt

$$(11) \quad f(x) = g(x)h(x) = \sum_{i=1}^i Q_i(x) p^{\alpha_i} \varphi(x)^{i-i}$$

ein Hauptpolygon S , das aus den Seiten der Hauptpolygone S' und S'' gebildet ist, und zwar so, daß man die Seiten nach steigender Neigung ordnet.

Für den Beweis dieses Satzes kann man annehmen, daß die Polygone von $g(x)$ und $h(x)$ beide Hauptpolygone sind. Denn wenn dies nicht der Fall ist, kann man nach § 1

$$g(x) \equiv \Pi_1(x) \Phi_1(x) \pmod{p^\alpha},$$

$$h(x) \equiv \Pi_2(x) \Phi_2(x) \pmod{p^\alpha}$$

schreiben, wobei $\Pi_1(x)$ und $\Pi_2(x) \pmod{p}$ nicht durch $\varphi(x)$ teilbar sind, während die Polygone von $\Phi_1(x)$ und $\Phi_2(x)$ Hauptpolygone sind und beziehungsweise gleich S' und S'' , wenn nur $\alpha > \alpha'_i + \alpha''_i$ gewählt wird. Daraus folgt aber

$$f(x) = g(x)h(x) \equiv [\Pi_1(x) \Pi_2(x)] [\Phi_1(x) \Phi_2(x)] \pmod{p^\alpha},$$

woraus man sieht, daß das Hauptpolygon von $f(x)$ gleich dem Polygone vom Produkte $\Phi_1(x) \Phi_2(x)$ ist.

Man kann also auch speziell voraussetzen, daß die Grade von $g(x)$ und $h(x)$ Multipla von m sind, also $Q'_0(x) = Q''_0(x) = 1$, und daher in (11) $Q_0(x) = 1$ und $t = t' + t''$. Multipliziert man nun die Polynome (10) miteinander, so kommt

$$f(x) = g(x)h(x) = \sum_{i,j=0}^{t',t''} Q'_i(x) Q''_j(x) p^{\alpha'_i + \alpha''_j} \varphi(x)^{t-i-j}.$$

Die Glieder

$$G = Q'_i(x) p^{\alpha'_i} \varphi(x)^{t'-i}, \quad H = Q''_j(x) p^{\alpha''_j} \varphi(x)^{t''-j}$$

werden durch die beiden Punkte

$$G = (i, \alpha'_i), \quad H = (j, \alpha''_j)$$

abgebildet, während das Produktglied

$$GH = Q'_i(x) Q''_j(x) p^{\alpha'_i + \alpha''_j} \varphi(x)^{t-i-j}$$

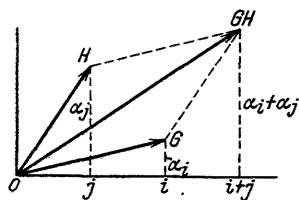
durch den Punkt

$$GH = (i + j, \alpha'_i + \alpha''_j)$$

abgebildet wird. Diesen Punkt GH kann man sich so entstanden denken, daß man die beiden Vektoren \overline{OG} und \overline{OH} bildet, der Vektor $\overline{OG} + \overline{OH}$ wird dann in GH seinen Endpunkt haben, also ist

$$\overline{OG} + \overline{OH} = \overline{O(GH)}.$$

Sowohl für das Polygon S' als das Polygon S'' liegen die Abbildungspunkte G und H für alle Glieder in (10) auf oder oberhalb des entsprechenden Newtonschen Polygons, nie unterhalb des Polygons. Man konstruiert nun aus den Seiten von S' und S'' ein neues Polygon S_1 , indem



man diese Seiten nach steigender Neigung ordnet. Aus der erwähnten Vektorendarstellung folgt dann sofort, daß ein Punkt GH immer oberhalb S_1 liegen muß, sobald nur entweder G oder H oberhalb des entsprechenden Polygons S' oder S'' liegen; weiter ist einleuchtend, daß ein Punkt GH nie unterhalb S_1 liegen kann.

Um daher zu beweisen, daß S_1 das Polygon von $f(x)(p, \varphi(x))$ ist, braucht man nur zu zeigen, daß es für jeden Eckpunkt genau ein entsprechendes Produktglied GH gibt, das durch diesen Punkt abgebildet wird.

Dies ist aber leicht nachweisbar. Denn wenn P ein Eckpunkt von S_1 ist, so besteht das Polygon S_1 von O bis zu P von gewissen Seiten

$$S'_1, S'_2, \dots, S'_r,$$

$$S''_1, S''_2, \dots, S''_{r'},$$

welche zu den Polygonen S' und S'' gehören, und nach dem Bildungsgesetz von S_1 sind diese Seiten die r' bzw. r'' ersten Seiten von S' und S'' . Bezeichnet man nun auf dem Polygone S' den Endpunkt der Seite S'_r mit G und ebenso auf S'' den Endpunkt von $S''_{r'}$ mit H , so wird der Punkt GH in P fallen. Denn die beiden Vektoren \overline{OG} und \overline{OH} können durch die Vektorpolygone

$$\overline{S}'_1 + \overline{S}'_2 + \dots + \overline{S}'_r,$$

$$\overline{S}''_1 + \overline{S}''_2 + \dots + \overline{S}''_{r'}$$

ersetzt werden und die Summe der beiden wird natürlich in P ihren Endpunkt haben. Sogleich bemerkt man auch ohne Schwierigkeit, daß dies die einzige Weise ist, wodurch ein Produkt GH durch den Punkt P abgebildet werden kann. Der Satz 2 ist daher bewiesen.

Vom Satz 2 wird sehr oft ein Spezialfall angewandt: Wenn nämlich ein Polynom $f(x)$ ein geradliniges Polygon besitzt, so folgt, daß die Faktoren von $f(x)$ auch alle geradlinige Polygone mit derselben Neigung haben müssen.

Aus dem Satze 1 habe ich in einer früheren Arbeit³⁾ einen allgemeinen *Irreduzibilitätssatz* abgeleitet, worin die Sätze von Schönemann, Eisenstein, Königsberger, Perron und Bauer, sowie die Untersuchungen von Dumas als Spezialfälle enthalten sind.

Weiter kann man aus diesem Satze unter Anwendung des ersten Hauptsatzes (I, Kap. 2, Satz 3) eine Reihe von Primitivitätskriterien für algebraische Gleichungen ableiten⁴⁾; Sätze von Schur, Bauer und Furtwängler sind darin enthalten.

³⁾ Ö. Ore, „Zur Theorie der Irreduzibilitätskriterien“, Math. Zeitschr. 18 (1923), S. 278—288. In dieser Arbeit findet man auch die zugehörigen Literaturangaben.

⁴⁾ Ö. Ore, „Kriterien für Gleichungen mit primitiven Gruppen.“ Akademie der Wissenschaften, Oslo, Mat.-nat. Klasse, Nr. 18 (1924). Hier findet man auch die Literaturangaben.

§ 3.

Die Faktoren der Seiten.

Es sei wieder angenommen, daß das Polygon $(p, \varphi(x))$ von dem Polynome $f(x)$ ein Hauptpolygon ist, und die Entwicklung $(p, \varphi(x))$ von $f(x)$ sei durch (5) gegeben. Man soll nun speziell die Eigenschaften derjenigen Glieder in dieser Entwicklung untersuchen, welche durch Punkte auf dem Polygone S von $f(x)$ abgebildet werden. Im folgenden wird der Kürze wegen oft gesagt, ein Glied in der Entwicklung (5) liegt auf S , dabei wird gemeint, daß der entsprechende abbildende Punkt in (6) auf S liegt.

Wie man leicht sieht, hat die Summe der Glieder in (5), welche auf der ersten Seite S_1 des Polygons liegen, die Form

$$\varphi(x)^t + Q_{\lambda_1}(x) p^{\lambda_1} \varphi(x)^{t-\lambda_1} + Q_{2\lambda_1}(x) p^{2\lambda_1} \varphi(x)^{t-2\lambda_1} \\ + \dots + Q_{h_1}(x) p^{h_1} \varphi(x)^{t-h_1},$$

und diese Summe kann man in der etwas kürzeren Form

$$\varphi(x)^{t-h_1} [\varphi(x)^{h_1} + R_{1,1}(x) p^{\lambda_1} \varphi(x)^{h_1-\lambda_1} + R_{1,2}(x) p^{2\lambda_1} \varphi(x)^{h_1-2\lambda_1} \\ + \dots + R_{1,\varepsilon_1}(x) p^{h_1}]$$

schreiben; man beachte dabei die Bezeichnungen (7). Ebenso wird die Summe der Glieder auf der zweiten Seite

$$\varphi(x)^{t-h_1-h_2} p^{h_1} [R_{2,0}(x) \varphi(x)^{h_2} + R_{2,1}(x) p^{\lambda_2} \varphi(x)^{h_2-\lambda_2} \\ + R_{2,2}(x) p^{2\lambda_2} \varphi(x)^{h_2-2\lambda_2} + \dots + R_{2,\varepsilon_2}(x) p^{h_2}]$$

und im allgemeinen für die i -te Seite

$$(12) \left\{ \varphi(x)^{t-h_1-\dots-h_i} p^{h_1+\dots+h_{i-1}} [R_{i,0}(x) \varphi(x)^{h_i} + R_{i,1}(x) p^{\lambda_i} \varphi(x)^{h_i-\lambda_i} \right. \\ \left. + R_{i,2}(x) p^{2\lambda_i} \varphi(x)^{h_i-2\lambda_i} + \dots + R_{i,\varepsilon_i}(x) p^{h_i}] \right\}.$$

Zwei aufeinanderfolgende Seiten werden immer ein gemeinsames Glied enthalten, nämlich dasjenige, das durch den gemeinsamen Eckpunkt abgebildet wird. Daraus folgt sofort, daß immer

$$R_{i,0}(x) = R_{i-1,\varepsilon_{i-1}}(x) \equiv 0 \pmod{p, \varphi(x)}$$

ist.

Man kann also auch für jede Seite ein solches Polynom $A_i(x)$ bestimmen, daß

$$(13) \quad R_{i,0}(x) A_i(x) \equiv 1 \pmod{p, \varphi(x)},$$

und für jedes Polynom $R_{i,j}(x)$ in (12) wird

$$(14) \quad S_{i,j}(x) = A_i(x) R_{i,j}(x) \quad (j = 1, 2, \dots, \varepsilon_i)$$

gesetzt. Das Polynom

$$(15) \quad f_i(x) = \varphi(x)^{i_i} + S_{i,1}(x)p^{s_i}\varphi(x)^{i_i-2i_i} + S_{i,2}(x)p^{2s_i}\varphi(x)^{i_i-2\cdot 2i_i} \\ + \dots + S_{i,e_i}(x)p^{h_i},$$

das man so abgeleitet hat, spielt im folgenden eine wichtige Rolle und heißt *der Faktor der i-ten Seite*.

Wie man sieht, hat $f_i(x)$ ein geradliniges Polygon S_i , welches dieselbe Länge und Neigung wie die i -te Seite des Polygons S hat.

Aus $f_i(x)$ leitet man weiter eine wichtige Hilfsgröße ab, nämlich das Polynom

$$(16) \quad F_i(x, y) = y^{e_i} + S_{i,1}(x)y^{e_i-1} + \dots + S_{i,e_i}(x),$$

das *das zugeordnete Polynom der i-ten Seite* heißt.

Zwischen dem zugeordneten Polynome $F_i(x, y)$ und dem Faktor $f_i(x)$ der Seite besteht die Relation

$$(17) \quad p^{h_i} F_i(x, \Theta_i(x)) = f_i(x),$$

wobei

$$(18) \quad \Theta_i(x) = \frac{\varphi(x)^{i_i}}{p^{s_i}}$$

gesetzt worden ist; die Richtigkeit von (17) folgt sofort aus (15) und (16), indem man die Bezeichnungen (7) beachtet.

Die Wichtigkeit dieser Hilfsbezeichnungen folgt leicht aus den weiteren Untersuchungen. Wenn zwei Polynome $f(x)$ und $\psi(x)$ dasselbe Polygon $S(p, \varphi(x))$ haben, so kann es vorkommen, daß alle Glieder in $f(x)$ und $\psi(x)$, welche auf S liegen, gleich sind, d. h. wenn man die Differenz $f(x) - \psi(x)$ bildet, so gibt es darin nur Glieder, welche oberhalb S liegen. Im folgenden wird dies kurz mit

$$(19) \quad f(x) \equiv \psi(x) \pmod{S}$$

bezeichnet.

Man kann nun genau die Bedingung dafür angeben, daß eine Kongruenz (19) besteht. Wenn nämlich

$$\psi(x) = \sum_{i=0}^t Q_i^{(1)}(x)p^{s_i'}\varphi(x)^{t-i}$$

die Entwicklung $(p, \varphi(x))$ ist, so besteht die Kongruenz (19) dann und nur dann, wenn

$$Q_i(x) \equiv Q_i^{(1)}(x) \pmod{p}$$

für alle i ist, wofür der entsprechende Punkt auf dem Polygone S liegt. Wenn man für die Entwicklung $(p, \varphi(x))$ von $f(x)$ und $\psi(x)$ die etwas

allgemeinere Form annimmt, wo die Grade der Koeffizientenpolynome $Q_i(x)$ auch größer als $m - 1$ sein dürfen, so lautet die Bedingung

$$Q_i(x) \equiv Q_i^{(1)}(x) \pmod{p, \varphi(x)}$$

für alle i , wofür der entsprechende Punkt auf S liegt. Aus der Definition der zugeordneten Polynome folgt dann sofort:

Satz 3. *Es seien $f(x)$ und $\psi(x)$ zwei Polynome mit demselben Polygone $S(p, \varphi(x))$ und weiter seien*

$$F_i(x, y), \quad F_i^{(1)}(x, y) \quad (i = 1, 2, \dots, k)$$

die entsprechenden zugeordneten Polynome der Seiten. Die notwendige und hinreichende Bedingung dafür, daß die Kongruenz

$$f(x) \equiv \psi(x) \pmod{S}$$

besteht, ist

$$F_i(x, y) \equiv F_i^{(1)}(x, y) \pmod{p, \varphi(x)} \quad (i = 1, 2, \dots, k).$$

Für den Spezialfall, daß das Polygon S eine Gerade L ist, existiert nur ein einziger zugeordneter Faktor $f_1(x)$, wofür also

$$(20) \quad f(x) \equiv f_1(x) \pmod{L}$$

ist; weiter gibt es nur ein einziges zugeordnetes Polynom $F(x, y)$.

Wenn ein solches $f(x)$ das Produkt von zwei Faktoren $g(x)$ und $h(x)$ ist, so müssen diese auch geradlinige Polygone mit derselben Neigung wie L haben. Da weiter in $f(x)$ Glieder auf L nur dadurch entstehen können, daß die entsprechenden Glieder in $h(x)$ und $g(x)$ auch auf den Polygonen liegen, so wird offenbar

$$(21) \quad f(x) \equiv g_1(x)h_1(x) \pmod{L},$$

wobei $g_1(x)$ und $h_1(x)$ die entsprechenden Faktoren der Seite von $g(x)$ und $h(x)$ sind. Aus der Definition der zugeordneten Polynome folgt dann nach (21)

$$F(x, y) \equiv G(x, y)H(x, y) \pmod{p, \varphi(x)},$$

wobei $G(x, y)$ und $H(x, y)$ die zugeordneten Polynome von $g(x)$ und $h(x)$ sind.

Satz 4. *Bildet man das Produkt mehrerer Polynome, welche alle geradlinige Polygone $(p, \varphi(x))$ mit derselben Neigung besitzen, so hat auch das Produkt ein geradliniges Polygon mit derselben Neigung, und das zugeordnete Polynom des Produkts ist $\pmod{p, \varphi(x)}$ kongruent dem Produkte der zugeordneten Polynome der Faktoren.*

Man kann auch im allgemeinen einen wichtigen Zusammenhang zwischen einem beliebigen $f(x)$ und den $f_i(x)$ der Seiten angeben, wo-

durch auch die Bezeichnung „Faktoren der Seiten“ für die $f_i(x)$ gerechtfertigt wird. Bildet man nämlich das Produkt

$$\psi(x) = \prod_{i=1}^k f_i(x),$$

so hat $\psi(x)$ nach Satz 2 dasselbe Polygon S wie $f(x)$; man kann aber auch den weiteren Satz beweisen:

Satz 5. *Es besteht die Kongruenz*

$$(22) \quad f(x) \equiv \prod_{i=1}^k f_i(x) \pmod{S}.$$

Die Richtigkeit dieser Kongruenz beweist man am einfachsten durch Induktion; für $k=1$ ist der Satz nach (20) selbstklar. Man kann daher voraussetzen, daß der Satz für alle Polynome bewiesen ist, wofür das entsprechende Polygon $k-1$ Seiten hat, und es soll bewiesen werden, daß er auch für alle Polynome mit k Polygonseiten richtig bleibt. Das Produkt

$$\varphi(x)^{l_k} \prod_{i=1}^{k-1} f_i(x)$$

hat daher, wie man voraussetzen darf, für die $k-1$ ersten Seiten von S dieselben darauf liegenden Glieder wie $f(x)$. Bildet man daher das Produkt

$$(23) \quad f_k(x) \prod_{i=1}^{k-1} f_i(x),$$

so ergibt also bei der Multiplikation

$$\varphi(x)^{l_k} \prod_{i=1}^{k-1} f_i(x)$$

diejenigen Glieder in $f(x)$, welche auf den $k-1$ ersten Seiten liegen. Diejenigen Glieder des Produkts (23), welche auf der k -ten Seite von S liegen, kann man aber, wie es leicht aus der früher angewandten Vektorendarstellung hervorgeht, nur dann erhalten, wenn man die Glieder in $f_k(x)$ mit dem letzten Gliede

$$p^{h_1 + \dots + h_{k-1}} R_{k-1, \varepsilon_{k-1}}(x) = p^{h_1 + \dots + h_{k-1}} R_{k,0}(x)$$

in $\prod_{i=1}^{k-1} f_i(x)$ multipliziert. Für diese Glieder auf S_k erhält man daher den Ausdruck

$$(24) \quad p^{h_1 + \dots + h_{k-1}} R_{k,0}(x) (\varphi(x)^{l_k} + S_{k,1}(x) p^{\varepsilon_k} \varphi(x)^{l_k - \lambda_k} + \dots + S_{k,\varepsilon_k}(x) p^{h_k}),$$

und da allgemein nach (13) und (14)

$$R_{k,0}(x) S_{k,j}(x) \equiv R_{k,0}(x) A_k(x) R_{k,j}(x) \equiv R_{k,j}(x) \pmod{p, \varphi(x)}$$

ist, so stimmt (24) mit (12) für $i=k$ überein. Der Satz 5 ist folglich bewiesen.

Aus dem Beweise folgt leicht in derselben Weise durch Induktion, daß bei jeder Zerlegung von der Form (22) eine gewisse Eindeutigkeit bestehen muß, und man erhält dadurch den Satz:

Satz 6. *Zerlegt man das Polynom $f(x)$ in Faktoren (mod S)*

$$f(x) \equiv \prod_{i=1}^k \psi_i(x) \pmod{S},$$

wobei das Polygon $(p, \varphi(x))$ eines Faktors $\psi_i(x)$ geradlinig ist, und dieselbe Neigung und Länge wie die Seite S_i von S hat, so ist diese Zerlegung für jede Seite eindeutig, d. h. es ist

$$\psi_i(x) \equiv f_i(x) \pmod{S_i},$$

und wenn das zugeordnete Polynom von $\psi_i(x)$ durch $\Psi_i(x, y)$ bezeichnet wird, so ist

$$F_i(x, y) \equiv \Psi_i(x, y) \pmod{p, \varphi(x)}.$$

Unter Anwendung der gegebenen Sätze kann man den Multiplikationssatz, Satz 2, wesentlich ergänzen. Es sei nämlich wie in § 2

$$(25) \quad f(x) = g(x)h(x),$$

wobei $f(x)$, $g(x)$ und $h(x)$ durch (10) und (11) gegeben sind. Nach Satz 2 ist also das Polygon S von $f(x)$ aus den Polygonen S' und S'' von $g(x)$ und $h(x)$ so entstanden, daß man die Seiten nach steigender Neigung ordnet; man kann dabei voraussetzen, daß es sich überall um Hauptpolygone handelt. Nach Satz 5 bestehen für $g(x)$ und $h(x)$ Kongruenzen von der Form

$$(26) \quad \begin{cases} g(x) \equiv \prod_{i=1}^{k'} g_i(x) \pmod{S'}, \\ h(x) \equiv \prod_{i=1}^{k''} h_i(x) \pmod{S''}, \end{cases}$$

wobei $g_i(x)$ und $h_i(x)$ die Faktoren der Seiten sind; weiter seien $G_i(x, y)$ und $H_i(x, y)$ die entsprechenden zugeordneten Polynome. Da nach (26) die rechtsstehenden Produkte dieselben Glieder auf den Polygonen wie $g(x)$ und $h(x)$ besitzen, so folgt leicht aus dem Beweise des Satzes 2, daß auch die Kongruenz

$$f(x) \equiv \prod_{i=1}^{k'} g_i(x) \prod_{i=1}^{k''} h_i(x) \pmod{S}$$

besteht.

Nach Satz 4 und Satz 6 folgt daraus weiter, daß wenn $G_m(x, y)$ und $H_n(x, y)$ zwei zugeordnete Polynome in $g(x)$ und $h(x)$ sind, welche zu Seiten S'_m und S''_n mit derselben Neigung gehören, so ist

$$F_i(x, y) \equiv G_m(x, y) H_n(x, y) \pmod{p, \varphi(x)},$$

wenn $F_i(x, y)$ zu der Seite S_i in S gehört, welche auch dieselbe Neigung besitzt. Wenn es aber z. B. in $g(x)$ keine Seite mit dieser Neigung gibt, so ist

$$F_i(x, y) \equiv H_m(x, y) \pmod{p, \varphi(x)}.$$

Man ersieht daraus die Richtigkeit des wichtigen Satzes:

Satz 7. *Wenn man das Produkt $f(x)$ von mehreren Polynomen bildet, so entsteht das Polygon $S(p, \varphi(x))$ des Produkts aus den Seiten der Polygone der Faktoren, indem man diese Seiten nach steigender Neigung ordnet; weiter ist für eine beliebige Seite von S das zugeordnete Polynom von $f(x)$ kongruent $\pmod{p, \varphi(x)}$ dem Produkte aller zugeordneten Polynome der Faktoren, welche Seiten mit derselben Neigung entsprechen.*

Kapitel 2.

Zusammenhang zwischen Polygonen und Idealeigenschaften.

§ 1.

Einleitende Untersuchungen.

Es soll nun allgemein der Zusammenhang zwischen den Newtonschen Polygonen einer beliebigen Primzahl p für ein gegebenes Polynom $f(x)$ und den Idealeigenschaften dieser Primzahl im entsprechenden Körper studiert werden. $f(x)$ soll daher im folgenden immer ein irreduzibles Polynom bezeichnen, und der entsprechende Körper K wird durch eine Wurzel ϑ der Gleichung

$$f(x) = x^n + a_1 x^{n-1} + \dots + a_n = 0$$

definiert.

Zunächst sei wie früher

$$(1) \quad f(x) \equiv \varphi_1(x)^{\alpha_1} \dots \varphi_s(x)^{\alpha_s} \pmod{p}$$

die Primfunktionzerlegung von $f(x) \pmod{p}$; aus dieser Zerlegung folgt bekanntlich, daß auch für alle $\alpha \geq 1$ eine Zerlegung

$$(2) \quad f(x) \equiv \Phi_1(x) \dots \Phi_s(x) \pmod{p^\alpha}$$

besteht, wie schon in Kap. 1 § 1 angegeben.

Aus dem ersten Hauptsatze I, Kap. 2 § 3 folgt aus (2) schon, daß für die Primzahl p in K eine Idealzerlegung

$$(3) \quad p = \mathfrak{A}_1 \dots \mathfrak{A}_s, \quad N(\mathfrak{A}_i) = p^{\alpha_i m_i}$$

besteht, wobei alle Ideale \mathfrak{A}_i zueinander relativ prim sind; das Produkt $\alpha_i m_i$ ist der Grad des entsprechenden Faktors $\Phi_i(x)$ in (2).

Man muß daher, um die Primidealzerlegung von p zu bestimmen, die weitere Zerlegung der Ideale \mathfrak{A} in (3) studieren.

Es sei

$$(4) \quad \Phi(x) \equiv \varphi(x)^a \pmod{p}$$

der entsprechende Faktor von \mathfrak{A} in (2). Weiter sei p ein beliebiges Primideal, das in \mathfrak{A} aufgeht; man kann dann sofort verschiedene Eigenschaften eines solchen Primideals bestimmen.

Nach dem Hauptsatze kann man nämlich erreichen, daß die Zahl $\Phi(\vartheta)$ durch eine beliebig hohe Potenz von p teilbar wird, und aus (4) folgt dann, daß auch die Zahl $\varphi(\vartheta)$ durch p teilbar sein muß.

Weiter folgt auch, daß wenn f der Grad von p ist, so muß f durch m teilbar sein, wenn m der Grad der Primfunktion $\varphi(x)$ ist. Denn eine Zahl $F(\vartheta)$, wobei $F(x)$ ein Polynom ist, kann offenbar nicht durch p teilbar sein, außer wenn

$$F(x) \equiv 0 \pmod{p, \varphi(x)}.$$

Andererseits ist aber sicher nach dem verallgemeinerten Fermatschen Satz

$$\vartheta^{p^f} - \vartheta \equiv 0 \pmod{p},$$

und das Polynom

$$x^{p^f} - x$$

muß daher \pmod{p} durch $\varphi(x)$ teilbar sein. Dies ist bekanntlich nur dann möglich, wenn m ein Teiler von f ist.

Zuletzt soll noch eine Eigenschaft des Primideals p abgeleitet werden, woraus die Bedeutung der Polygone klar hervortritt.

Aus dem Faktor $\Phi(x)$ bildet man die zugehörige Entwicklung $(p, \varphi(x))$

$$(5) \quad \Phi(x) = \sum_{i=1}^a Q_i(x) p^{\alpha_i} \varphi(x)^{\alpha-i}$$

und konstruiert das zugehörige Polygon S von $\Phi(x)$. Das Polygon S wird ein Hauptpolygon sein und ist nach Kap. 1 mit dem Hauptpolygone von $f(x) \pmod{p, \varphi(x)}$ identisch. Aus dem früher Bewiesenen folgt auch noch, daß die beiden Polygone auch dieselben zugehörigen Faktoren und Polynome der Seiten haben werden. Für das Polygon S werden die früher gebrauchten Bezeichnungen angewandt.

Weiter sei das Ideal \mathfrak{A} und also auch die Primzahl p genau durch p^e teilbar, während die Zahl $\varphi(\vartheta)$ genau durch p^t teilbar sein soll. Es soll nun gezeigt werden, wie die Exponenten e und t von der Form des Polygons S abhängen.

Wird in der Entwicklung (5) $x = \vartheta$ gesetzt, so wird ein Glied

$$Q_i(\vartheta) p^{\alpha_i} \varphi(\vartheta)^{\alpha-i}$$

genau durch die Potenz

$$p^{\alpha_i e + (a-i)t}$$

teilbar. Zu jedem Gliede gibt es daher einen zugehörigen Exponenten

$$(6) \quad \alpha_i e + (a - i)t \quad (i = 0, 1, \dots, \alpha),$$

der durch die Lage des entsprechenden Punktes (i, α_i) vollständig bestimmt ist.

Die Zahl $\Phi(\vartheta)$ kann aber nach dem Hauptsatze durch eine beliebig hohe Potenz von p teilbar werden, wenn man nur in (2) den Exponenten α passend groß wählt. Im folgenden wird jedenfalls vorausgesetzt, daß α fest und größer als die Endordinate α_a vom Polygone S gewählt wird.

Unter den Exponenten (6) wird es gewiß einige geben, welche einen absolut kleinsten Wert haben. Da aber $\Phi(\vartheta)$, wie bemerkt, bei einer passenden Wahl von α durch eine beliebig hohe Potenz von p teilbar wird, so muß es unter den Zahlen (6) mindestens zwei oder mehrere geben, welche diesen Minimalwert haben. Die entsprechenden Punkte müssen alle auf einer Geraden liegen, denn aus

$$\alpha_i e + (a - i)t = \alpha_j e + (a - j)t$$

folgt

$$(7) \quad \frac{\alpha_i - \alpha_j}{i - j} = \frac{t}{e}.$$

Weiter muß aber diese Gerade mit einer Seite des Polygons zusammenfallen. Denn zunächst liegen auf dieser Geraden immer Punkte (i, α_i) , und weiter kann es unterhalb dieser Geraden keine abbildenden Punkte geben, indem dann für die entsprechenden Glieder der Exponent (6) noch kleiner würde.

Nimmt man an, daß die Gerade etwa mit der ν -ten Seite zusammenfällt, so wird gesagt, das *Primideal* p gehöre zur ν -ten Seite S_ν . Da die Neigungszahl von S_ν nach Kap. 1, § 1 gleich $\frac{\kappa_\nu}{\lambda_\nu}$ ist, so wird nach (7)

$$\frac{t}{e} = \frac{\kappa_\nu}{\lambda_\nu},$$

oder da κ_ν zu λ_ν relativ prim ist, erhält man

$$e = \varrho \lambda_\nu, \quad t = \varrho \kappa_\nu,$$

wobei ϱ ganz rational ist. Man kann diese vorläufigen Resultate folgendermaßen zusammenfassen:

Satz 1. *Hat $f(x) \pmod{p}$ die Primfunktionzerlegung (1), so hat die Primzahl p in K die Idealzerlegung*

$$(8) \quad p = \mathfrak{A}_1 \dots \mathfrak{A}_s, \quad N(\mathfrak{A}_i) = p^{a_i m_i},$$

wobei die Ideale \mathfrak{A}_i zueinander relativ prim sind. Wenn \mathfrak{p} ein beliebiges Primideal vom Grade f und Ordnung e ist, das in einem Ideale \mathfrak{A} aufgeht, so geht \mathfrak{p} gleichzeitig in der entsprechenden Zahl $\varphi(\vartheta)$ auf, etwa in einer Potenz \mathfrak{p}^t , und die Gradzahl f von \mathfrak{p} ist durch den Grad m von $\varphi(x)$ teilbar. Konstruiert man weiter das Polygon $(\mathfrak{p}, \varphi(x))$ von $f(x)$, so muß \mathfrak{p} zu einer bestimmten Seite S_ν dieses Polygons gehören, indem also Relationen von der Form

$$(9) \quad e = \varrho \lambda_\nu, \quad t = \varrho \kappa_\nu$$

bestehen, wobei ϱ eine ganze rationale Zahl ist.

Später wird auch gezeigt, daß umgekehrt zu jeder Seite S_ν des Polygons entsprechende Primideale existieren.

Man sieht leicht ein, daß wenn ein Primideal \mathfrak{p} zur ν -ten Seite gehört, so sind alle Glieder in der Entwicklung $(\mathfrak{p}, \varphi(x))$ von $f(x)$, welche auf der ν -ten Seite liegen, für $x = \vartheta$ durch dieselbe Potenz von \mathfrak{p} teilbar. Die Summe dieser Glieder ist nämlich in (12), Kap. 1 angegeben, und hier ist ein Glied

$$R_{\nu,j}(\vartheta) \mathfrak{p}^{j\kappa_\nu} \varphi(\vartheta)^{l_\nu - j\lambda_\nu}$$

durch eine Potenz von \mathfrak{p} mit dem Exponenten

$$j\kappa_\nu e + (l_\nu - j\lambda_\nu)t = tl_\nu = e h_\nu$$

teilbar; die letzten Reduktionen folgen aus (9) und (7), Kap. 1. In der Summe

$$A_\nu(\vartheta) = R_{\nu,0}(\vartheta) \varphi(\vartheta)^{l_\nu} + R_{\nu,1}(\vartheta) \mathfrak{p}^{\kappa_\nu} \varphi(\vartheta)^{l_\nu - \lambda_\nu} + \dots + R_{\nu,\varepsilon_\nu}(\vartheta) \mathfrak{p}^{h_\nu}$$

sind also alle Glieder sicher durch $\mathfrak{p}^{tl_\nu} = \mathfrak{p}^{e h_\nu}$ teilbar. Da aber die Summe der Glieder auf der ν -ten Seite aus denjenigen Gliedern von $f(x)$ besteht, welche für $x = \vartheta$ durch die kleinste Potenz von \mathfrak{p} teilbar sind, so wird die Summe selbst durch eine höhere Potenz von \mathfrak{p} teilbar sein müssen, indem sonst $\Phi(\vartheta)$ (oder $f(\vartheta)$) auch nur durch diese minimale Potenz teilbar sein würde. Es folgt daraus, daß für $A_\nu(\vartheta)$ die Kongruenz

$$A_\nu(\vartheta) \equiv 0 \pmod{\mathfrak{p}^{e h_\nu + 1}}$$

bestehen muß. Aus der Weise, worin man in § 3, Kap. 1 den Faktor $f_\nu(x)$ der ν -ten Seite aus $A_\nu(x)$ ableitete, folgt sofort, daß auch für $f_\nu(\vartheta)$ die Kongruenz

$$f_\nu(\vartheta) \equiv 0 \pmod{\mathfrak{p}^{e h_\nu + 1}}$$

besteht.

Satz 2. Wenn \mathfrak{p} ein Primideal der ν -ten Seite und

$$f_\nu(x) = \varphi(x)^{l_\nu} + S_{\nu,1}(x) \mathfrak{p}^{\kappa_\nu} \varphi(x)^{l_\nu - \lambda_\nu} + \dots + S_{\nu,\varepsilon_\nu}(x) \mathfrak{p}^{h_\nu}$$

der entsprechende Faktor dieser Seite ist, so ist die Zahl $f_\nu(\vartheta)$ mindestens durch $\mathfrak{p}^{e h_\nu + 1} = \mathfrak{p}^{t l_\nu + 1}$ teilbar.

§ 2.

Eigenschaften der unzerlegbaren Polynome $(\text{mod } p^\alpha)$.

Um nun den allgemeinen Satz über den Zusammenhang zwischen Polygonen und Idealeigenschaften zu beweisen, müssen einige allgemeine Eigenschaften von solchen Polynomen abgeleitet werden, welche $(\text{mod } p^\alpha)$ für ein genügend großes α irreduzibel sind. Die notwendige und hinreichende Bedingung dafür, daß ein Polynom $F(x) (\text{mod } p^\alpha)$ $\alpha > \delta$ irreduzibel ausfällt, ist, nach I, Kap. 1, daß $F(x)$ schon $(\text{mod } p^{\delta+1})$ irreduzibel ist, wobei die Diskriminante von $F(x)$ genau durch p^δ teilbar ist.

Wenn $F(x)$ vom Grade q ist, so definiert die Gleichung

$$F(\vartheta') = 0$$

einen Körper $k(\vartheta')$ q -ten Grades; in k ist p die Potenz eines Primideals

$$(10) \quad p = \mathfrak{p}^e, \quad N(\mathfrak{p}) = p^f, \quad ef = q,$$

und umgekehrt folgt auch aus dem ersten Hauptsatze, daß in jedem Körper, wo p eine Potenz eines Primideals ist, wird die entsprechende definierende Gleichung $(\text{mod } p^\alpha)$ irreduzibel.

Es sollen nun verschiedene Eigenschaften solcher $(\text{mod } p^\alpha)$ unzerlegbaren Polynome abgeleitet werden. Zunächst folgt aus I, Kap. 1, § 3, daß $F(x) (\text{mod } p)$ kongruent einer Potenz einer Primfunktion $(\text{mod } p)$

$$(11) \quad F(x) \equiv \varphi(x)^\alpha (\text{mod } p)$$

sein muß. Nach Satz 1 ist der Grad m der Primfunktion $\varphi(x)$ ein Teiler von f , und weiter geht \mathfrak{p} in der Zahl $\varphi(\vartheta')$ in einer Potenz \mathfrak{p}^t auf. Wenn in (11) speziell $\alpha = 1$ ist, wird die Irreduzibilität von $F(x) (\text{mod } p^\alpha)$ selbstklar.

Man konstruiert nun das Polygon $S(p, \varphi(x))$ von $F(x)$. Aus (11) folgt sofort, daß dies Polygon ein Hauptpolygon ist; es soll aber bewiesen werden, daß S geradlinig ist, d. h. aus einer einzigen Seite besteht.

Aus Satz 1 folgt jedenfalls, daß das Primideal \mathfrak{p} in (10) zu einer bestimmten Seite S_ν von S gehört, und es bestehen also die Relationen (9). Für den entsprechenden Faktor $f_\nu(x)$ dieser Seite besteht nach Satz 2 die Kongruenz

$$(12) \quad f_\nu(\vartheta') = \varphi(\vartheta')^{l_\nu} + S_{\nu,1}(\vartheta') p^{s_\nu} \varphi(\vartheta')^{l_\nu - l_{\nu-1}} + \dots + S_{\nu,s_\nu}(\vartheta') p^{h_\nu} \\ \equiv 0 (\text{mod } \mathfrak{p}^{e h_\nu + 1}),$$

oder wenn man durch p^{h_ν} dividiert, erhält man die neue Kongruenz

$$(13) \quad F(\vartheta', \Theta_\nu(\vartheta')) = \Theta_\nu^{s_\nu} + S_{\nu,1}(\vartheta') \Theta_\nu^{s_\nu - 1} + \dots + S_{\nu,s_\nu}(\vartheta') \equiv 0 (\text{mod } \mathfrak{p}),$$

wobei also nach Kap. 1, § 3 $F(x, y)$ das zugeordnete Polynom der ν -ten Seite ist, und wo weiter wie in (18), Kap. 1

$$\Theta_v = \Theta_v(\vartheta') = \frac{\varphi(\vartheta')^{\lambda_v}}{p^{\alpha_v}}$$

gesetzt worden ist. Die Zahl Θ_v ist offenbar ganz und nicht durch p teilbar, denn nach (9) enthält der Zähler genau dieselbe Potenz von p wie der Nenner. Der Kürze wegen wird in (13)

$$(14) \quad \Phi = F(\vartheta', \Theta(\vartheta'))$$

gesetzt, und man bildet die Gleichung

$$(15) \quad \Phi^q + e_1 \Phi^{q-1} + \dots + e_q = 0,$$

welcher die ganze Zahl Φ genügt⁵⁾. Da nach (13) Φ durch p teilbar ist, folgt leicht, daß in (15) alle Koeffizienten e_i durch p teilbar sind⁶⁾.

Wird in (15) der Wert (14) von Φ eingesetzt, und multipliziert man mit $p^{q h_v}$, so folgt, indem man die Identität

$$f_v(\vartheta') = p^{h_v} F(\vartheta', \Theta_v(\vartheta'))$$

anwendet,

$$(16) \quad f_v(\vartheta')^q + e_1 p^{h_v} f_v(\vartheta')^{q-1} + \dots + e_q p^{q h_v} = 0,$$

und diese Gleichung (16) ist mit einer Identität

$$(17) \quad f_v(x)^q + e_1 p^{h_v} f_v(x)^{q-1} + \dots + e_q p^{q h_v} = F(x) G(x)$$

gleichbedeutend, wobei $G(x)$ ein Polynom ist.

Das Polygon $(p, \varphi(x))$ der linken Seite in (17) ist leicht bestimmbar. Das Polygon von $f_v(x)^q$ ist offenbar eine Gerade L mit derselben Neigung wie S_v . Die Glieder eines Ausdrucks

$$p^{i k_v} f_v(x)^{q-i}$$

liegen alle auf derselben Geraden L , wie man leicht einsieht. Da aber diese in (17) mit der durch p teilbaren Zahl e_i multipliziert werden sollen, ergeben sich daraus nur Glieder, welche oberhalb L liegen. Die linke Seite von (17) hat also ein geradliniges Polygon, und nach dem Multiplikationssatz für Polygone folgt daraus aus der Identität (17), daß auch $F(x)$ ein geradliniges Polygon hat, w. z. b. w.

⁵⁾ Man kann hierbei voraussetzen, daß Φ eine primitive Zahl des Körpers ist; denn wenn dies nicht der Fall ist, bildet man nur die Gleichung der Zahl $\Phi' = \Phi + r p \vartheta'$, wo bekanntlich die ganze rationale Zahl r so gewählt werden kann, daß Φ' primitiv ist. Im folgenden kommen ähnliche Fälle öfters vor, und man kann dann immer analog eine primitive Zahl erhalten.

⁶⁾ Man vergleiche den allgemeinen Satz 18 in meiner Arbeit: „Zur Theorie der algebraischen Körper“, Acta math. 44 (1923), S. 219–314.

Für ein (mod p^α) irreduzibles Polynom $F(x)$ gibt es also nur einen einzigen Faktor der Seite

$$f_1(x) = \varphi(x)^l + p^\kappa S_1(x) \varphi(x)^{l-1} + \dots + p^h S_\varepsilon(x),$$

wobei die Relationen

$$l = \varepsilon \lambda, \quad h = \varepsilon \kappa, \quad (\lambda, \kappa) = 1, \quad lm = q$$

bestehen müssen. Das zugeordnete Polynom der Seite heißt

$$(18) \quad F(x, y) = y^\varepsilon + S_1(x) y^{\varepsilon-1} + \dots + S_\varepsilon(x),$$

wobei wie früher die Identität

$$p^h F(x, \Theta(x)) = f_1(x), \quad \Theta(x) = \frac{\varphi(x)^\lambda}{p^\kappa}$$

besteht. Aus einer früheren Bemerkung folgt, daß die Zahl $\Theta = \Theta(\vartheta')$ ganz und nicht durch p teilbar ist.

Es sollen nun die Eigenschaften des Primideals p genauer studiert werden. Nach Satz 1 bestehen die Relationen

$$e = \varrho \lambda, \quad t = \varrho \kappa,$$

wenn die Zahl $\varphi(\vartheta')$ genau durch p^t teilbar ist. Durch Betrachtung der Zahlen des Ringes $R(\vartheta')$ ist in § 1 bewiesen, daß der Grad f von p durch m teilbar ist. Wenn man nun alle Zahlen von der Form

$$(19) \quad A(\vartheta', \Theta) = A_0(\vartheta') \Theta^a + A_1(\vartheta') \Theta^{a-1} + \dots + A_\alpha(\vartheta')$$

betrachtet, wobei die $A_i(\vartheta')$ Polynome in ϑ' sind, kann man einen schärferen Satz über f beweisen.

Speziell ist von Interesse, wann eine Zahl (19) durch p teilbar ist. Dies ist natürlich immer der Fall, wenn in (19) alle Zahlen $A_i(\vartheta')$ durch p teilbar sind, d. h. wenn die identische Kongruenz

$$(20) \quad A(x, y) \equiv 0 \pmod{p, \varphi(x)}$$

besteht. Aus der Kongruenz (13) folgt aber, daß es auch solche durch p teilbare Zahlen (19) gibt, wofür die Kongruenz (20) nicht erfüllt ist. Es gibt daher sicher ein Polynom kleinsten Grades in y

$$B(x, y) = y^m + B_1(x) y^{m-1} + \dots + B_m(x),$$

das die Eigenschaft hat, daß die entsprechende Zahl $B(\vartheta', \Theta)$ durch p teilbar ist, während $B(x, y) \not\equiv 0 \pmod{p, \varphi(x)}$. $B(x, y)$ wird dann offenbar eine Primfunktion (mod $p, \varphi(x)$), und wenn eine Zahl (19) durch p teilbar ist, muß notwendigerweise $A(x, y) \pmod{p, \varphi(x)}$ durch $B(x, y)$ teilbar sein.

Man beweist weiter analog wie in § 1, daß der Grad f von p durch $m_1 m$ teilbar ist; denn nach dem Fermatschen Satz ist

$$\Theta^{p^f} - \Theta = \Theta^{p^{mf_1}} - \Theta \equiv 0 \pmod{p},$$

wobei im letzten Gliede $f = m f_1$ gesetzt ist. Wie eben bemerkt folgt daraus, daß das Polynom

$$(21) \quad y^{p^{f_1 m}} - y \pmod{p, \varphi(x)}$$

durch $B(x, y)$ teilbar ist. Das Polynom (21) ist aber bekanntlich $\pmod{p, \varphi(x)}$ kongruent dem Produkte aller Primfunktionen $P(x, y)$, deren Grade (in y) Teiler von f_1 sind; also ist m_1 ein Teiler von f_1 , w. z. b. w.

Zwischen der Primfunktion $B(x, y)$ und dem definierenden Polynom $F(x)$ besteht eine enge Verbindung. Bildet man nämlich die Gleichung, welcher die Zahl $B = B(\vartheta', \Theta)$ genügt:

$$B^q + b_1 B^{q-1} + \dots + b_q = 0,$$

so sind nach einer früheren Bemerkung alle Koeffizienten b_i durch p teilbar. Multipliziert man die ganze Gleichung mit $p^{m_1 \times q}$, um die Nenner wegzubringen, und wird

$$(22) \quad C(x) = p^{m_1 \times} B(x_1 \Theta(x))$$

gesetzt, so folgt leicht wie früher, daß eine Identität

$$C(x)^q + p^{m_1 \times} b_1 C(x)^{q-1} + \dots + b_q p^{m_1 \times q} = F(x) G(x)$$

bestehen muß, wobei beide Seiten Polynome sind. Wie früher zeigt man, daß das Polygon der linken Seite mit dem Polygon des Gliedes $C(x)^q$ zusammenfällt. Das Polygon der linken Seite ist folglich nach (22) eine Gerade, und das zugeordnete Polynom dieser Seite ist gleich $B(x, y)^q$. Nach Satz 7, Kap. 1 schließt man daraus, daß das zugeordnete Polynom $F(x, y)$ (18) von $F(x)$ ein Teiler von $B(x, y)^q \pmod{p, \varphi(x)}$ ist, und $F(x, y)$ ist daher selbst $\pmod{p, \varphi(x)}$ eine Potenz von $B(x, y)$:

$$F(x, y) \equiv B(x, y)^{\alpha_1} \pmod{p, \varphi(x)}.$$

Satz 3. Wenn $F(x) \pmod{p^\alpha}$ $\alpha > \delta$ ein irreduzibles Polynom ist, so wird

$$F(x) \equiv \varphi(x)^a \pmod{p},$$

wobei $\varphi(x)$ eine Primfunktion m -ten Grades \pmod{p} ist. Bildet man das Polygon $(p, \varphi(x))$, so ist dies eine Gerade, und wenn $F(x, y)$ das zugeordnete Polynom von $F(x)$ für diese Seite ist, besteht die Kongruenz

$$(23) \quad F(x, y) \equiv F_1(x, y)^{\alpha_1} \pmod{p, \varphi(x)},$$

wobei $F_1(x, y)$ eine Primfunktion m_1 -ten Grades in $y \pmod{p, \varphi(x)}$ ist.

Dieser Satz enthält notwendige, aber keineswegs hinreichende Bedingungen für ein unzerlegbares Polynom $(\text{mod } p^\alpha)$. Nur in dem Falle, daß in (23) $a_1 = 1$ ist, muß auch nach Satz 7, Kap. 1 das entsprechende Polynom $F(x) (\text{mod } p^\alpha)$ irreduzibel sein.

Über den Primidealteiler \mathfrak{p} von p im entsprechenden Körper weiß man also, daß die Ordnung e durch λ teilbar ist, während der Grad f durch $m m_1$ teilbar wird. In Satz 3 ist $F_1(x, y)$ an der Stelle der früheren Bezeichnung $B(x, y)$ geschrieben worden; es besteht daher die wichtige Kongruenz

$$(24) \quad F_1(\vartheta', \Theta) \equiv 0 \pmod{\mathfrak{p}}.$$

Im Spezialfalle $a_1 = 1$ ist der Grad von $F(x)$ gleich $\lambda m m_1$, und in diesem Falle ist einfach nach dem Hauptsatze $e = \lambda$, $f = m m_1$.

§ 3.

Der allgemeine Satz über Polygone und Idealeigenschaften.

Nach diesen vorbereitenden Untersuchungen gehe ich wieder zum allgemeinen Teil über: Das Polynom $f(x)$ soll die Primfunktionzerlegung (1) $(\text{mod } p)$ haben, woraus wieder die Zerlegung (2) $(\text{mod } p^\alpha)$ folgt. Für einen beliebigen Faktor $\Phi(x)$ in (2), wofür also die Kongruenz (4) besteht, bildet man das entsprechende Polygon $(p, \varphi(x))$, wofür die gewöhnlichen Bezeichnungen angewandt werden. Speziell ist also $f_i(x)$ der Faktor und weiter $F_i(x, y)$ das zugeordnete Polynom der i -ten Seite. Im folgenden spielt die Primfunktionzerlegung von $F_i(x, y) (\text{mod } p, \varphi(x))$

$$(25) \quad F_i(x, y) = F_1^{(\delta)}(x, y)^{a_1^{(\delta)}} \dots F_{\varepsilon_i}^{(\delta)}(x, y)^{a_{\varepsilon_i}^{(\delta)}} \pmod{p, \varphi(x)}$$

eine wichtige Rolle; hier ist allgemein eine Primfunktion $F_i(x, y) (\text{mod } p, \varphi(x))$ von der Form

$$F_j^{(\delta)}(x, y) = y^{m_j^{(\delta)}} + A_{j,1}^{(\delta)}(x) y^{m_j^{(\delta)}-1} + \dots + A_{j,m_j^{(\delta)}}^{(\delta)}(x),$$

wobei nach (16), Kap. 1 die Relation

$$\varepsilon_i = a_1^{(\delta)} m_1^{(\delta)} + \dots + a_{\varepsilon_i}^{(\delta)} m_{\varepsilon_i}^{(\delta)}$$

bestehen muß, indem ε_i der Grad von $F_i(x, y)$ in y ist.

Es ist nun möglich eine weitere Zerlegung eines Faktors $\Phi(x) (\text{mod } p^\alpha)$ anzugeben; denn aus Satz 3 folgt jedenfalls, daß $\Phi(x)$ kein unzerlegbares Polynom ist. Wenn daher $\Psi(x)$ ein $(\text{mod } p^\alpha)$ unzerlegbarer Faktor von $\Phi(x) (\text{mod } p^\alpha)$ ist, so muß nach dem Multiplikationssatze das geradlinige Polygon von $\Psi(x)$ dieselbe Neigung wie eine Seite S_i des Polygons von $\Phi(x)$ haben. Wenn aber weiter $\Psi(x, y)$ das zugeordnete Polynom von $\Psi(x)$ ist, so ist nach Satz 3 $\Psi(x, y) (\text{mod } p, \varphi(x))$ kon-

gruent einer Potenz einer Primfunktion, und da nach Satz 7, Kap. 1 $\Psi(x, y) \pmod{p, \varphi(x)}$ ein Teiler von $F_i(x, y)$ sein muß, so folgt nach (25), daß $\Psi(x, y)$ die Form

$$(26) \quad \Psi(x, y) \equiv F_j^{(i)}(x, y)^b \pmod{p, \varphi(x)}$$

hat, wo $b \leq a_j^{(i)}$ sein muß. Bildet man nun das Produkt aller Faktoren $\Psi(x)$ von $\Phi(x) \pmod{p^\alpha}$, welche Polygone mit derselben Neigung haben, und wofür in (26) dieselbe Primfunktion $F_j^{(i)}(x, y)$ vorkommt, so erhält man einen Faktor $\Phi_j^{(i)}(x)$ von $\Phi(x) \pmod{p^\alpha}$, der auch ein geradliniges Polygon mit derselben Neigung hat, und wofür das zugeordnete Polynom gleich $F_j^{(i)}(x, y)^{a_j^{(i)}}$ ist; der Grad von $\Phi_j^{(i)}(x)$ ist also gleich $a_j^{(i)} m_j^{(i)} \lambda_i m$. Man hat daher den Satz bewiesen:

Satz 4. *Es sei $f(x)$ ein Polynom mit der Primfunktionzerlegung*

$$(27) \quad f(x) \equiv \varphi_1(x)^{a_1} \dots \varphi_s(x)^{a_s} \pmod{p}.$$

Daraus folgt sofort eine Zerlegung

$$(28) \quad f(x) \equiv \Phi_1(x) \dots \Phi_s(x) \pmod{p^\alpha},$$

wobei allgemein

$$(29) \quad \Phi_i(x) \equiv \varphi_i(x)^{a_i} \pmod{p}$$

ist. Um die weitere Zerlegung eines Faktors $\Phi(x) \pmod{p^\alpha}$ in (28) zu bestimmen, konstruiert man das entsprechende Polygon $(p, \varphi(x))$ von $\Phi(x)$ (oder $f(x)$). Wenn dann allgemein das entsprechende zugeordnete Polynom $F_i(x, y)$ der i -ten Seite die Primfunktionzerlegung

$$(30) \quad F_i(x, y) \equiv F_1^{(i)}(x, y)^{a_1^{(i)}} \dots F_{t_i}^{(i)}(x, y)^{a_{t_i}^{(i)}} \pmod{p, \varphi(x)}$$

hat, so besteht die Zerlegung

$$(31) \quad \Phi(x) \equiv \prod_{i=1}^k \prod_{j=1}^{t_i} \Phi_j^{(i)}(x) \pmod{p^\alpha},$$

wobei ein Faktor $\Phi_j^{(i)}(x)$ ein geradliniges Polygon mit derselben Neigung wie die i -te Seite des Polygons von $\Phi(x)$ hat. Der Grad von $\Phi_j^{(i)}(x)$ ist gleich $m m_j^{(i)} \lambda_i a_j^{(i)}$ und für das zugeordnete Polynom $\Phi_j^{(i)}(x, y)$ von $\Phi_j^{(i)}(x)$ besteht die Kongruenz

$$\Phi_j^{(i)}(x, y) \equiv F_j^{(i)}(x, y)^{a_j^{(i)}} \pmod{p, \varphi(x)}.$$

Dieser Satz ist auch für reduzible Polynome richtig. Im allgemeinen sind die Faktoren $\Phi_j^{(i)}(x)$ in (31) weiter zerlegbar $\pmod{p^\alpha}$; im Falle $a_j^{(i)} = 1$ ist aber nach § 2 die Irreduzibilität des entsprechenden $\Phi_j^{(i)}(x)$ sofort ersichtlich.

Unter Anwendung des ersten Hauptsatzes folgt nun aus Satz 4 leicht der Fundamentalsatz über den Zusammenhang zwischen dem Polygone

und den Idealeigenschaften der Primzahl p . Schon aus der Zerlegung (28) folgt wie in Satz 1 die Idealzerlegung

$$p = \mathfrak{A}_1 \dots \mathfrak{A}_s, \quad N(\mathfrak{A}_i) = p^{a_i m_i},$$

und aus der weiteren Zerlegung (31) eines Faktors $\Phi(x)$ folgt auch entsprechend eine weitere Zerlegung des zugehörigen Ideals \mathfrak{A}

$$\mathfrak{A} = \prod_{i=1}^k \prod_{j=1}^{t_i} \mathfrak{B}_j^{(i)}, \quad N(\mathfrak{B}_j^{(i)}) = p^{m_j^{(i)} a_j^{(i)} \lambda_i},$$

wobei alle Ideale $\mathfrak{B}_j^{(i)}$ zueinander relativ prim sind. Ein Ideal $\mathfrak{B}_j^{(i)}$ entspricht also dem Faktor $\Phi_j^{(i)}(x)$; wenn daher $p_j^{(i)}$ ein Primidealteiler von $\mathfrak{B}_j^{(i)}$ ist, so gibt es nach dem Hauptsatze einen entsprechenden $(\text{mod } p^a)$ irreduziblen Faktor $\Psi(x)$ von $f(x)$, dem $p_j^{(i)}$ entspricht. $\Psi(x)$ wird dann auch $(\text{mod } p^a)$ ein Teiler von $\Phi_j^{(i)}(x)$ und hat folglich ein geradliniges Polygon mit der Neigung $\frac{x_i}{\lambda_i}$. Wenn daher p und also auch $\mathfrak{B}_j^{(i)}$ genau durch $p_j^{(i)e}$ teilbar ist, so folgt wie früher, daß e durch λ_i teilbar sein muß, und folglich kann man, da dies für alle Primidealteiler von $\mathfrak{B}_j^{(i)}$ gilt, ein Ideal $\mathfrak{C}_j^{(i)}$ so bestimmen, daß

$$\mathfrak{B}_j^{(i)} = \mathfrak{C}_j^{(i)\lambda_i}, \quad N(\mathfrak{C}_j^{(i)}) = p^{m_j^{(i)} a_j^{(i)}}$$

ist. Man hat daher den Fundamentalsatz bewiesen:

Satz 5. *Es sei $f(x)$ ein irreduzibles Polynom von der Form des Satzes 4 und weiter sei $K(\vartheta)$ der entsprechende Körper. Aus der Zerlegung (27) folgt dann zunächst, daß die Primzahl p in $K(\vartheta)$ die Idealzerlegung*

$$(32) \quad p = \mathfrak{A}_1 \dots \mathfrak{A}_s, \quad N(\mathfrak{A}_i) = p^{a_i m_i}$$

hat, wobei die Ideale \mathfrak{A}_i zueinander relativ prim sind. Um die weitere Zerlegung eines Ideals \mathfrak{A} zu bestimmen, konstruiert man das entsprechende Polygon $(p, \varphi(x))$ von $f(x)$. Wenn dann allgemein das zugeordnete Polynom einer Seite die Primfunktionzerlegung (30) $(\text{mod } p, \varphi(x))$ hat, so besteht für ein Ideal \mathfrak{A} in (32) die weitere Zerlegung

$$\mathfrak{A} = \prod_{i=1}^k \prod_{j=1}^{t_i} \mathfrak{C}_j^{(i)\lambda_i}, \quad N(\mathfrak{C}_j^{(i)}) = p^{m_j^{(i)} a_j^{(i)}},$$

wobei alle Ideale $\mathfrak{C}_j^{(i)}$ zueinander relativ prim sind.

Wenn $p_j^{(i)}$ ein Primidealteiler von $\mathfrak{C}_j^{(i)}$ ist, so folgt wie früher leicht, daß der Grad von $p_j^{(i)}$ durch $m_j^{(i)}$ teilbar ist. Die Ideale $\mathfrak{C}_j^{(i)}$ sind im allgemeinen keine Primideale; nur im Falle $a_j^{(i)} = 1$ weiß man sicher, daß der entsprechende Faktor $\Phi_j^{(i)}(x) (\text{mod } p^a)$ unzerlegbar ist, und dann wird $\mathfrak{C}_j^{(i)} = p_j^{(i)}$ ein Primideal, und zwar vom Grade $m_j^{(i)}$ und Ordnung λ_i .

§ 4.

Über die regulären Gleichungen des Körpers.

Wie schon bemerkt, gibt der Satz 5 allgemein nicht die volle Zerlegung, d. h. die Primidealzerlegung von p ; unter speziellen Voraussetzungen kann man aber auch wirklich die Primidealzerlegung erhalten. Ein irreduzibles Polynom $f(x)$ soll *regulär in bezug auf p* heißen, wenn in der Primfunktionzerlegung (30) der zugeordneten Polynome der Seiten alle $a_j^{(s)} = 1$ sind, d. h. wenn mehrfache Faktoren (mod $p, \varphi(x)$) nicht vorkommen, und dies soll für alle die s verschiedenen Polygone von $f(x)$ gelten, die es nach (27) gibt. In diesem Falle gibt der Satz 5 auch wirklich die Primidealzerlegung von p , indem alle $\mathfrak{C}_j^{(s)}$ Primideale sind.

Satz 6. *Es sei $f(x)$ ein reguläres Polynom, wofür die Primfunktionzerlegung (27) (mod p) besteht; im entsprechenden Körper ist also*

$$p = \mathfrak{A}_1 \dots \mathfrak{A}_s, \quad N(\mathfrak{A}_i) = p^{m_i a_i}.$$

Hat dann allgemein im Polygone $(p, \varphi(x))$ von $f(x)$ das zugeordnete Polynom $F_i(x, y)$ der i -ten Seite die Primfunktionzerlegung

$$F_i(x, y) \equiv F_1^{(i)}(x, y) \dots F_{t_i}^{(i)}(x, y) \pmod{p, \varphi(x)},$$

wobei $m_j^{(i)}$ der Grad von $F_j^{(i)}(x, y)$ in y ist, so hat ein Ideal \mathfrak{A} die Primidealzerlegung

$$\mathfrak{A} = \prod_{i=1}^k \prod_{j=1}^{t_i} \mathfrak{p}_j^{(i) \lambda_i}, \quad N(\mathfrak{p}_j^{(i)}) = p^{m_j^{(i)}}.$$

Für die Zahlen $m_j^{(i)}$ besteht nach § 3 die Relation

$$(33) \quad \varepsilon_i = m_1^{(i)} + \dots + m_{t_i}^{(i)}.$$

Der Dedekindsche Satz⁷⁾ über den Zusammenhang zwischen den höheren Kongruenzen und der Primidealzerlegung ist, wie man leicht sieht, ein einfacher Spezialfall dieses Satzes. Wie man aber weiß, gibt es Körper, worin keine Gleichungen von der Dedekindschen Form existieren, indem die Primzahl p ein sogenannter gemeinsamer, außerwesentlicher Diskriminantenteiler sein kann, und durch diese Tatsache verliert der Dedekindsche Satz viel an Bedeutung. Man kann aber hier den folgenden wichtigen Existenzsatz beweisen, wodurch man auch eine neue Normalform der Gleichungen des Körpers erhält:

Satz 7. *Bei einer gegebenen Primzahl p existieren in jedem Körper entsprechende reguläre Gleichungen.*

⁷⁾ R. Dedekind, „Über den Zusammenhang zwischen der Theorie der Ideale und der höheren Kongruenzen“, Göttinger Abhandlungen 1878, S. 15. Man vergleiche auch I, Kap. 3, § 1.

Man nimmt an, daß die Form der Primidealzerlegung von p im Körper gegeben ist, und es seien

$$p_1^{(f)}, p_2^{(f)}, \dots, p_{k_f}^{(f)}$$

die verschiedenen Primideale f -ten Grades, welche in p aufgehen, und weiter sollen

$$(34) \quad e_1^{(f)}, e_2^{(f)}, \dots, e_{k_f}^{(f)}$$

die entsprechenden Ordnungszahlen sein. Wird dann

$$\alpha_f = p_1^{(f)e_1^{(f)}} \dots p_{k_f}^{(f)e_{k_f}^{(f)}}$$

gesetzt, so ist also

$$(35) \quad p = \prod_f \alpha_f, \quad N(\alpha_f) = p^{fL_f},$$

wobei die Relationen

$$(36) \quad L_f = \sum_{i=1}^{k_f} e_i^{(f)}, \quad \sum_f fL_f = n$$

bestehen, indem man über alle vorkommende Gradzahlen f in der Primidealzerlegung von p summiert.

Zu jeder Reihe (34) von Ordnungszahlen kann man in unendlich vielen Weisen eine neue Reihe von ganzen rationalen Zahlen

$$t_1^{(f)}, t_2^{(f)}, \dots, t_{k_f}^{(f)}$$

so zuordnen, daß allgemein $e_i^{(f)}$ zu $t_i^{(f)}$ relativ prim ist, und weiter

$$(37) \quad \frac{t_1^{(f)}}{e_1^{(f)}} < \frac{t_2^{(f)}}{e_2^{(f)}} < \dots < \frac{t_{k_f}^{(f)}}{e_{k_f}^{(f)}}.$$

Aus einem Hilfssatze in I, Kap. 3, § 3 (Fußnote) folgt dann, daß man eine solche Zahl ω bestimmen kann, daß allgemein die Zahl $\varphi_f(\omega)$ genau durch $p_i^{(f)t_i^{(f)}}$ ($i = 1, 2, \dots, k_f$) teilbar ist, wobei $\varphi_f(x)$ für jedes f eine beliebig gewählte Primfunktion (mod p) f -ten Grades ist. Nach einer früheren Bemerkung kann man ω als primitiv voraussetzen und man zeigt dann leicht, daß ω einer regulären Gleichung $F(x) = 0$ n -ten Grades genügt.

Aus der Idealzerlegung (35) von p folgt nämlich nach dem Hauptsatze eine entsprechende Zerlegung von $F(x)$

$$F(x) \equiv \prod_f F_f(x) \pmod{p^a},$$

wobei ein Faktor $F_f(x)$ vom Grade fL_f ist. Weiter folgt leicht aus den Eigenschaften von ω , daß eine Kongruenz

$$(38) \quad F_f(x) \equiv \varphi_f(x)^{L_f} \pmod{p}$$

bestehen muß. Um die Regularität von $F(x)$ nachzuweisen, muß man nun allgemein das Hauptpolygon $(p, \varphi_f(x))$ von $F(x)$ oder, was dasselbe ist, von $F_f(x)$ untersuchen; aus (38) folgt schon nach (9), Kap. 1, daß die Projektion dieses Polygons auf die X -Achse die Länge L_f hat. Nach Satz 1 folgt aus (37), daß das Polygon $(p, \varphi_f(x))$ mindestens k_f verschiedene Seiten hat, und aus Satz 5 leitet man dann sofort ab, daß das Polygon genau k_f Seiten mit den Neigungen (37) hat. Weiter ist nach (36) schon die Summe der Zahlen $e_i^{(f)}$ für das Polygon gleich L_f , und da $e_i^{(f)}$ nach (9) immer ein Multiplum der entsprechenden Projektion l_i ist, so wird allgemein für die i -te Seite

$$l_i = \lambda_i = e_i^{(f)}, \quad h_i = \kappa_i = t_i^{(f)}$$

und $\varepsilon_i = 1$, indem $e_i^{(f)}$ zu $t_i^{(f)}$ relativ prim ist. Das zugeordnete Polynom $F_i(x, y)$ wird daher nach (16), Kap. 1 vom ersten Grade, und kann daher keine mehrfache Faktoren (mod $p, \varphi_f(x)$) enthalten. Der Satz 7 ist also bewiesen. Aus dem Beweise geht auch hervor, daß man immer solche spezielle reguläre Gleichungen bestimmen kann, daß für jede Seite $\varepsilon_i = 1$ ist und daher das zugeordnete Polynom dieser Seite vom ersten Grade wird.

Schon früher habe ich die Eigenschaften der regulären Gleichung eingehend studiert⁸⁾ und gezeigt, wie man charakteristische Zahlen, Resultantenexponenten und auch Körperdifferente und Diskriminante bestimmen kann. Wegen Anwendungen sollen hier kurz einige Resultate erwähnt werden. Der Partialführer $f_{p_j^{(i)}}$ des Ringes $R(\vartheta)$ ist gleich

$$p_j^{(i)m_j^{(i)\kappa_i \lambda_i - \kappa_i - \lambda_i + 1}},$$

und aus der Bestimmung der charakteristischen Zahlen folgt aus I, Kap. 3, § 1, daß der Führer f des Ringes genau durch

$$p_j^{(i)\lambda_i(h_1 + \dots + h_i) + \kappa_i(l_{i+1} + \dots + l_k) - \kappa_i - \lambda_i + 1}$$

teilbar ist.

Bildet man die Norm des Führers, so folgt nach I, Kap. 3 (13) ein neuer Ausdruck für die Potenz, wenn p im Index aufgeht. Man erhält dadurch den Satz⁹⁾:

Satz 8. *Der Index einer regulären Gleichung ist genau durch $p^{\sum m_k}$ teilbar, wobei die Summe über alle verschiedene Primfunktionenteiler*

⁸⁾ Ö. Ore, „Bestimmung der Diskriminanten algebraischer Körper“, Acta math. 45 (1925), S. 303–344. — Ö. Ore, „Bestimmung der Differente eines algebraischen Zahlkörpers“, Acta math. 46 (1926), S. 363–392.

⁹⁾ Die erste Abhandlung in ⁸⁾, Satz 8.

$\varphi(x) \pmod{p}$ von $f(x)$ in (27) zu nehmen ist, und wo K für jede Primfunktion durch

$$(39) \quad K = l_2 h_1 + l_3 (h_1 + h_2) + \dots + l_k (h_1 + \dots + h_{k-1}) \\ + \frac{1}{2} \sum_{i=1}^k (h_i l_i - h_i - l_i + \varepsilon_i)$$

bestimmt ist. K ist auch die Anzahl der Gitterpunkte, welche in dem Gebiete zwischen dem Polygon, der X -Achse und der Endordinate des Polygons liegen; Punkte auf der X -Achse und auf der Endordinate werden nicht mitgerechnet, wohl aber Gitterpunkte, welche auf dem Polygon selbst liegen.

Der Index der speziellen Zahl ω , welche im Beweise des Satzes 7 angewandt wurde, ist daher genau durch $p^{\sum f K_f}$ teilbar, wobei für jedes f

$$K_f = e_2 t_1 + e_3 (t_1 + t_2) + \dots + e_k (t_1 + \dots + t_{k-1}) + \frac{1}{2} \sum_{i=1}^k (e_i - 1) (t_i - 1),$$

wobei die Indizes (f) weggelassen sind. Die Zahlen t_i waren aber durch die Ordnungszahlen e_i bestimmt, und man kann daher für K_f eine obere Grenze angeben, welche nur von den Exponenten e_i abhängt.

Aus dieser Tatsache folgt eine interessante Bemerkung über gemeinsame außerwesentliche Diskriminantenteiler eines Körpers. Nach dem Kriterium von Dedekind (vgl. I, Kap. 3, § 1) kann man bei gegebener Form der Primidealzerlegung immer entscheiden, ob eine vorgelegte Primzahl p ein g. a. D -Teiler ist. Durch eine einfache Abschätzung folgt aus dem Dedekindschen Kriterium, daß in diesem Falle immer $p < n$ ist¹⁰⁾; und wenn umgekehrt $p < n$ ist, so folgt aus dem Existenzsatze I, Kap. 3, Satz 12 leicht, daß es entsprechende Körper n -ten Grades gibt, worin p ein g. a. D -Teiler ist¹¹⁾.

Dagegen hat man keine Formel für die Bestimmung des größten gemeinsamen Teilers aller Zahl- oder Gleichungsdiskriminanten des Körpers. Aus der eben gemachten Bemerkung über K_f folgt aber, daß man immer bei gegebener Form der Primidealzerlegung eine obere Grenze für die Potenz angeben kann, worin p in allen Diskriminanten aufgeht. Es wäre sogar denkbar, daß diese größte gemeinsame Potenz von p in den Indizen nur von der Form der Primidealzerlegung abhinge, ich halte aber diese Vermutung für unwahrscheinlich. Durch passende Zahlenbeispiele würde man wohl ziemlich leicht die Unrichtigkeit nachweisen können.

¹⁰⁾ E. v. Zyliński, „Zur Theorie der außerwesentlichen Diskriminantenteiler algebraischer Körper“, *Math. Ann.* 73 (1913), S. 273–274.

¹¹⁾ Man vergleiche auch die Arbeit von Herrn M. Bauer: „Über die außerwesentlichen Diskriminantenteiler einer Gattung“, *Math. Annalen* 64 (1907), S. 573–576.

§ 5.

Bestimmung der Primidealzerlegung im allgemeinen Falle.

Durch den Satz 5 hat man eine Idealzerlegung der Primzahl p erreicht, welche aber nur unter speziellen Voraussetzungen, wie bei den regulären Gleichungen, die vollständige Primidealzerlegung liefert. Es soll nun zuletzt gezeigt werden, wie man auch im allgemeinen Falle die Primidealzerlegung bestimmen kann. Es genügt dabei, wie ich später zeige, wenn man den Spezialfall behandelt, wo p die Potenz eines Primideals ist und also das entsprechende Polynom $F(x) \pmod{p^e}$ irreduzibel wird. Wie ich zeige, kann man durch sukzessive Bildung von neuen Gleichungen zuletzt ein $F(x)$ erhalten, das die Normalform

$$F(x) = \psi(x)^e + p M(x), \quad M(x) \not\equiv 0 \pmod{p, \psi(x)}$$

hat, wobei $\psi(x)$ eine Primfunktion $(\text{mod } p)$ f -ten Grades ist.

Im folgenden werden nun die Bezeichnungen des § 2 angewandt; durch Satz 3 sind die allgemeinen Eigenschaften des gegebenen $F(x)$ charakterisiert, und weiter weiß man, daß, wenn p die Primidealzerlegung (10) hat, e durch λ und f durch $m m_1$ teilbar ist. Weiter genügt nach (24)

die Zahl $\Theta = \frac{\varphi(\vartheta')^\lambda}{p^r}$ der Kongruenz

$$(40) \quad F_1(\vartheta', \Theta) \equiv 0 \pmod{p},$$

wobei $F_1(x, y)$ eine Primfunktion m_1 -ten Grades $(\text{mod } p, \varphi(x))$ ist.

Wie schon früher angewandt, ist das Polynom

$$\Phi(y) = y^{p^{m m_1}} - y$$

$(\text{mod } p, \varphi(x))$ kongruent dem Produkte aller Primfunktionen $F_1(x, y)$ $(\text{mod } p, \varphi(x))$, deren Grade Teiler von m_1 sind, und weiter ist $\Phi(y) \pmod{p}$ kongruent dem Produkt aller Primfunktionen $\varphi_1(y) \pmod{p}$, deren Grade Teiler von $m m_1$ sind. Nach (40) ist daher $\Phi(\Theta) \equiv 0 \pmod{p}$, und man beweist leicht, daß die Zahlen

$$A(\vartheta', \Theta) = A_1(\vartheta') \Theta^{m_1-1} + \dots + A_{m_1}(\vartheta'),$$

wobei $A_i(x)$ beliebige Polynome sind, die sämtlichen Wurzeln der Kongruenz $\Phi(y) \equiv 0 \pmod{p}$ ist. Man bestimmt daraus leicht eine solche Zahl $\vartheta_1 = A(\vartheta', \Theta)$, die einer Kongruenz $\varphi_1(x) \equiv 0 \pmod{p}$ genügt, wobei $\varphi_1(x)$ eine Primfunktion $(\text{mod } p)$ vom Grade $m m_1$ ist.

Bildet man nun die Gleichung $F_1(x) = 0$, welcher ϑ_1 genügt, so wird

$$F_1(x) \equiv \varphi_1(x)^{m_1} \pmod{p},$$

und das Polygon $(p, \varphi_1(x))$ von $F_1(x)$ ist wieder geradlinig. Man kann dann weiter eine neue Zahl ϑ_2 bestimmen, welche der Kongruenz

$\varphi_2(x) \equiv 0 \pmod{p}$ genügt, wobei $\varphi_2(x)$ eine Primfunktion \pmod{p} mit Grad größer als dem Grad von $\varphi_1(x)$ ist, usw. Da man dies nicht unbegrenzt fortsetzen kann, muß man zuletzt zu einer Gleichung $F(x) = 0$ kommen, wofür

$$F(x) \equiv \varphi(x)^a \pmod{p},$$

und das zugeordnete Polynom von $F(x)$ für sein geradliniges Polygon ist eine Potenz $F_1(x, y)^b$, wobei $F_1(x, y)$ vom ersten Grade in y ist. Man kann auch annehmen, daß hier $\varkappa = 1$ ist; denn wenn dies nicht der Fall ist, bestimmt man c und d so, daß

$$c\varkappa - d\lambda = 1,$$

und dann ist die Zahl

$$\vartheta_1 = \vartheta + \frac{\varphi(\vartheta)^c}{p^d}$$

sicher ganz, und wie man leicht sieht, ist für die entsprechende Gleichung $\varkappa = 1$.

Man kann daher annehmen, daß man eine Gleichung $F(x) = 0$ abgeleitet hat, welche die Form

$$(41) \quad F(x) = \varphi(x)^{b\lambda} + A_1(x)p\varphi(x)^{(b-1)\lambda} + \dots + A_b(x)p^b$$

hat, und wenn L das entsprechende geradlinige Polygon ist, wird

$$(42) \quad F(x) \equiv (\varphi(x)^\lambda + pA(x))^b \pmod{L}.$$

Es kommt also darauf an, eine solche weitere Reduktion zu finden, daß $b = 1$ wird. Es soll hier kurz eine Methode angegeben werden, welche immer zum Ziele führt; möglicherweise gibt es auch ein einfacheres Verfahren.

Nach den Voraussetzungen ist mit den Bezeichnungen von § 1 die Zahl $\varphi(\vartheta)$ genau durch p^e und p genau durch $p^e = p^{e_2}$ teilbar. Wie früher wird $\Theta(\vartheta) = \frac{\varphi(\vartheta)^\lambda}{p}$ gesetzt, und aus (42) folgt wie in § 2, daß die Kongruenz

$$(43) \quad \Delta(\vartheta) = \Theta(\vartheta) + A(\vartheta) \equiv 0 \pmod{p}$$

besteht; die Hauptaufgabe im folgenden ist, zu untersuchen, welche Potenz p^x von p in der Zahl $\Delta(\vartheta)$ aufgeht.

Wird nun (41) durch p^b dividiert, erhält man

$$\frac{1}{p^b} F(x) = \Theta(x)^b + A_1(x)\Theta(x)^{b-1} + \dots + A_b(x),$$

und daraus folgt nach (43) auch eine Darstellung in der Form

$$(44) \quad \frac{1}{p^b} F(x) = \Delta(x)^b + B_1(x)\Delta(x)^{b-1} + \dots + B_b(x),$$

wobei $B_i(x)$ Polynome von höchstens $(\lambda m - 1)$ -tem Grade sind. Wenn hier $x = \vartheta$ gesetzt wird, werden die Glieder $B_i(\vartheta) \Delta(\vartheta)^{b-i}$ durch verschiedene leicht bestimmbare Potenzen von p teilbar. Schreibt man $B_i(x)$ in der Form

$$B_i(x) \equiv p^{\beta_i} \varphi(x)^{\gamma_i} C_i(x) \pmod{p^{\beta_i+1}},$$

wobei $C_i(x) \not\equiv 0 \pmod{p, \varphi(x)}$, so ist $B_i(\vartheta)$ genau durch p^{α_i} teilbar, wenn $\alpha_i = \lambda \beta_i + \gamma_i$ gesetzt wird; das Glied $B_i(\vartheta) \Delta(\vartheta)^{b-i}$ ist also genau durch $p^{e\alpha_i + (b-i)T}$ teilbar.

Man kann nun in (44) die entsprechenden Zahlen (α_i, i) als Gitterpunkte abbilden und erhält daraus ein neues Newtonsches Polygon für $\frac{1}{p^b} F(x)$, das man „das Polygon zweiter Ordnung“ nennen kann. Für das neue Polygon gelten ähnliche Sätze wie für die Polygone $(p, \varphi(x))$, und man beweist analog, daß das Polygon von $\frac{1}{p^b} F(x)$ geradlinig ist. Man sucht entsprechend diejenigen Glieder auf, welche durch Punkte auf dem Polygon abgebildet sind, und die Summe dieser Glieder hat die Form

$$(45) \quad \Delta(x)^{EL} + D_1(x) \Delta(x)^{(E-1)L} + \dots + D_E(x),$$

wobei EL und EH die Projektionen der Seite auf die X -Achse und Y -Achse sind, und wo L zu H relativ prim ist. In (45) hat $D_i(x)$ allgemein die Form

$$(46) \quad D_i(x) \equiv p^{\beta_i} \varphi(x)^{\gamma_i} G_i(x) \pmod{p^{\beta_i+1}},$$

wobei $\alpha_i = \lambda \beta_i + \gamma_i = iH$ ist.

Wird nun in (44) $x = \vartheta$ gesetzt, so folgt leicht wie in § 1

$$(47) \quad LT = H\varrho$$

und alle Glieder auf der Seite werden durch p^{eEH} teilbar; dagegen ist die Summe (45) sicher durch eine höhere Potenz teilbar und man erhält

$$(48) \quad \Delta(\vartheta)^{EL} + D_1(\vartheta) \Delta(\vartheta)^{(E-1)L} + \dots + D_E(\vartheta) \equiv 0 \pmod{p^{eEH+1}}.$$

Man kann nun ein Analogon zum zugeordneten Polynom der Seite einführen, wobei die Primfunktion $\varphi(x)$ eine ähnliche Rolle wie früher die Primzahl p spielt. Nach (43) bestimmt man leicht ein Polynom $B(x)$, so daß

$$p \equiv B(\vartheta) \varphi(\vartheta)^{\lambda} \pmod{p^{e\lambda+1}},$$

und wird dies in den Ausdruck (46) für $D_i(\vartheta)$ eingeführt, so kann man die Kongruenz (48) in der Form

$$(49) \quad \Delta(\vartheta)^{EL} + P_1(\vartheta) \varphi(\vartheta)^H \Delta(\vartheta)^{(E-1)L} + \dots + P_E(\vartheta) \varphi(\vartheta)^{EH} \equiv 0 \pmod{p^{eEH+1}}$$

schreiben, wobei $P_i(x)$ einfach bestimmbare Polynome sind. Dann ist

$$(50) \quad F(x, y) = y^E + P_1(x) y^{E-1} + \dots + P_E(x)$$

das zugeordnete Polynom der Seite, und wie früher beweist man, daß eine Zerlegung

$$(51) \quad F(x, y) \equiv F_1(x, y)^{a_1} \pmod{p, \varphi(x)}$$

besteht, wobei $F_1(x, y) \pmod{p, \varphi(x)}$ eine Primfunktion ist.

Es sei nun $N = N(\varphi(\vartheta))$ die Norm von $\varphi(\vartheta)$; man kann dann $N = p^{ef} M$ schreiben, wobei M eine rationale, nicht durch p teilbare Zahl ist. Wird nun (49) durch $\varphi(\vartheta)^{eH}$ dividiert, so folgt leicht nach (50) und (51), daß die ganze Zahl

$$\Theta_1 = M^H \frac{\Delta(\vartheta)^L}{\varphi(\vartheta)^H}$$

einer Kongruenz

$$(52) \quad F_2(\vartheta, \Theta_1) \equiv 0 \pmod{p}$$

genügt, wobei die Primfunktion $F_2(x, y)$ aus $F_1(x, y)$ dadurch hervorgeht, daß man y durch $\frac{y}{M^H}$ ersetzt und mit M^{eH} multipliziert. Aus (52) kann man dann wie früher eine neue Zahl ϑ_1 ableiten, die einer Kongruenz $\varphi_1(\vartheta_1) \equiv 0 \pmod{p}$ genügt, wobei der Grad der Primfunktion $\varphi_1(x) \pmod{p}$ größer als der Grad von $\varphi(x)$ ist.

Nimmt man an, daß die Reduktion so weit geführt ist, daß eine Gradvergrößerung von $\varphi(x)$ nicht mehr möglich ist, kann man voraussetzen, daß der Grad von $F_1(x, y)$ in (51) gleich eins ist. Aus (50) und (51) folgt dann leicht, daß eine Kongruenz

$$(53) \quad \Delta(\vartheta)^L + P(\vartheta) \varphi(\vartheta)^H \equiv 0 \pmod{p^{eH+1}},$$

besteht, und hier kann man weiter voraussetzen, daß $L = 1$ ist. Denn da in (47) L zu H relativ prim ist, kann man

$$\varrho = \varrho_1 L, \quad T = \varrho_1 H, \quad \varrho_1 < \varrho$$

setzen, und wenn $aL - bH = 1$ ist, so wird die Zahl

$$M^b \frac{\Delta(\vartheta)^a}{\varphi(\vartheta)^b}$$

ganz und durch eine kleinere Potenz als p^e teilbar. Wie früher leitet man daraus eine Zahl ϑ_1 ab, wofür $\varphi(\vartheta_1)$ bei derselben $\varphi(x)$ durch eine kleinere Potenz von p als $\varphi(\vartheta)$ teilbar ist. Diese Reduktion kann man aber nicht unbegrenzt vornehmen, und man kann daher zuletzt $L = 1$ voraussetzen. Dann ist nach (47) $T = H\varrho$, und wenn man wieder die Potenzen von $\varphi(\vartheta)^2$ reduziert, erhält man nach (53) eine Kongruenz

$$\Delta(\vartheta) + G(\vartheta) \equiv 0 \pmod{p^{T+1}}$$

oder nach (43) folgt daraus eine Kongruenz von der Form

$$\Delta_1(\vartheta) = \Theta(\vartheta) + A_1(\vartheta) \equiv 0 \pmod{p^{T+1}},$$

wobei $A_1(x)$ sowohl als $G(x)$ Polynome von höchstens $(m\lambda - 1)$ -tem Grade sind. $A_1(\vartheta)$ ist also durch p^{T_1} teilbar, wobei $T_1 > T$ ist.

Man entwickelt nun in (44) nicht nach Potenzen von $\Delta(x)$, sondern nach Potenzen von $A_1(x)$ und erhält wie früher

$$A_2(\vartheta) = \Theta(\vartheta) + A_2(\vartheta) \equiv 0 \pmod{p^{T_1+1}},$$

und wenn man so fortsetzt, kommt man zu Kongruenzen

$$\Theta(\vartheta) + A_r(\vartheta) \equiv 0 \pmod{p^{T_{r-1}+1}},$$

oder durch Multiplikation mit p

$$\varphi(\vartheta)^i + p A_r(\vartheta) \equiv 0 \pmod{p^{\alpha_r}},$$

wobei α_r mit r beliebig groß wird. Dies ist aber nur möglich, wenn in (42) $b = 1$ ist, denn die Kongruenz $F(x) \equiv 0 \pmod{p^\alpha}$ ist für hinreichend große α die Kongruenz kleinsten Grades, welcher ϑ genügt. $F(x)$ ist daher in die gewünschte Normalform gebracht.

Es ist nun leicht im allgemeinen und für ein beliebiges Primideal dieselbe Reduktion vorzunehmen. Es sei nämlich

$$f(x) \equiv f_1(x) \dots f_r(x) \pmod{p^\alpha}$$

die Primfunktionzerlegung der gegebenen definierenden Gleichung. Im Abbildungskörper $K_i(\vartheta')$ des Primideals \mathfrak{p}_i , der durch die Gleichung $f_i(\vartheta') = 0$ definiert ist (I, Kap. 3, § 2), kann man also, wie eben bewiesen, eine solche Zahl $\omega(\vartheta')$ angeben, daß das entsprechende Polynom in $K_i(\vartheta')$ die Normalform

$$(54) \quad \Phi(x) = \varphi(x)^{\omega} + p M(x), \quad M(x) \not\equiv 0 \pmod{p, \varphi(x)}$$

hat, wobei $\varphi(x)$ eine Primfunktion f_i -ten Grades ist. Bildet man die entsprechende Zahl $\omega(\vartheta)$ in $K(\vartheta)$, so wird auch der Faktor $f_i(x)$ die Form (54) erhalten, vorausgesetzt, daß $\omega(\vartheta)$ ganz ist. Dies braucht aber aber allgemein nicht der Fall zu sein, aber man leitet leicht eine ganze Zahl in $K(\vartheta)$ ab, welche dieselbe Eigenschaft hat. Wird nämlich

$$\Pi_i(x) = f_1(x) \dots f_{i-1}(x) f_{i+1}(x) \dots f_r(x)$$

gesetzt, so kann man zwei solche Polynome $A(x)$ und $B(x)$ angeben, daß

$$(55) \quad A(x) \Pi_i(x) + B(x) f_i(x) \equiv p^e \pmod{p^\alpha},$$

wobei mit den Bezeichnungen von I

$$e = e_{1,i} + \dots + e_{i-1,i} + e_{i+1,i} + \dots + e_{i,r} \leq \frac{\delta}{2}$$

ist. Wird (55) durch p^e dividiert und $x = \vartheta$ gesetzt, erhält man

$$\frac{\Pi_i(\vartheta)}{p^e} A(\vartheta) + \frac{f_i(\vartheta)}{p^e} B(\vartheta) \equiv 1 \pmod{p^{\alpha-e}}.$$

Es folgt leicht, daß die Zahl

$$\Theta = \frac{1}{p^e} \Pi_i(\vartheta) A(\vartheta)$$

ganz ist und daß die Kongruenzen

$$(56) \quad \Theta \equiv 0 \pmod{p_j^{e_j(a-e)}}, \quad j \neq i, \quad \Theta \equiv 1 \pmod{p_i^{e_i(a-e)}}$$

bestehen. Die Zahl $\vartheta_1 = \omega(\vartheta)\Theta$ ist daher sicher ganz, und wegen (56) überzeugt man sich leicht, daß in der entsprechenden Gleichung der Faktor $f_i(x)$ die Normalform (54) hat. Unsere Aufgabe ist daher vollständig gelöst.

Aus (56) folgt weiter, daß die Zahl $\varphi(\vartheta_1)$ nur durch das Primideal p_i und durch keinen anderen Primidealteiler von p teilbar sein kann. Man hat daher für p_i die Normaldarstellung $p_i = (p, \varphi(\vartheta_1))$. Aus diesen Resultaten folgt auch, daß man in endlich vielen Schritten Fundamentalsysteme für alle Primideale und Primzahlen aufstellen kann, weiter kann man nun nach I, Kap. 3, § 3 die Körperdifferente und Körperdiskriminante vollständig bestimmen.

Zuletzt soll als eine letzte Anwendung der Newtonschen Polygone gezeigt werden, wie man sehr einfach die folgende, oft vorkommende Aufgabe löst: *Man soll die Primidealzerlegung einer gegebenen Zahl ω des Körpers bestimmen.*

Man bildet zunächst die Norm $N = N(\omega)$ und bestimmt dadurch, welche Primzahlen p überhaupt mit ω gemeinsame Primidealfaktoren haben können. Weiter bildet man die Gleichung $F(x) = 0$, welcher ω genügt, und kann dabei voraussetzen, daß ω primitiv ist. Denn sonst bildet man die Zahl

$$\omega_1 = \omega + r N^2 \vartheta,$$

wobei die ganze rationale Zahl r so gewählt wird, daß ω_1 primitiv ist; ω_1 enthält dann genau dieselbe Potenz wie ω von allen Primidealen, welche in ω aufgehen.

Wenn dann p eine beliebige Primzahl ist, die in N aufgeht, bildet man das Polygon (p, x) von $F(x)$. Gehört nun ein Primideal \mathfrak{p} von p zur i -ten Seite dieses Polygons, so bestehen nach Satz 1 die Relationen

$$e = \rho \lambda_i, \quad t = \rho \kappa_i,$$

wenn p durch p^e und ω durch p^t genau teilbar sind; t ist daher einfach durch

$$t = \frac{\kappa_i}{\lambda_i} e$$

bestimmt.

Zum Hilbertschen Aufbau der reellen Zahlen.

Von

Wilhelm Ackermann in Göttingen.

Um den Beweis für die von Cantor aufgestellte Vermutung zu erbringen, daß sich die Menge der reellen Zahlen, d. h. der zahlentheoretischen Funktionen, mit Hilfe der Zahlen der zweiten Zahlklasse auszählen läßt, benutzt Hilbert einen speziellen Aufbau der zahlentheoretischen Funktionen. Wesentlich bei diesem Aufbau ist der Begriff des Typs einer Funktion. Eine Funktion vom Typ 1 ist eine solche, deren Argumente und Werte ganze Zahlen sind, also eine gewöhnliche zahlentheoretische Funktion. Die Funktionen vom Typ 2 sind die Funktionenfunktionen. Eine derartige Funktion ordnet jeder zahlentheoretischen Funktion eine Zahl zu. Eine Funktion vom Typ 3 ordnet wieder den Funktionenfunktionen Zahlen zu, usw. Die Definition der Typen läßt sich auch ins Transfinite fortsetzen, für den Gegenstand dieser Arbeit ist das aber nicht von Belang¹⁾.

Sämtliche Funktionen werden durch Rekursion nach einer gewöhnlichen Zahlenvariablen definiert. Z. B. ist

$$\varphi(a, 0) = 0,$$

$$\varphi(a, n + 1) = \varphi(a, n) + a$$

ein Beispiel einer Definition einer Funktion vom Typ 1. $\varphi(a, b)$ stellt offenbar die Funktion $a \cdot b$ dar. Die folgende Rekursion definiert eine Funktion vom Typ 2.

$$\varrho_c(f(c), a, 0) = a,$$

$$\varrho_c(f(c), a, n + 1) = f(\varrho_c(f(c), a, n)).$$

$\varrho_c(f(c), a, n)$ stellt die n -malige Iteration der Funktion f , genommen für das Argument a , dar. (Der Index c von ϱ soll bedeuten, daß $\varrho_c(f(c), a, n)$ von f abhängt und nicht von c .) Neben der Rekursion ist natürlich auch

¹⁾ Siehe des näheren D. Hilbert, Über das Unendliche, Math. Annalen 95 (1926).

die Bildung von Funktionen durch Einsetzung zulässig. Sind z. B. $\varphi(a, b)$ und $\psi(a, b)$ zahlentheoretische Funktionen von zwei Variablen, so ist auch $\varphi(\varphi(a, a), \psi(a, b))$ eine derartige Funktion.

Der Aufbau der zahlentheoretischen Funktionen geschieht nun in der Weise, daß man sie in gewisse Gruppen ordnet. In die erste Gruppe kommen diejenigen Funktionen, bei deren Definition nur der Typ 1 gebraucht wird, in die zweite die Funktionen, bei deren Definition die Benutzung einer durch Rekursion definierten Funktion vom Typ 2 wesentlich ist, während sich ein höherer Typ umgehen läßt, usf..

Damit der geschilderte Aufbau wirklich alle zahlentheoretischen Funktionen liefert, ist es wesentlich, daß jede Gruppe auch eine Erweiterung des Funktionenbestandes bringt. Es wäre ja z. B. denkbar, daß man alle Funktionen, bei deren Definition man den ersten und zweiten Typ braucht, auch mit Hilfe des ersten Typs allein erhalten kann.

Wir stellen nun im folgenden eine Funktion auf, von der wir beweisen, daß sie nicht ohne den zweiten oder einen höheren Typ definiert werden kann. Mit Hilfe der dabei benutzten Methode wird man auch für andere Gruppen unschwer den Beweis führen können, daß sie den Funktionenbestand erweitern²⁾.

Es sollen übrigens die Rekursionen sämtlich nur nach einer Variablen gehen. Simultane Rekursionen sind ausgeschlossen. Eine Funktion, die durch simultane Rekursion definiert ist, läßt sich auch durch gewöhnliche Rekursion darstellen, falls man eventuell höhere Typen benutzt.

1.

Wir wollen die erwähnte Funktion zunächst angeben. Bei dem Aufbau der Funktionen vermittelt Rekursion muß die Funktion $a + 1$ als bekannt vorausgesetzt werden und kann nicht wieder durch Rekursion definiert werden. Die Funktion $a + b$ wird dann durch die Rekursion

$$a + 0 = a,$$

$$a + (b + 1) = (a + b) + 1$$

definiert. Der Einfachheit halber wollen wir noch zwei weitere Funktionen als bekannte Ausgangsfunktionen nehmen, nämlich $\lambda(a, b)$ und $\iota(a, b)$. $\lambda(a, b)$ ist gleich 1, falls $a = b$, und gleich 0, falls $a \neq b$. Bei $\iota(a, b)$ ist es umgekehrt. Man kann aber auch λ und ι durch Rekursion definieren.

²⁾ Eine Arbeit, die mit der vorliegenden manche Berührungspunkte hat, wird von Herrn G. Sudan publiziert werden. Es handelt sich bei ihr um die Definition von Zahlen der zweiten Zahlklasse, die man in ähnlicher Weise klassifizieren kann wie die Definitionen der reellen Zahlen.

$\alpha(a, n)$ soll für $\iota(n, 1) \cdot \iota(n, 0) \cdot a + \lambda(n, 1)$ geschrieben werden. Hierbei ist $a \cdot b$ definiert durch

$$\begin{aligned} a \cdot 0 &= a, \\ a \cdot (b + 1) &= a \cdot b + a. \end{aligned}$$

$\alpha(a, n)$ ist also immer gleich a , ausgenommen für $n = 0$ und $n = 1$. Es ist $\alpha(a, 0) = 0$ und $\alpha(a, 1) = 1$. ϱ bedeute die schon vorher erwähnte Iterationsfunktion. Endlich wird eine zahlentheoretische Funktion φ mit drei Variablen gegeben durch:

$$\begin{aligned} \varphi(a, b, 0) &= a + b, \\ \varphi(a, b, n + 1) &= \varrho_c(\varphi(a, c, n), \alpha(a, n), b). \end{aligned}$$

Wie man daraus erkennt, ist $\varphi(a, b, 1)$ mit $a \cdot b$ identisch. $\varphi(a, b, 2)$ stimmt mit a^b überein. $\varphi(a, b, 3)$ ist die b -malige Iteration von a^b , genommen für a , usw. Unsere gesuchte Funktion erhalten wir nun, wenn wir in $\varphi(a, b, c)$ alle drei Argumente gleichnehmen. Wir behaupten also, $\varphi(a, a, a)$ kann nicht ohne Benutzung des zweiten Typs definiert werden³⁾.

Die Ausschließung von simultanen Rekursionen ist für unsere Behauptung wesentlich. Es gelten nämlich die folgenden Formeln:

$$\begin{aligned} \varphi(a, b, 0) &= a + b, \\ \varphi(a, 0, n + 1) &= \alpha(a, n), \\ \varphi(a, b + 1, n + 1) &= \varphi(a, \varphi(a, b, n + 1), n). \end{aligned}$$

(Diese Formeln werden übrigens gleich bewiesen.)

2.

Der Beweis für unsere Behauptung wird darin bestehen, daß wir zeigen, daß die Funktion $\varphi(a, a, a)$ schneller wächst als jede Funktion, bei deren Definition man nur den ersten Typ benutzt. Um das zeigen zu können, müssen wir eine Reihe von Eigenschaften und Formeln für unsere Funktion beweisen.

I. $\varphi(a, 0, n + 1) = \alpha(a, n)$.

Beweis.

$$\varphi(a, 0, n + 1) = \varrho_c(\varphi(a, c, n), \alpha(a, n), 0) = \alpha(a, n).$$

II. $\varphi(a, b + 1, n + 1) = \varphi(a, \varphi(a, b, n + 1), n)$.

Beweis.

$$\varphi(a, b + 1, n + 1) = \varrho_c(\varphi(a, c, n), \alpha(a, n), b + 1),$$

³⁾ Diese von mir aufgestellte Funktion ist schon in der Hilbertschen Abhandlung „Über das Unendliche“ erwähnt; vgl. S. 185 dort.

$$\varrho_c(\varphi(a, c, n), \alpha(a, n), b+1) = \varphi(a, \varrho_c(\varphi(a, c, n), \alpha(a, n), b), n),$$

$$\varrho_c(\varphi(a, c, n), \alpha(a, n), b) = \varphi(a, b, n+1),$$

woraus die Behauptung ersichtlich ist.

Wir brauchen ferner gewisse Monotonitätseigenschaften von φ .

III. Falls $a \geq 2$, ist $\varphi(a, b+1, n) > \varphi(a, b, n)$.

Beweis. Für $n = 0, 1$ ist diese Behauptung richtig; denn

$$a + 2 + b + 1 > a + 2 + b$$

und

$$(a+2) \cdot (b+1) > (a+2) \cdot b.$$

Es bleibt zu beweisen:

$$A) \quad \varphi(a+2, b+1, n+2) > \varphi(a+2, b, n+2).$$

Wir beweisen diese Formel zusammen mit

$$B) \quad \varphi(a+2, b, n+2) > b$$

durch Induktion nach n .

$$\varphi(a+2, b+1, 2) > \varphi(a+2, b, 2),$$

denn

$$(a+2)^{b+1} > (a+2)^b,$$

$$\varphi(a+2, b, 2) = (a+2)^b > b.$$

Für n seien beide Formeln schon bewiesen. Wir beweisen zunächst B) für $n+1$ durch Induktion nach b . Für $b=0$ und $n+1$ ist B) richtig, denn es gilt für jedes n :

$$\varphi(a+2, 0, n+3) = a+2 > 0,$$

$$\varphi(a+2, b+1, n+3) = \varphi(a+2, \varphi(a+2, b, n+3), n+2).$$

Wenn also $\varphi(a+2, b, n+3) \geq b+1$, so ist wegen der Gültigkeit von A) für n :

$$\varphi(a+2, \varphi(a+2, b, n+3), n+2) \geq \varphi(a+2, b+1, n+2),$$

also

$$\varphi(a+2, b+1, n+3) \geq \varphi(a+2, b+1, n+2),$$

$$\varphi(a+2, b+1, n+2) > b+1,$$

$$\varphi(a+2, b+1, n+3) > b+1.$$

B) ist damit für $n+1$ bewiesen.

$$\varphi(a+2, b+1, n+3) = \varphi(a+2, \varphi(a+2, b, n+3), n+2)$$

$$> \varphi(a+2, b, n+3) \quad (\text{nach B}).$$

Damit sind die Formeln A) und B), also auch III bewiesen.

In ähnlicher Weise beweisen sich die beiden Formeln

$$\text{IV. } \varphi(a+1, b, n) \geq \varphi(a, b, n),$$

$$\text{V. } \varphi(a+2, b+3, n+1) > \varphi(a+2, b+3, n).$$

Beweis von IV. Man überzeugt sich leicht, daß die Formel IV für $a=0$ und $a=1$ gilt. Für $n=0, 1, 2$ läßt sich das leicht nachprüfen. Für $n \geq 3$ gelten die Formeln:

$$\varphi(0, b, n) \leq 1,$$

$$\varphi(1, b, n) = 1,$$

$$\varphi(2, b, n) \geq 1,$$

die man durch Induktion nach n und Unterscheidung der Fälle $b=0$ und $b \neq 0$ beweist. Wir müssen, um IV zu beweisen, noch zeigen, daß

$$\varphi(a+3, b, n) \geq \varphi(a+2, b, n).$$

Für $n=0$ stimmt das, desgl. für $b=0$. Für kleineres n sowie für dasselbe n und kleineres b sei es schon gezeigt.

$$\varphi(a+3, b+1, n+1) = \varphi(a+3, \varphi(a+3, b, n+1), n),$$

$$\varphi(a+3, \varphi(a+3, b, n+1), n) \geq \varphi(a+2, \varphi(a+3, b, n+1), n),$$

$$\varphi(a+3, b, n+1) \geq \varphi(a+2, b, n+1).$$

Aus III folgt dann:

$$\varphi(a+2, \varphi(a+3, b, n+1), n) \geq \varphi(a+2, \varphi(a+2, b, n+1), n).$$

$\varphi(a+2, \varphi(a+2, b, n+1), n)$ ist aber dasselbe wie $\varphi(a+2, b+1, n+1)$.

Beweis von V. Für $n=0$ und $n=1$ stimmt die Formel. Es bleibt zu zeigen:

$$\varphi(a+2, b+3, n+3) > \varphi(a+2, b+3, n+2).$$

Nehmen wir statt dessen die allgemeinere:

$$\varphi(a+2, b+1, n+3) > \varphi(a+2, b+1, n+2).$$

Für $b=0$ ist das richtig, denn

$$\begin{aligned} \varphi(a+2, 1, n+3) &= \varphi(a+2, \varphi(a+2, 0, n+3), n+2) \\ &= \varphi(a+2, a+2, n+2), \end{aligned}$$

$$\varphi(a+2, a+2, n+2) > \varphi(a+2, 1, n+2) \quad [\text{nach III}].$$

Für $n=0$ ebenfalls. Es sei nun für kleineres n sowie für dasselbe n und kleineres b schon gezeigt.

$$\varphi(a+2, b+2, n+4) = \varphi(a+2, \varphi(a+2, b+1, n+4), n+3),$$

$$\varphi(a+2, b+1, n+4) > \varphi(a+2, b+1, n+3),$$

$$\varphi(a+2, b+1, n+3) \geq b+2.$$

(Nach Formel B), S. 121.)

$$\begin{aligned} \varphi(a+2, \varphi(a+2, b+1, n+4), n+3) &> \varphi(a+2, b+2, n+3), \\ \varphi(a+2, b+2, n+4) &> \varphi(a+2, b+2, n+3). \end{aligned}$$

Damit ist auch V bewiesen.

Die folgende Eigenschaft ist die wichtigste, die bei unserem Beweis gebraucht wird.

VI. Von den beiden Formeln

$$\begin{aligned} \varrho_c(\varphi(c, c, n), a, a) &\leq \varphi(a, a, n+3), \\ \varrho_c(\varphi(c, c, n), a, a) &\leq 2 \end{aligned}$$

ist mindestens eine richtig.

[Die erste der beiden Formeln gilt übrigens immer, ausgenommen für $a=1$ und $n=0$.]

Wir brauchen hier eine Reihe von Hilfsformeln.

Formel 1: Wir beweisen zunächst:

$$\varphi(\varphi(a+2, c, 3), \varphi(a+2, c, 3), 1) \leq \varphi(a+2, c+1, 3)$$

durch Induktion nach c . Für $c=0$ ist diese Formel richtig, denn die linke Seite wird gleich $\varphi(a+2, a+2, 1)$, d. h. $(a+2) \cdot (a+2)$, und die rechte $(a+2)^{(a+2)}$. Benutzt man, daß für $c=1, 2, 3$ die Funktion $\varphi(a, b, c)$ die bekannten Funktionen darstellt, so ergeben sich leicht die folgenden Formeln:

$$2 \leq \varphi(a+2, c, 3),$$

$$\begin{aligned} \varphi(\varphi(a+2, c+1, 3), \varphi(a+2, c+1, 3), 1) \\ &= \varphi(a+2, c+1, 3)^2 = (a+2)^{2 \cdot \varphi(a+2, c, 3)} \\ &\leq (a+2)^{\varphi(a+2, c, 3) \varphi(a+2, c, 3)} \\ &\leq (a+2)^{\varphi(a+2, c+1, 3)}, \\ (a+2)^{\varphi(a+2, c+1, 3)} &= \varphi(a+2, c+2, 3). \end{aligned}$$

Formel 2:

$$\varphi(\varphi(a+2, b, 3), \varphi(a+2, c, 3), 2) \leq \varphi(a+2, c+2, 3),$$

falls $c \geq b$.

Beweis. a) Es sei $b=0$.

$$\varphi(a+2, 0, 3) = a+2,$$

$$\begin{aligned} \varphi(a+2, \varphi(a+2, c, 3), 2) \\ &= (a+2)^{\varphi(a+2, c, 3)} = \varphi(a+2, c+1, 3) < \varphi(a+2, c+2, 3) \\ &\quad \text{(nach III)} \end{aligned}$$

b) $b \geq 1$.

$$\begin{aligned} & \varphi(\varphi(a+2, b, 3), \varphi(a+2, c, 3), 2) \\ &= (a+2)^{\varphi(a+2, b-1, 3) \cdot \varphi(a+2, c, 3)} \leq (a+2)^{\varphi(a+2, c, 3) \cdot \varphi(a+2, c, 3)} \\ & \leq (a+2)^{\varphi(a+2, c+1, 3)}, \\ & \quad (\text{nach Formel 1}) \\ & (a+2)^{\varphi(a+2, c+1, 3)} = \varphi(a+2, c+2, 3). \end{aligned}$$

Formel 3:

Für alle $a \geq 2$, $c \geq b$ und $b \neq 0$ und für ein festes n möge die Beziehung gelten:

$$\varphi(\varphi(a, b, n+3), \varphi(a, c, n+3), n+2) \leq \varphi(a, c+2, n+3).$$

Wir behaupten dann, es gilt die Beziehung:

$$\varphi(\varphi(a, b, n+4), m, n+3) \leq \varphi[a, \varphi(a, b-1, n+4) + 2m, n+3]$$

für alle derartigen a, b, c und für beliebiges m .

Für $m=0$ sind die rechten und linken Seiten gleich.

$$\begin{aligned} & \varphi(\varphi(a, b, n+4), m+1, n+3) \\ &= \varphi\{\varphi(a, b, n+4), \varphi[\varphi(a, b, n+4), m, n+3], n+2\}. \end{aligned}$$

Gilt die Behauptung schon für kleineres m , so folgt mit Hilfe von Formel III, daß die rechte Seite nicht größer ist als

$$\begin{aligned} & \varphi\{\varphi(a, b, n+4), \varphi[a, \varphi(a, b-1, n+4) + 2m, n+3], n+2\}, \\ & \varphi(a, b, n+4) = \varphi[a, \varphi(a, b-1, n+4), n+3]. \end{aligned}$$

Aus der Voraussetzung ergibt sich dann, daß der vorletzte Ausdruck kleiner oder gleich

$$\varphi[a, \varphi(a, b-1, n+4) + 2m+2, n+3]$$

ist. Damit ist die Behauptung bewiesen.

Formel 4:

Es sei $a \geq 2$ und $c \geq b$. Dann ist

$$\varphi(\varphi(a, b, n+3), \varphi(a, c, n+3), n+2) \leq \varphi(a, c+2, n+3).$$

Für $n=0$ stimmt diese Formel mit Formel 2 überein. Wir beweisen sie nun für $n+1$, unter der Voraussetzung, daß sie für n gilt. Zunächst sei b von 0 verschieden. Nach Formel 3 ist

$$\begin{aligned} & \varphi(\varphi(a, b, n+4), \varphi(a, c, n+4), n+3) \\ & \leq \varphi[a, \varphi(a, b-1, n+4) + 2\varphi(a, c, n+4), n+3] \\ & \leq \varphi[a, 3 \cdot \varphi(a, c, n+4), n+3]. \end{aligned}$$

Durch Induktion nach c läßt sich für $c \geq 2$ leicht beweisen

$$3c \leq \varphi(2, c, 3).$$

Da ferner $\varphi(a, c, n+4) \geq 2$, weil $a \geq 2$, so ist

$$\begin{aligned} 3 \cdot \varphi(a, c, n+4) &\leq \varphi(2, \varphi(a, c, n+4), 3) \leq \varphi(a, \varphi(a, c, n+4), n+3), \\ \varphi(a, \varphi(a, c, n+4), n+3) &= \varphi(a, c+1, n+4), \\ \varphi(\varphi(a, b, n+4), \varphi(a, c, n+4), n+3) &\leq \varphi(a, \varphi(a, c+1, n+4), n+3), \\ \varphi(a, \varphi(a, c+1, n+4), n+3) &= \varphi(a, c+2, n+4). \end{aligned}$$

Damit ist die Formel für $b \neq 0$ bewiesen.

$$\begin{aligned} \varphi(\varphi(a, 0, n+3), \varphi(a, c, n+3), n+2) &= \varphi(a, \varphi(a, c, n+3), n+2) \\ &= \varphi(a, c+1, n+3) \leq \varphi(a, c+2, n+3). \end{aligned}$$

Für $b = 0$ ist sie also ebenfalls richtig.

Formel 5:

$$\varrho_c(\varphi(c, c, n+2), a+2, b) \leq \varphi(a+2, 2b, n+3).$$

Beweis durch Induktion nach b . Für $b = 0$ bedeutet die Formel

$$a+2 \leq a+2$$

$$\begin{aligned} &\varrho_c(\varphi(c, c, n+2), a+2, b+1) \\ &= \varphi\{\varrho_c[\varphi(c, c, n+2), a+2, b], \varrho_c[\varphi(c, c, n+2), a+2, b], n+2\}. \end{aligned}$$

Der letzte Ausdruck ist nach Formel IV und III

$$\leq \varphi(\varphi(a+2, 2b, n+3), \varphi(a+2, 2b, n+3), n+2).$$

Nach Formel 4 ist der letzte Term wieder

$$\leq \varphi(a+2, 2b+2, n+3).$$

Damit ist Formel 5 bewiesen.

Beweis für VI. a) $a = 0$.

$$\varrho_c(\varphi(c, c, n), 0, 0) = 0 \leq \varphi(0, 0, 3).$$

b) $a = 1$.

$$\begin{aligned} \varrho_c(\varphi(c, c, n), 1, 1) &= \varphi(1, 1, n), \\ \varphi(1, 1, 0) &= 2, \\ \varphi(1, 1, n+1) &= 1. \end{aligned}$$

Das letztere beweist sich leicht durch Induktion nach n .

c) $a \geq 2$. Für $n = 0$ und $n = 1$ ist VI leicht zu verifizieren. Es bleibt für beliebige a und n zu beweisen:

$$\varrho_c(c, c, n+2), a+2, a+2 \leq \varphi(a+2, a+2, n+5).$$

Ersetzen wir in Formel 5 b durch $a + 2$, so erhalten wir:

$$\varrho_c(\varphi(c, c, n + 2), a + 2, a + 2) \leq \varphi(a + 2, 2 \cdot (a + 2), n + 3),$$

$$2(a + 2) \leq (a + 2)^{a+2} = \varphi(a + 2, 1, 3),$$

$$\varphi(a + 2, 1, 3) \leq \varphi(a + 2, a + 1, n + 4),$$

$$\varrho_c(\varphi(c, c, n + 2), a + 2, a + 2)$$

$$\leq \varphi(a + 2, \varphi(a + 2, a + 1, n + 4), n + 3),$$

$$\varphi(a + 2, \varphi(a + 2, a + 1, n + 4), n + 3) = \varphi(a + 2, a + 2, n + 4),$$

$$\varphi(a + 2, a + 2, n + 4) \leq \varphi(a + 2, a + 2, n + 5),$$

$$\varrho_c(\varphi(c, c, n + 2), a + 2, a + 2) \leq \varphi(a + 2, a + 2, n + 5),$$

w. z. b. w.

3.

Die angegebenen Formeln werden uns den Beweis ermöglichen, daß $\varphi(a, a, a)$ nicht ohne Benutzung des zweiten Typs definiert werden kann. Bezeichnen wir zur Abkürzung die Aussage, daß die zahlentheoretische Funktion $f(\cdot)$ von einem gewissen Argument an nicht kleiner ist als $g(\cdot)$, mit $\mathfrak{R}_a(g(a), f(a))$, so ist der Gang des Beweises der folgende:

Sei $\psi(a)$ irgendeine zahlentheoretische Funktion, die ohne Benutzung des zweiten Typs definiert ist. Wir werden dann zeigen, daß es ein n gibt, so daß $\mathfrak{R}_a(\psi(a), \varphi(a, a, n))$, oder, wie wir zur Abkürzung schreiben, wir zeigen, daß $\mathfrak{G}_{a_n}(\psi(a), \varphi(a, a, n))$ der Fall ist. Damit wäre der Beweis geführt, da $\mathfrak{G}_{a_n}(\varphi(a, a, a), \varphi(a, a, n))$ nicht zutrifft, wie sich aus Formel V ergibt.

Unsere Ausgangsfunktionen $\lambda, c, a + 1$ haben die Eigenschaft \mathfrak{G} . Aus diesen Ausgangsfunktionen gewinnt man sukzessiv neue Funktionen, indem man die schon gebildeten Funktionen wieder bei den Rekursionen verwendet. Wir werden durch Induktion beweisen, daß die neugebildeten Funktionen immer wieder die Eigenschaft \mathfrak{G} haben.

Genauer gesagt beweisen wir die folgenden Sätze: Wir mögen einen endlichen Bereich von Funktionen, von einer, von zwei und mehreren Zahlenvariablen haben mit den folgenden Eigenschaften:

1. Wenn $\psi(a_1, a_2, \dots, a_n)$ eine Funktion unseres Bereiches ist, so enthält der Bereich auch eine in bezug auf alle Variablen monoton steigende Funktion $\psi'(a_1, \dots, a_n)$ (Monotonie hier mit Einschluß der Gleichheit verstanden), so daß für beliebige a_1, a_2, \dots, a_n

$$\psi'(a_1, a_2, \dots, a_n) \geq \psi(a_1, a_2, \dots, a_n)$$

gilt. Es darf natürlich auch ψ' mit ψ identisch sein, wenn ψ schon selbst monoton ist.

2. Für jede Funktion $\psi(a_1, a_2, \dots, a_n)$ des Bereiches gilt

$$\mathfrak{G}_{a_n}(\psi(a, a, \dots, a), \varphi(a, a, n)).$$

Unsere beiden Sätze lauten dann folgendermaßen:

Satz 1. *Fügt man unserem Bereich eine neue Funktion hinzu, die man durch Einsetzung aus den Funktionen unseres Bereiches gewinnt, so hat der neue Bereich entweder wieder die beiden Eigenschaften, oder man kann ihn durch Hinzufügung noch weiterer Funktionen wieder in einen solchen Bereich verwandeln.*

Satz 2. *Bildet man unter Benützung der Funktionen unseres Bereichs durch Rekursion eine neue Funktion ψ , und fügt man diese Funktion und eventuell noch andere dem Bereich hinzu, so treffen auf den neu entstehenden Bereich wieder die beiden Eigenschaften zu.*

Da unsere Ausgangsfunktionen die Eigenschaften 1. und 2. haben, so ist mit dem Beweis dieser beiden Sätze alles erledigt.

Beweis von Satz 1. Man habe eine neue Funktion durch Einsetzung gebildet. Falls die Funktionen, aus denen man die neue Funktion kombinierte, alle monotone Funktionen waren, so ist auch die neue Funktion monoton in bezug auf sämtliche Variablen (monoton hier immer im Sinne von monoton steigend mit Einschluß der Gleichheit verstanden). Dann ist also die erste Eigenschaft erfüllt. — Sind bei der Zusammensetzung der neuen Funktion nicht lauter monotone Funktionen verwandt, so ersetzt man in dieser Kombination jede Funktion durch die zugehörige monotone, und fügt auch die so entstehende Funktion noch dem Bereich hinzu. Die erste Eigenschaft ist dann erfüllt. — Was die zweite Eigenschaft anbelangt, so genügt es zu zeigen, daß sie für alle Funktionen gilt, die man durch Einsetzung aus den monotonen Funktionen unseres Bereiches erhält. Ferner kann man sich darauf beschränken zu zeigen, daß, wenn $\psi(a_1, a_2, \dots, a_m)$ eine Funktion unseres Bereiches ist, die m Argumente hat, und wenn $\chi_1(a), \dots, \chi_m(a)$ m Funktionen mit den Eigenschaften

$$\mathfrak{G}_{a_n}(\chi_i(a), \varphi(a, a, n)) \mathfrak{G}_{a_n} \text{ sind, } (\psi(\chi_1(a), \dots, \chi_m(a)), \varphi(a, a, n))$$

der Fall ist. Der allgemeine Fall ergibt sich nämlich daraus durch Induktion nach dem Grade der Ineinanderschaltung der aus den Funktionen unseres Bereichs durch Einsetzung entstehenden Funktion. Da nun für $n \geq m$ $\mathfrak{H}_a(\varphi(a, a, m), \varphi(a, a, n))$ und da ψ monoton und für ein beliebiges i $\mathfrak{G}_{a_n}(\chi_i(a), \varphi(a, a, n))$ richtig ist, so hat man

$$\mathfrak{G}_{a_n}(\psi(\chi_1(a), \dots, \chi_m(a)), \psi(\varphi(a, a, n), \dots, \varphi(a, a, n))).$$

Da ferner

$$\mathfrak{G}_{a_n}(\psi(a, a, \dots, a), \varphi(a, a, n)),$$

so ist auch

$$\mathfrak{G}_{an}(\psi(\chi_1(a), \dots, \chi_m(a)), \varphi(\varphi(a, a, n), \varphi(a, a, n), n)), \\ \varphi(\varphi(a, a, n), \varphi(a, a, n), n) = \varrho_c(\varphi(c, c, n), a, 2).$$

Da

$$\mathfrak{K}_a(\varrho_c(\varphi(c, c, n), a, 2), \varrho_c(\varphi(c, c, n), a, 2))$$

und wegen Formel VI ergibt sich schließlich:

$$\mathfrak{G}_{an}(\psi(\chi_1(a), \dots, \chi_m(a)), \varphi(a, a, n)), \text{ q. e. d.}$$

Beweis von Satz 2. a) Der Beweis von Satz 2 ist bei weitem komplizierter als der von Satz 1 und wird den Rest der Arbeit ausmachen. Zunächst wird eine nähere Erörterung des Rekursionsschemas, das den zweiten Typ nicht benutzt, notwendig. Bei einer Funktion einer Variablen sieht die Rekursion folgendermaßen aus:

$$\psi(0) = a,$$

$$\psi(a+1) = b(a, \psi(a)),$$

a ist hier ein Funktional, das keine Variablen enthält. b eine eventuell durch Ineinanderschachtung mehrerer Rekursionsfunktionen entstehende Funktion zweier Variablen. In a und b darf keine Rekursionsfunktion vom zweiten Typ und nur bereits definierte Funktionen vorkommen. Analog lautet das Schema bei Funktionen zweier Variablen:

$$\psi(a, 0) = a(a)$$

$$\psi(a, b+1) = b_c(a, b, \psi(c, b)).$$

Hier gelten dieselben Bemerkungen. Es ist besonders zu beachten, daß die Funktionenfunktion $b_c(a, b, f(c))$ ohne eine Rekursionsfunktion vom zweiten Typ gebildet sein muß. Man kann natürlich auch ohne Rekursion, durch bloße Ineinanderschachtung Funktionenfunktionen bilden, z. B. $f(f(a))$, $f(a) + f(b)$, usw. Die folgende Rekursion ist aber verboten:

$$\psi(a, 0) = a(a),$$

$$\psi(a, b+1) = \varrho_c(\psi(c, b), b, a).$$

Dagegen ist diese erlaubt:

$$\psi(a, 0) = a + 1,$$

$$\psi(a, b+1) = 1 + \psi[\psi(a+b, b), b].$$

Entsprechend hat man für Funktionen von n Variablen das Schema:

$$\psi(a, b, \dots, m, 0) = a(a, b, \dots, m),$$

$$\psi(a, b, \dots, m, n+1) = b_{a, b, \dots, m_1}(a, b, \dots, m, n, \psi(a_1, \dots, m_1, n)).$$

Wir können hier annehmen, daß sich $a(\dots)$ und $b(\dots)$ aus lauter monotonen Funktionen zusammensetzen. Hat das Schema für ψ nicht direkt diese Eigenschaft, so möge $a'(a, b, \dots, m)$ aus $a(a, b, \dots, m)$ dadurch entstehen, daß man alle in $a(a, b, \dots, m)$ vorkommenden Funktionen durch die zugehörigen größeren monotonen Funktionen ersetzt. Ebenso gelangen wir von $b(\dots)$ zu $b'(\dots)$. Die folgende Funktion

$$\psi'(a, b, \dots, m, 0) = a'(a, b, \dots, m),$$

$$\psi'(a, b, \dots, m, n+1) = b'_{a_1 \dots a_m}(a, \dots, m, n, \psi(a_1, \dots, m_1, n))$$

ist dann von der gewünschten Art. Sie kann an Stelle von ψ betrachtet werden, da sie niemals einen kleineren Wert hat als ψ . ψ' ist monoton in bezug auf a, b, \dots, m , wie man sich leicht durch Induktion nach n überzeugt. Bei allen folgenden Überlegungen denken wir uns immer stillschweigend den Übergang von ψ zu ψ' schon vollzogen. Den Strich lassen wir der Einfachheit halber fort.

b) Wir zeigen nun, daß man statt unseres allgemeinen Rekursionschemas ein spezielleres der Betrachtung zugrunde legen kann. Wir machen übrigens die folgenden Überlegungen nur für Funktionen von zwei Variablen. Zum Schluß werden wir zeigen, daß diese Einschränkung unwesentlich ist. Wir betrachten also das Schema:

$$\psi(a, 0) = a(a)$$

$$\psi(a, b+1) = b_c(a, b, \psi(c, b)).$$

Wir dürfen hier annehmen, daß $b_c(a, b, \psi(c, b))$ wirklich das Zeichen ψ enthält, sonst handelt es sich ja um bloße Einsetzung und wir kommen auf den Satz 1 zurück. Nehmen wir zunächst an, daß in $b_c(a, b, \psi(c, b))$ kein ψ vorkommt, das in der zu einem anderen ψ gehörigen Klammer steht. In diesem Falle enthält jedes ψ von b in seinem ersten Argument eine Funktion $\alpha_i(a, b)$, die durch Einsetzung von Funktionen unseres Bereiches zustande kommt, die also die beiden Eigenschaften unseres Bereiches hat. Diese Funktionen α seien $\alpha_1(a, b), \dots, \alpha_m(a, b)$. Es sei ferner $\alpha(a, b) = \text{Max}(\alpha_1(a, b), \dots, \alpha_m(a, b))$. Offenbar hat auch α die Eigenschaften der Funktionen unseres Bereiches. Setzen wir nun in $b_c(a, b, \psi(c, b))$ in die ersten Argumente aller ψ statt der $\alpha_i(a, b)$ die Funktion $\alpha(a, b)$ ein, und bilden wir hiermit eine Rekursion, so erhalten wir eine Funktion, die für kein Wertepaar kleiner ist als $\psi(a, b)$. Hat aber in $b_c(a, b, \psi(c, b))$ die Funktion ψ als erstes Argument immer $\alpha(a, b)$, so kann man $b_c(a, b, \psi(c, b))$ auch in der Form schreiben $\omega(a, b, \psi(\alpha(a, b), b))$, wo ω eine Funktion ist, die zu unserem Bereich gehört bzw. durch Einsetzung

aus den Funktionen unseres Bereiches zustande kommt. Wir können uns also auf die Betrachtung des folgenden Schemas beschränken:

$$\begin{aligned}\psi(a, 0) &= a(a), \\ \psi(a, b+1) &= \omega(a, b, \psi(\alpha(a, b), b)).\end{aligned}$$

Nun komme in $\mathfrak{b}_c(a, b, \psi(c, b))$ eine zweimalige Ineinanderschachtelung der ψ vor. Die innersten ψ in \mathfrak{b} enthalten dann im ersten Argument wieder Funktionen $\alpha_i(a, b)$, die wir wieder durch eine einzige Funktion $\alpha(a, b)$ ersetzen können. Die nächstinnersten ψ haben als erstes Argument Funktionen μ_i von a, b und $\psi(\alpha(a, b), b)$. Ebenso wie wir vorher das Maximum der $\alpha_i(a, b)$ nahmen, können wir jetzt das Maximum der $\mu_i(a, b, c)$ benutzen. Falls also in $\mathfrak{b}_c(a, b, \psi(c, b))$ das ψ höchstens zweimal ineinandergeschachtelt vorkommt, so dürfen wir das folgende Rekursionsschema betrachten:

$$\begin{aligned}\psi(a, 0) &= a(a), \\ \psi(a, b+1) &= \omega\{a, b, \psi[\alpha(a, b), b], \psi(\mu(a, b, \psi[\alpha(a, b), b]), b)\}.\end{aligned}$$

Entsprechende Überlegungen können wir anstellen, falls es sich um eine n -fache Ineinanderschachtelung handelt. Hier genügt es, das folgende Schema zu betrachten:

$$\begin{aligned}\psi(a, 0) &= a(a), \\ \psi(a, b+1) &= \omega(a, b, \dots).\end{aligned}$$

Hier hat die Funktion ω außer a und b noch n , also im ganzen $(n+2)$ Argumente. Diese Argumente sehen folgendermaßen aus: Das dritte Argument hat die Gestalt $\psi[\alpha(a, b), b]$. Sind ferner die ersten n Argumente von ω c_1, c_2, \dots, c_n , so lautet das $(n+1)$ te Argument:

$$\psi(\mu(c_1, c_2, \dots, c_n), b),$$

Hier ist μ eine Funktion, die unserem Bereich angehört bzw. sich aus den Funktionen unseres Bereiches zusammensetzt.

c) Wir müssen nun zeigen, daß die Funktion $\psi(a, b)$ oder aber eine größere, auch bei festgehaltenem a in bezug auf b monoton steigt, und daß $\mathfrak{G}_{a_n}(\psi(a, a), \varphi(a, a, n))$ der Fall ist. Wir geben zunächst die größere monotone Funktion an. Wir ersetzen in der Rekursion $\alpha(a, b)$ durch die Funktion $\alpha'(a, b) = \text{Max}(a, \alpha(a, b))$. Statt jeder Funktion μ benutzen wir $\mu'(a, b, \dots, k) = \text{Max}(\mu(a, b, \dots, k), k)$ und statt $\omega(a_1, \dots, a_{n+2})$ nehmen wir $\omega'(a_1, \dots, a_{n+2}) = \text{Max}(\omega(a_1, \dots, a_{n+2}), a_{n+2})$. $\alpha(a)$ wird ersetzt durch $\text{Max}(a, \alpha(a)) = \sigma(a)$. Die neue Funktion ist natürlich wieder monoton in bezug auf a , da alle die Funktionen ω', σ, α' und die Funktionen μ' monoton sind. Diese Funktion ist aber auch monoton in bezug auf b . Wir beweisen zunächst: Wenn für alle a $\psi(a, b) \geq a$, so ist für alle a $\psi(a, b+1) \geq \psi(a, b)$ und $\psi(a, b+1) \geq a$.

Unter der Voraussetzung gelten nämlich die folgenden Überlegungen:

$$\psi(a, b + 1) = \omega'(a, b, \psi[\alpha'(a, b), b], \dots).$$

Für das dritte Argument ist $\psi[\alpha'(a, b), b] \geq a$ und $\psi[\alpha'(a, b), b] \geq \psi(a, b)$. Da nämlich $\psi(a, b)$ monoton ist in bezug auf das erste Argument und da $\alpha'(a, b) \geq a$, so ist $\psi[\alpha'(a, b), b] \geq \psi(a, b)$ und $\psi[\alpha'(a, b), b] \geq a$. — Es sei für ein Argument c von ω' schon gezeigt: $c \geq \psi(a, b)$ und $c \geq a$. Es gelten diese Beziehungen dann auch für das nächste Argument δ . Denn

$$\delta = \psi[\mu(a, b, \dots, c), b]$$

$$\psi[\mu(a, b, \dots, c), b] \geq \psi[\mu(a, b, \dots, \psi(a, b)), b],$$

$$\mu(a, b, \dots, \psi(a, b)) \geq \psi(a, b),$$

$$\psi[\mu(a, b, \dots, \psi(a, b)), b] \geq \psi[\psi(a, b), b],$$

$$\psi(a, b) \geq a; \psi[\psi(a, b), b] \geq \psi(a, b) \text{ (nach Vorauss.)}$$

also $\delta \geq \psi(a, b)$ und $\delta \geq a$.

Da nun $\omega'(a, b, \dots)$ größer oder gleich dem letzten Argument von ω' ist, so ist auch

$$\psi(a, b + 1) \geq \psi(a, b) \text{ und } \psi(a, b + 1) \geq a, \text{ w. z. b. w.}$$

Nun ist $\psi(a, 0) = \sigma(a) \geq a$. Infolgedessen gilt allgemein

$$\psi(a, b + 1) \geq \psi(a, b).$$

Damit ist gezeigt, daß die Funktion ψ die erste Eigenschaft unseres Bereiches hat.

Es bleibt jetzt noch zu zeigen, daß $\mathfrak{G}_{an}(\psi(a, a), \varphi(a, a, n))$. Solange bei $\psi(a, b)$ $b \leq a$ bleibt, genügt es, wenn wir eine vereinfachte Rekursion betrachten:

$$\psi(a, 0) = \sigma(a),$$

$$\psi(a, b + 1) = \omega'(a, a, \dots).$$

Hier ist überall für b a eingesetzt, ausgenommen, wenn b in dem zweiten Argument von ψ steht. Daß diese Funktion nicht kleiner ist als die frühere, folgt einfach aus der Monotonität aller auftretenden Funktionen. Wir bezeichnen nun $\omega'(a, a, \dots, a)$ mit $\omega''(a)$, $\alpha'(a, a)$ mit $\alpha''(a)$ und $\mu'_i(a, \dots, a)$ mit $\mu''_i(a)$. In der Formel $\psi(a, b + 1) = \omega'(a, a, \dots)$ sind die Argumente von ω' alle kleiner als das letzte Argument. Zunächst ist nämlich $\psi(\alpha'(a), b) \geq \psi(a, b) \geq a$.

Es sei schon gezeigt, daß die Argumente bis zum k -ten nie fallen. Das k -te Argument sei δ . Das $(k + 1)$ -te Argument hat dann die Gestalt: $\psi(\mu(a, a, \dots, \delta), b)$. Nun ist $\mu(a, a, \dots, \delta) \geq \delta$

$$\psi(\mu(a, \dots, \delta), b) \geq \psi(\delta, b) \geq \delta.$$

Wir erhalten also keine kleinere Funktion, wenn wir alle Argumente von ω' ,

gleich dem letzten nehmen. Das letzte Argument ist $\psi(\mu'_{n-1}(a, \dots, \delta), b)$, wenn jetzt δ das vorletzte ist. Da alle Argumente von $\mu'_{n-1} \leq \delta$ sind, so ist

$$\psi(\mu'_{n-1}(a, \dots, \delta), b) \leq \psi(\mu''_{n-1}(\delta), b).$$

δ ist wieder kleiner oder gleich $\psi(\mu'_{n-2}(e), b)$, wo e das vorangehende Argument ist. Schließlich erhalten wir keine kleinere Funktion, wenn wir die folgende Rekursion nehmen:

$$\psi(a, 0) = \sigma(a),$$

$$\psi(a, b+1) = \omega''(\psi\{\mu''_{n-1} \dots \psi[\mu''_2(\psi\{\mu'_1[\psi(a''(a), b)], b)\}, b], \dots, b\}).$$

Wir müssen nun feststellen, aus einer wievielfachen Ineinanderschachtelung der Funktionen ω'' , μ''_i , a'' und σ $\psi(a, b+1)$ besteht, falls die Rekursion aufgelöst wird. Man sieht, daß, wenn $\psi(a, b)$ aus einer x -fachen Ineinanderschachtelung besteht, $\psi(a, b+1)$ aus einer $((n)(x+1)+1)$ -fachen Ineinanderschachtelung besteht. Die Ineinanderschachtelung von $\psi(a, b)$ ist also gleich $2(n^b + n^{b-1} + \dots + 1) - 1$. Statt dieser Zahl können wir die nicht kleinere $2(b+1)n^b$ nehmen. Ferner gibt es ein m , so daß alle die Funktionen μ'' , ω'' , a'' , σ' von einem gewissen a an kleiner sind als $\varphi(a, a, m)$. Es gilt also

$$\mathfrak{G}_{am}(\psi(a, a)), \varrho_c(\varphi(c, c, m), a, 2(a+1)n^a).$$

$2(a+1) \cdot n^a$ ist eine Funktion von a , die durch Einsetzung aus Funktionen entsteht, die die Eigenschaften unseres Bereiches haben. Folglich ist:

$$\mathfrak{G}_{ak}(2(a+1)n^a, \varphi(a, a, k)).$$

Für jedes m und k ist ferner

$$\mathfrak{R}_a(\varrho_c(\varphi(c, c, m), a, \varphi(a, a, k)), \varrho_c(\varphi(c, c, m), \varphi(a, a, k), \varphi(a, a, k))).$$

Weiter hat man nach Formel VI:

$$\varrho_c(\varphi(c, c, m), \varphi(a, a, k), \varphi(a, a, k)) \leq \varphi(\varphi(a, a, k), \varphi(a, a, k), m+3)$$

bzw.

$$\varrho_c(\varphi(c, c, m), \varphi(a, a, k), \varphi(a, a, k)) \leq 2.$$

$\varphi(\varphi(a, a, k), \varphi(a, a, k), m+3)$ ist aber eine Funktion, die die Eigenschaften unseres Bereiches hat, da sie durch Einsetzung aus derartigen Funktionen entsteht. Wir haben also

$$\mathfrak{G}_{an}(\psi(a, a), \varphi(a, a, n)), \text{ q. e. d.}$$

d) Für Funktionen von zwei Variablen ist damit der Satz 2 bewiesen. Auf die Funktionen einer Variablen brauchen wir nicht besonders einzugehen. Eine Funktion $\psi(a)$ einer Variablen kann als Spezialfall $\psi(a, 0)$ einer Funktion zweier Variablen aufgefaßt werden. Nun möge die neu

definierte Funktion mehr als zwei Argumente haben. Dieser Fall läßt sich auf den für zwei Variablen zurückführen. Es sei

$$\psi(a, b, \dots, m, 0) = \sigma(a, b, \dots, m),$$

$$\psi(a, b, \dots, m, n+1) = \mathfrak{b}_{a_1, \dots, m_1}(a, b, \dots, m, n, (a_1, \dots, m_1, n)).$$

In $\mathfrak{b}_{a_1, \dots, m_1}(\dots)$ stehen in den Argumenten der innersten ψ Funktionen von a, b, \dots, m . Die ψ haben also eine folgende Gestalt

$$\psi(\alpha_1(a, b, \dots, m, n), \dots, \alpha_r(a, \dots, m, n), n).$$

Hier kann man in alle Argumente von ψ mit Ausnahme des letzten das Maximum aller dieser Funktionen einsetzen. Auf dieselbe Weise kann man erreichen, daß bei den nächstinnersten ψ alle Argumente, mit Ausnahme des letzten, gleich werden, usw. Nun betrachten wir die Rekursion

$$\psi(a, \dots, a, 0) = \sigma(a, a, \dots, a),$$

$$\psi(a, a, \dots, a, b+1) = \mathfrak{b}_{a_1, b_1, \dots, m_1}(a, a, \dots, a, b, \psi(a_1, \dots, m_1, b)).$$

Das ist jetzt ein Rekursionsschema für eine Funktion $\psi'(a, b)$ von zwei Variablen. Es gilt also $\mathfrak{G}_{a_n}(\psi(a, a, \dots, a), \varphi(a, a, n))$. Auch die zu ψ gehörige monotone Funktion läßt sich angeben. Die Monotonität in bezug auf a, b, \dots, m ist schon dadurch gegeben, daß man statt σ und statt der in \mathfrak{b} vorkommenden Funktionen immer die zugehörigen monotonen nimmt. Nun hatten wir ein $\psi'(a, b)$ gefunden, so daß $\psi(a, a, \dots, a, n) \leq \psi'(a, n)$. Es ist $\psi'(a, n) \leq \psi''(a, n)$, wo jetzt ψ'' auch in bezug auf n monoton ist. Die zu $\psi(a, b, \dots, m, n)$ gehörige monotone Funktion ist nun $\psi''(\text{Max}(a, b, \dots, m), n)$. Denn

$$\begin{aligned} \psi(a, b, \dots, m, n) &\leq \psi(\text{Max}(a, b, \dots, m), \text{Max}(\dots), \dots, \text{Max}(\dots), n) \\ &\leq \psi''(\text{Max}(a, b, \dots, m), n). \end{aligned}$$

(Eingegangen am 20. 1. 1927.)

Ein System algebraisch unabhängiger Zahlen.

Von

J. v. Neumann in Berlin.

I. Wir sagen: Die m Zahlen a_1, a_2, \dots, a_m bilden ein algebraisch unabhängiges System, wenn sie die folgende Eigenschaft besitzen: Eine Gleichung

$$\Phi(a_1, a_2, \dots, a_m) = 0,$$

wobei $\Phi(x_1, x_2, \dots, x_m)$ ein Polynom mit rationalen Koeffizienten ist, besteht nur dann, wenn identisch

$$\Phi(x_1, x_2, \dots, x_m) \equiv 0$$

ist.

Das Ziel dieser Arbeit ist, *eine Menge M effektiv anzugeben, von der jede endliche Teilmenge ein algebraisch unabhängiges System ist, und welche die Mächtigkeit des Kontinuums hat.*

Hierzu ist zu bemerken, daß H. Lebesgue¹⁾ und E. Steinitz²⁾ die Existenz einer Menge M^* bewiesen haben, die die zwei folgenden Eigenschaften hat:

1. Jede endliche Teilmenge von M^* ist ein algebraisch unabhängiges System.

2. Zu jeder nicht zu M^* gehörenden Zahl x gibt es eine endliche Teilmenge von M^* , die zusammen mit x kein algebraisch unabhängiges System bildet.

Nun kann man aus 2. nach Steinitz (a. a. O.) leicht schließen, daß M^* (eine „algebraische Basis der Zahlen“) die Mächtigkeit des Kontinuums haben muß. Indessen gelingt die Konstruktion der Lebesgue-Steinitzchen Menge nur durch Anwendung des Wohlordnungssatzes, während wir die

¹⁾ H. Lebesgue, Sur les transformations ponctuelles . . . , Atti Ac. Torino 1906—1907.

²⁾ E. Steinitz, Algebraische Theorie der Körper, Journ. f. r. u. a. Mathematik **137** (1910).

Menge M effektiv angegeben werden — also ohne Benutzung des Auswahlpostulats.

Demgegenüber ist aber festzustellen, daß unserer Menge M die Eigenschaft 2. abgeht (vgl. Fußnote 5). Es ist überhaupt höchst unwahrscheinlich, daß die Konstruktion einer „algebraischen Basis der Zahlen“ ohne Benutzung des Wohlordnungssatzes gelinge.

II. Die Menge M , von der bewiesen werden soll, daß sie die unter I. erwähnte Eigenschaft besitzt, umfaßt die Zahlen

$$A_\varrho = \sum_{\nu=0}^{\infty} \frac{2^{2^{[\varrho \nu]}}}{2^{2^{\nu^2}}} \quad (\varrho > 0), \quad ^3)$$

und nur diese ($[x]$ ist, wie üblich, die größte ganze Zahl $\leq x$). Daß diese Reihe für alle $\varrho > 0$ konvergiert, ist klar, und ebenso, daß aus $\varrho < \sigma$ $A_\varrho < A_\sigma$ folgt. Also hat M die Mächtigkeit des Kontinuums⁴⁾. Es gilt nun zu zeigen, daß jede endliche Teilmenge von M ein algebraisch unabhängiges System ist.

D. h. es muß gezeigt werden: Wenn $\Phi(x_1, x_2, \dots, x_m)$ ein Polynom mit rationalen Koeffizienten ist, und $\varrho_1, \varrho_2, \dots, \varrho_m$ irgendwelche paarweise voneinander verschiedene Zahlen > 0 sind, so folgt aus

$$\Phi(A_{\varrho_1}, A_{\varrho_2}, \dots, A_{\varrho_m}) = 0,$$

daß identisch

$$\Phi(x_1, x_2, \dots, x_m) \equiv 0$$

ist.

Es ist klar, daß wir uns dabei auf Polynome mit lauter ganzen Koeffizienten beschränken dürfen.

III. Wir beweisen zuerst einen Hilfssatz.

Hilfssatz. $\varrho_1, \varrho_2, \dots, \varrho_m$ seien irgendwelche paarweise voneinander verschiedenen Zahlen > 0 . Wenn $\varphi(x_1, x_2, \dots, x_m)$ ein Polynom ist, das nicht identisch $\equiv 0$ ist, so gibt es ein N und ein $\delta > 0$, daß aus $\nu \geq N$

$$|\varphi(2^{2^{[\varrho_1 \nu]}}, 2^{2^{[\varrho_2 \nu]}}, \dots, 2^{2^{[\varrho_m \nu}]})| \geq \delta$$

folgt.

Der Beweis wird durch vollständige Induktion in bezug auf m geführt. Für $m=0$ ist φ eine Konstante, falls $\varphi \neq 0$ ist, so können wir also $N=1, \delta = |\varphi| > 0$ setzen.

³⁾ Damit ist die dyadische Entwicklung der Zahlen A direkt angegeben. Es sei übrigens bemerkt, daß der eigentliche Grund der algebraischen Unabhängigkeit der A_ϱ analog zum Grunde der Transzendenz der Liouvilleschen Zahlen ist. Ein A_ϱ ist durch rationale Zahlen viel besser approximierbar als alle A_σ ($0 < \sigma < \varrho$), und kann mit ihnen darum algebraisch nicht zusammenhängen.

⁴⁾ Daß aus $\varrho \neq \sigma$ immer $A_\varrho \neq A_\sigma$ folgt, kann auch aus der algebraischen Unabhängigkeit geschlossen werden: sonst wäre für $\Phi(x_1, x_2) = x_1 - x_2$ nämlich $\Phi(A_{\varrho_1}, A_{\varrho_2}) = 0$.

Nun sei die Behauptung für m bereits bewiesen, wir wollen sie dann auf $m + 1$ übertragen.

Wir können zunächst ohne Beschränkung der Allgemeinheit annehmen, daß ϱ_{m+1} die größte der Zahlen $\varrho_1, \varrho_2, \dots, \varrho_m, \varrho_{m+1}$ ist. Es ist ferner offenbar

$$\varphi(x_1, x_2, \dots, x_m, x_{m+1}) = \psi_0(x_1, x_2, \dots, x_m) \cdot x_{m+1}^s + \psi_1(x_1, x_2, \dots, x_m) \cdot x_{m+1}^{s-1} + \dots + \psi_s(x_1, x_2, \dots, x_m),$$

wobei nicht identisch $\psi_0(x_1, x_2, \dots, x_m) \equiv 0$ ist. Für $s = 0$ ist $\varphi(x_1, x_2, \dots, x_m, x_{m+1}) = \psi_0(x_1, x_2, \dots, x_m)$, und da die Behauptung für m gilt, nichts zu beweisen. Also sei $s \geq 1$.

Die Behauptung gilt für $\psi_0(x_1, x_2, \dots, x_m)$; wir wählen also N und $\vartheta > 0$ so, daß für $\nu \geq N$

$$\left| \psi_0(2^{2[\varrho_1\nu]}, 2^{2[\varrho_2\nu]}, \dots, 2^{2[\varrho_m\nu]}) \right| \geq \vartheta$$

wird.

Nun sei $\text{Max}(\varrho_1, \varrho_2, \dots, \varrho_m) = \varrho < \varrho_{m+1}$, der höchste der Grade der Polynome ψ_1, \dots, ψ_s sei t , die Summe der Absolutwerte aller ihrer Koeffizienten sei C . Dann ist für $\nu \geq N$

$$\begin{aligned} & \left| \varphi(2^{2[\varrho_1\nu]}, 2^{2[\varrho_2\nu]}, \dots, 2^{2[\varrho_m\nu]}, 2^{2[\varrho_{m+1}\nu]}) \right| \\ & \geq \left| \psi_0(2^{2[\varrho_1\nu]}, \dots, 2^{2[\varrho_m\nu]}) \right| \cdot 2^{s \cdot 2[\varrho_{m+1}\nu]} \\ & - \left\{ \left| \psi_0(2^{2[\varrho_1\nu]}, \dots, 2^{2[\varrho_m\nu]}) \right| \cdot 2^{(s-1) \cdot 2[\varrho_{m+1}\nu]} + \dots + \left| \psi_s(2^{2[\varrho_1\nu]}, \dots, 2^{2[\varrho_m\nu]}) \right| \right\} \\ & \geq \vartheta \cdot 2^{s \cdot 2[\varrho_{m+1}\nu]} - C \cdot 2^{t \cdot 2[\varrho\nu]} \cdot 2^{(s-1) \cdot 2[\varrho_{m+1}\nu]} \\ & = \vartheta \cdot 2^{s \cdot 2[\varrho_{m+1}\nu]} \cdot \left\{ 1 - \frac{C}{\vartheta} 2^{t \cdot 2[\varrho\nu] - 2[\varrho_{m+1}\nu]} \right\}. \end{aligned}$$

Der zweite Faktor ist stets ≥ 1 , der dritte strebt, wegen $\varrho_{m+1} > \varrho$, für $\nu \rightarrow \infty$ gegen 1.

Es gibt also ein N' ($\geq N$), so daß für $\nu \geq N'$

$$\left| \varphi(2^{2[\varrho_1\nu]}, 2^{2[\varrho_2\nu]}, \dots, 2^{2[\varrho_m\nu]}, 2^{2[\varrho_{m+1}\nu]}) \right| \geq \vartheta \cdot 1 \cdot \frac{1}{2} = \frac{\vartheta}{2}$$

wird, und unser Hilfssatz ist restlos bewiesen.

IV. Wir gehen nun zum Beweise der eigentlichen Behauptung über. Also seien $\varrho_1, \varrho_2, \dots, \varrho_m$ wieder irgendwelche paarweise voneinander verschiedene Zahlen > 0 , und $\Phi(x_1, x_2, \dots, x_m)$ ein Polynom mit ganzen Koeffizienten. Dabei sei

$$\Phi(A_{\varrho_1}, A_{\varrho_2}, \dots, A_{\varrho_m}) = 0.$$

Es soll gezeigt werden, daß dann identisch

$$\Phi(x_1, x_2, \dots, x_m) = 0$$

ist. Wir nehmen also das Gegenteil an.

Der Grad von $\Phi(x_1, x_2, \dots, x_m)$ sei s , sein homogener Teil s -ten Grades sei $\Psi(x_1, x_2, \dots, x_m)$. Natürlich ist $\Psi(x_1, x_2, \dots, x_m)$ nicht identisch $\equiv 0$. Ferner werde eine Zahl C so gewählt, daß aus

$$0 \leq x_1 \leq y_1 \leq A_{e_1}, 0 \leq x_2 \leq y_2 \leq A_{e_2}, \dots, 0 \leq x_m \leq y_m \leq A_{e_m}$$

folgt

$$|\Phi(x_1, x_2, \dots, x_m) - \Phi(y_1, y_2, \dots, y_m)| \leq C \text{Max}(y_1 - x_1, y_2 - x_2, \dots, y_m - x_m).$$

Schließlich werde eine ganze Zahl r gewählt, die

$$\geq \text{Max}(e_1, e_2, \dots, e_m)$$

ist; und die Summe der Absolutwerte der Koeffizienten von $\Psi(x_1, x_2, \dots, x_m)$ sei D .

l sei nun eine bestimmte ganze Zahl (≥ 1), über die wir erst später verfügen werden. Es ist wesentlich, daß s, C, D, r von l unabhängig sind.

Es gilt offenbar die Abschätzung

$$\left| \Phi(A_{e_1}, A_{e_2}, \dots, A_{e_m}) - \Phi\left(\sum_{\nu=0}^l \frac{2^{2[e_1, \nu]}}{2^{2\nu^2}}, \sum_{\nu=0}^l \frac{2^{2[e_2, \nu]}}{2^{2\nu^2}}, \dots, \sum_{\nu=0}^l \frac{2^{2[e_m, \nu]}}{2^{2\nu^2}}\right) \right| \\ \leq C \text{Max}\left(\sum_{\nu=l+1}^{\infty} \frac{2^{2[e_1, \nu]}}{2^{2\nu^2}}, \dots, \sum_{\nu=l+1}^{\infty} \frac{2^{2[e_m, \nu]}}{2^{2\nu^2}}\right) \leq C \sum_{\nu=l+1}^{\infty} \frac{2^{2r\nu}}{2^{2\nu^2}}.$$

Weiter ist für $\nu \geq r$

$$\frac{2^{2r(\nu+1)}}{2^{2(\nu+1)^2}} : \frac{2^{2r\nu}}{2^{2\nu^2}} = 2^{-2(\nu+1)^2 - 2r\nu + 2\nu^2 + 2r(\nu+1)} \\ \leq 2^{-2(\nu+1)^2 + 2\nu(\nu+1) + 2\nu^2} < 2^{-2(\nu+1)^2 + 2\nu^2 + \nu + 1} < 2^{-1} = \frac{1}{2},$$

also gilt für $l \geq r$

$$\sum_{\nu=l+1}^{\infty} \frac{2^{2r\nu}}{2^{2\nu^2}} \leq \frac{2^{2r(l+1)}}{2^{2(l+1)^2}} \left(1 + \frac{1}{2} + \frac{1}{4} + \dots\right) \leq \frac{2^{2r(l+1)} \cdot 2}{2^{2(l+1)^2}} \\ = 2^{-2(l+1)^2 + 2r(l+1) + 1} \leq 2^{-2(l+1)^2 + 2l(l+1) + 1} \leq 2^{-2(l+1)^2 + 2l^2 + l + 1} \\ = 2^{-2l^2 + l + 1(2l-1)} \leq 2^{-2l^2 + l + 1}$$

und dies ist für $2^{l+1} \geq s + 1$ (was z. B. für $l \geq s$ sicher der Fall ist)

$$\leq 2^{-(s+1)2^{l^2}} = \frac{1}{2^{(s+1)2^{l^2}}}.$$

Wir können also unser Resultat so zusammenfassen: Für $l \geq r, s$ ist

$$\left| \Phi(A_{e_1}, \dots, A_{e_m}) - \Phi\left(\sum_{\nu=0}^l \frac{2^{2[e_1, \nu]}}{2^{2\nu^2}}, \dots, \sum_{\nu=0}^l \frac{2^{2[e_m, \nu]}}{2^{2\nu^2}}\right) \right| \leq \frac{C}{2^{(s+1) \cdot 2^{l^2}}}.$$

Wegen des Verschwindens von $\Phi(A_{e_1}, \dots, A_{e_m})$ gilt also

$$\left| 2^{s \cdot 2^{l^2}} \cdot \Phi \left(\sum_{\nu=0}^l \frac{2^{2^{[e_1 \nu]}}}{2^{2^{\nu^2}}}, \dots, \sum_{\nu=0}^l \frac{2^{2^{[e_m \nu]}}}{2^{2^{\nu^2}}} \right) \right| \leq \frac{C}{2^{2^{l^2}}}.$$

V. Nun soll der Ausdruck

$$\Phi \left(\sum_{\nu=0}^l \frac{2^{2^{[e_1 \nu]}}}{2^{2^{\nu^2}}}, \dots, \sum_{\nu=0}^l \frac{2^{2^{[e_m \nu]}}}{2^{2^{\nu^2}}} \right)$$

ausmultipliziert werden. Es entstehen dabei endlich viele Glieder, die übrigens alle rational sind, mit Zweierpotenzen als Nennern. Wir teilen sie in drei Gruppen ein, die nacheinander betrachtet werden sollen.

Zur ersten Gruppe zählen wir die Glieder, in denen beim Ausmultiplizieren der Potenzprodukte der

$$\sum_{\nu=0}^l \frac{2^{2^{[e_1 \nu]}}}{2^{2^{\nu^2}}}, \dots, \sum_{\nu=0}^l \frac{2^{2^{[e_m \nu]}}}{2^{2^{\nu^2}}}$$

nicht stets das letzte Glied

$$\frac{2^{2^{[e_1 l]}}}{2^{2^{l^2}}}, \dots, \frac{2^{2^{[e_m l]}}}{2^{2^{l^2}}}$$

genommen wurde. Diese Glieder haben, wie wir soeben bemerkten, Nenner von der Form 2^u , und es ist hier

$$u \leq (s-1)2^{l^2} + 2^{(l-1)^2}.$$

Wenn wir sie mit $2^{s \cdot 2^{l^2}}$ multiplizieren, so entstehen also aus ihnen ganze Zahlen, die durch

$$2^{s \cdot 2^{l^2} - (s-1) \cdot 2^{l^2} - 2^{(l-1)^2}} = 2^{2^{l^2} - 2^{(l-1)^2}}$$

teilbar sind. Da der Exponent

$$\geq 2^{l^2} - 2^{l^2-1} = 2^{l^2-1}$$

ist, sind sie auch durch $2^{2^{l^2-1}}$ teilbar.

Zur zweiten Gruppe zählen wir die Glieder, bei denen bei der Ausmultiplikation zwar stets

$$\frac{2^{2^{[e_1 l]}}}{2^{2^{l^2}}}, \dots, \frac{2^{2^{[e_m l]}}}{2^{2^{l^2}}}$$

genommen wurde, die aber aus einem Gliede von $\Phi(x_1, x_2, \dots, x_m)$ entstanden sind, das nicht den Grad s hat (also einen Grad $< s$). Wieder sind es rationale Zahlen mit einem Nenner 2^u , wobei

$$u \leq (s-1) \cdot 2^{l^2}$$

ist; nach Multiplikation mit $2^{s \cdot 2^{l^2}}$ entsteht also eine ganze Zahl, die durch

$$2^{s \cdot 2^{l^2} - (s-1) \cdot 2^{l^2}} = 2^{2^{l^2}}$$

und um so mehr durch $2^{2^{l^2-1}}$ teilbar ist,

Zur dritten Gruppe schließlich zählen wir die Glieder, bei denen bei der Ausmultiplikation stets

$$\frac{2^{2^{[e_1 l]}}}{2^{2^{l^2}}}, \dots, \frac{2^{2^{[e_m l]}}}{2^{2^{l^2}}}$$

genommen wurde und die aus einem Gliede von $\Phi(x_1, x_2, \dots, x_m)$ entstanden sind, das den Grad s hat. Ihre Summe können wir sofort angeben, sie ist:

$$\frac{\Psi(2^{2^{[e_1 l]}}, \dots, 2^{2^{[e_m l]}})}{2^{s \cdot 2^{l^2}}},$$

und nach der Multiplikation mit $2^{s \cdot 2^{l^2}}$ wird daraus: $\Psi(2^{2^{[e_1 l]}}, \dots, 2^{2^{[e_m l]}})$.

VI. Es wurde also gefunden: Es ist

$$2^{s \cdot 2^{l^2}} \Phi\left(\sum_{\nu=0}^l \frac{2^{2^{[e_1 \nu]}}}{2^{2^{\nu^2}}}, \dots, \sum_{\nu=0}^l \frac{2^{2^{[e_m \nu]}}}{2^{2^{\nu^2}}}\right) = p_l + q_l,$$

wobei p_l (Summe der Glieder der zwei ersten Gruppen) eine durch $2^{2^{l^2-1}}$ teilbare ganze Zahl ist und q_l (Summe der Glieder der dritten Gruppe)

$$= \Psi(2^{2^{[e_1 l]}}, \dots, 2^{2^{[e_m l]}}),$$

also auch eine ganze Zahl ist. Dies gilt für alle hinreichend großen l .

Nun gilt aber nach dem Schlußresultate aus IV. für alle hinreichend großen l

$$|p_l + q_l| \leq \frac{C}{2^{2^{l^2}}}.$$

Die rechte Seite konvergiert gegen 0, wird also schließlich < 1 ; da aber p_l, q_l ganze Zahlen sind, so muß für alle hinreichend großen l

$$p_l + q_l = 0, \quad q_l = -p_l$$

sein. Folglich müßte q_l auch durch $2^{2^{l^2-1}}$ teilbar sein, wir werden aber zeigen, daß für alle hinreichend großen l

$$q_l \neq 0, \quad |q_l| < 2^{2^{l^2-1}}$$

ist, was unmöglich ist.

Die erste Behauptung folgt wegen

$$q_l = \Psi(2^{2^{[e_1 l]}}, \dots, 2^{2^{[e_m l]}})$$

ohne weiteres aus unserem Hilfssatze, es gilt also nur noch die zweite zu beweisen. Nun ist

$$|q_l| = \left| \Psi \left(2^{2^{\lfloor \ell_1 l \rfloor}}, \dots, 2^{2^{\lfloor \ell_m l \rfloor}} \right) \right| \leq D \cdot \left(2^{2^{r_l}} \right)^s = D \cdot 2^{s \cdot 2^{r_l}}.$$

Es ist aber klar, daß dieser Ausdruck schließlich $< 2^{2^{l^2-1}}$ wird.

Damit ist unsere Behauptung restlos bewiesen⁵⁾.

VII. Wir bemerken noch: Wenn jedem $\varrho > 0$ eine rationale Zahl R_ϱ zugeordnet wird, so besitzt die Menge M' aller $A_\varrho + R_\varrho$ auch die in I. formulierten Eigenschaften. In der Tat: M' hat die Mächtigkeit des Kontinuums, weil aus $\varrho \neq \sigma$ $A_\varrho + R_\varrho \neq A_\sigma + R_\sigma$ folgt (sonst wäre für $\Phi(x_1, x_2) = x_1 - x_2 + (R_\varrho - R_\sigma)$ $\Phi(A_\varrho, A_\sigma) = 0$); und aus der algebraischen Unabhängigkeit eines Systems $A_{\varrho_1}, A_{\varrho_2}, \dots, A_{\varrho_m}$ folgt offenbar auch diejenige des Systems $A_{\varrho_1} + R_{\varrho_1}, A_{\varrho_2} + R_{\varrho_2}, \dots, A_{\varrho_m} + R_{\varrho_m}$.

Die Freiheit in der Wahl der R_ϱ können wir zur Erzeugung solcher Mengen M' benutzen, die noch gewissen weiteren Bedingungen genügen.

So können wir z. B. M' in ein (beliebig kleines) Intervall $a < x < b$ ($a < b$) hineinzwängen: es genügt R_ϱ mit $a - A_\varrho < R_\varrho < b - A_\varrho$ zu wählen. Oder wir können auch erreichen, daß jedes Intervall continuum-viele Punkte von M' enthalte. Es sei nämlich I_1, I_2, \dots die Folge aller Intervalle mit rationalen Endpunkten; I_p habe die Endpunkte a_p, b_p ($a_p < b_p$). Wir wählen dann für $p-1 < \varrho \leq p$ $a_p - A_\varrho < R_\varrho < b_p - A_\varrho$.

Bemerkung.

Unsere Menge M erlaubt die Lösung gewisser Aufgaben, die von Herrn M. Mazurkiewicz⁶⁾ und St. Ruziewicz⁷⁾ gestellt worden sind.

⁵⁾ Unser Beweis gilt fast wörtlich für irgendein System von Zahlen C_1, C_2, \dots, C_m

$$C_n = \sum_{\nu=0}^{\infty} \frac{2^{2^{\varphi_n(\nu)}}}{2^{2^{\nu^2}}} \quad (n = 1, 2, \dots, m),$$

wenn die Funktionen $\varphi_n(\nu)$ den folgenden Bedingungen genügen: für $\nu \rightarrow \infty$ gilt

$$\varphi_n(\nu) \rightarrow \infty, \quad \varphi_{n+1}(\nu) - \varphi_n(\nu) \rightarrow \infty, \quad \nu^2 - \varphi_n(\nu) \rightarrow \infty.$$

(Dies gilt natürlich für $\varphi_n(\nu) = [\ell_n \nu]$, wenn $0 < \ell_1 < \ell_2 < \dots < \ell_m$ ist.) Infolgedessen ist es leicht continuum-viele weitere, voneinander und von M algebraisch unabhängige Zahlen anzugeben:

$$B_\varrho = \sum_{\nu=0}^{\infty} \frac{2^{2^{\lfloor \varrho \sqrt{\nu} \rfloor}}}{2^{2^{\nu^2}}} \quad (\varrho > 0),$$

usw. — Übrigens könnten unsere Abschätzungen teilweise dadurch abgekürzt werden, daß die, die A_ϱ definierenden, Reihen mit alternierendem Vorzeichen angesetzt werden (diese Bemerkung verdanke ich Herrn Grandjot). Damit ginge aber der Vorteil verloren, daß direkt die dyadische Entwicklung der Zahlen vorliegt.

⁶⁾ M. Mazurkiewicz: Fund. Math. 1 (1920), Problème 8, S. 224.

⁷⁾ St. Ruziewicz, Sur un ensemble dénombrable ..., Fund. Math. 2 (1921).

Die Aufgabe von Mazurkiewicz ist die: Ein nicht abzählbarer Körper reeller Zahlen werde effektiv angegeben, der nicht alle reellen Zahlen umfaßt. Diese Bedingungen werden offenbar durch die Menge aller rationaler Funktionen mit ganzen Koeffizienten und den A_ϱ ($0 < \varrho < 1$) als Argumenten erfüllt.

Die Aufgabe von Ruziewicz lautet so: Eine nicht abzählbare Menge komplexer Zahlen werde effektiv angegeben derart, daß zwischen zwei verschiedenen Elementen z_1, z_2 desselben niemals eine Relation von der folgenden Form besteht:

$$P(e^i)z_1 + Q(e^i)z_2 + R(e^i) = 0,$$

wobei $P(x), Q(x), R(x)$ irgendwelche Polynome mit rationalen Koeffizienten sind (es sei denn, daß identisch $P(x) \equiv Q(x) \equiv R(x) \equiv 0$ ist). Diese Aufgabe können wir folgendermaßen erledigen:

Es sei etwa M_1 die Menge aller A_ϱ , $0 < \varrho \leq 1$, und M_2 die Menge aller A_ϱ , $1 < \varrho \leq 2$. Beide haben die Mächtigkeit des Kontinuums und eine von beiden genügt der Ruziewicz'schen Bedingung. Denn wäre

$$\begin{aligned} P(e^i)z_1 + Q(e^i)z_2 + R(e^i) &= 0 & (z_1 \neq z_2, z_1, z_2 \text{ aus } M_1), \\ U(e^i)z_3 + V(e^i)z_4 + W(e^i) &= 0 & (z_3 \neq z_4, z_3, z_4 \text{ aus } M_2), \end{aligned}$$

so könnten wir hieraus e^i eliminieren, und es ergäbe sich

$$\Phi(z_1, z_2, z_3, z_4) = 0,$$

wobei $\Phi(z_1, z_2, z_3, z_4)$ offenbar nicht identisch verschwinden kann. Dies ist aber unmöglich, denn z_1, z_2, z_3, z_4 sind verschiedene Elemente von M .

Es mag als ein Mangel erscheinen, daß nicht entschieden wurde, ob M_1 oder M_2 die gewünschte Menge ist. Auch dem ist abzuhelfen, wenn wir e^i durch eine andere transzendente Zahl vom Absolutwerte 1 ersetzen, was für Ruziewicz' Zwecke gleichgültig ist.

Hierzu werde ein rationales R so ausgewählt, daß $-1 < A_1 - R < 1$ ist, und es sei

$$\varepsilon = (A_1 - R) + i \sqrt{1 - (A_1 - R)^2}.$$

ε ist transzendent und $|\varepsilon| = 1$, und wir können die gesuchte Menge etwa als Menge aller A_ϱ , $\varrho > 0$, $\varrho \neq 1$, ansetzen.

(Eingegangen am 5. 2. 1927.)

Über Funktionen mit positivem Realteil.

Von

G. Szegö in Königsberg.

Im Bd. 64 der Math. Annalen hat Herr Carathéodory zum erstenmal die analytischen Funktionen, welche im Innern eines Kreises regulär sind und dort einen positiven Realteil haben, eingehender untersucht¹⁾. Durch eine anschließende Bemerkung von O. Toeplitz ist es ihm weiter gelungen, für diese Klasse von Funktionen mit Hilfe der Koeffizienten ihrer Potenzreihenentwicklung gewisse Kriterien anzugeben, welche ihm und Herrn Fejér später die Möglichkeit lieferten, einige hierher gehörige Extremumsaufgaben in befriedigender Weise zu lösen²⁾.

Diese Fragestellungen und Resultate fanden in den Arbeiten mehrerer Autoren einen großen Widerhall; namentlich sind für den Hauptsatz der Carathéodoryschen Theorie (vgl. unten § 1) verschiedene Beweise gegeben worden³⁾. Die vorliegende Note enthält einen neuen, wie mir scheint recht einfachen Beweis, der von den meisten bisherigen Beweisen ab-

¹⁾ C. Carathéodory, Über den Variabilitätsbereich der Koeffizienten von Potenzreihen, die gegebene Werte nicht annehmen [Math. Annalen 64 (1907), S. 95—115].

²⁾ O. Toeplitz, Über die Fouriersche Entwicklung positiver Funktionen [Rendiconti del Circolo Matematico di Palermo 32 (1911), S. 191—192]; C. Carathéodory, Über den Variabilitätsbereich der Fourierschen Konstanten von positiven harmonischen Funktionen [ebenda, S. 193—217]; C. Carathéodory und L. Fejér, Über den Zusammenhang der Extremen von harmonischen Funktionen mit ihren Koeffizienten und über den Picard-Landauschen Satz [ebenda, S. 218—239].

³⁾ Vgl. die Literaturzusammenstellung in den Enzyklopädieartikeln: L. Lichtenstein, Neuere Entwicklung der Potentialtheorie. Konforme Abbildung [II C 3, S. 177—377], S. 229—230; L. Bieberbach, Neuere Untersuchungen über Funktionen von komplexen Variablen [II C 4, S. 379—532], S. 501—504. — Außer der hier zitierten Literatur sei noch die Arbeit: G. Szegö, Über einen Satz des Herrn Carathéodory [Jahresbericht der Deutschen Math.-Vereinigung 28 (1919), S. 131—137] genannt.

weichend formentheoretische Betrachtungen gänzlich vermeidet und auf allgemeinere Probleme verwandter Art übertragbar ist⁴⁾.

Der Vollständigkeit halber skizziere ich in § 2 einen einfachen Weg von dem Hauptsatz zur Lösung der wichtigsten Carathéodory-Fejérschen Aufgabe. § 3 enthält eine Bemerkung zu dem in § 1 gegebenen Beweise vom Standpunkt der in § 2 behandelten Extremumsaufgabe.

§ 1.

Beweis des Hauptsatzes.

Wie von mehreren Verfassern hervorgehoben worden ist, kann der ganzen Carathéodoryschen Theorie der folgende Satz zugrunde gelegt werden:

Es seien c_1, c_2, \dots, c_n gegebene komplexe Konstanten, die nicht sämtlich verschwinden. Man kann dann die ganze positive Zahl m , $m \leq n$, die positiven Konstanten

$$\gamma_1, \gamma_2, \dots, \gamma_m,$$

schließlich die voneinander verschiedenen Zahlen vom Betrage 1

$$\varepsilon_1 = e^{i\theta_1}, \varepsilon_2 = e^{i\theta_2}, \dots, \varepsilon_m = e^{i\theta_m}$$

derart bestimmen, daß für $\nu = 1, 2, \dots, n$

$$(1) \quad c_\nu = \gamma_1 \varepsilon_1^\nu + \gamma_2 \varepsilon_2^\nu + \dots + \gamma_m \varepsilon_m^\nu$$

wird. Die Zahl m , sowie die Zahlen γ_k, ε_k sind eindeutig bestimmt⁵⁾.

Beweis. 1. Es gibt gewisse positive Zahlen

$$g_1, g_2, \dots, g_{2n+1},$$

sowie Zahlen vom absoluten Betrage 1

$$\eta_1 = e^{it_1}, \eta_2 = e^{it_2}, \dots, \eta_{2n+1} = e^{it_{2n+1}},$$

so daß für $\nu = 1, 2, \dots, n$

$$(1') \quad c_\nu = g_1 \eta_1^\nu + g_2 \eta_2^\nu + \dots + g_{2n+1} \eta_{2n+1}^\nu$$

wird. Man kann sogar, behaupte ich,

$$t_k = k \frac{2\pi}{2n+1} \quad (k = 1, 2, \dots, 2n+1)$$

wählen.

⁴⁾ Der von Herrn Carathéodory a. a. O. ²⁾ gegebene Beweis scheint übrigens auch verallgemeinerungsfähig zu sein (vgl. die Andeutung a. a. O. ²⁾, S. 202). — Auf derartige Verallgemeinerungen möchte ich bei einer anderen Gelegenheit zurückkommen.

⁵⁾ Die Zahlenpaare γ_k, ε_k ($k = 1, 2, \dots, m$) können natürlich untereinander vertauscht werden.

Setzt man nämlich

$$f(t) = 2\Re(c_1 e^{-it} + c_2 e^{-2it} + \dots + c_n e^{-nit}) + C,$$

wo C eine beliebige Konstante bezeichnet, so ist

$$c_\nu = \frac{1}{2\pi} \int_0^{2\pi} f(t) e^{i\nu t} dt \quad (\nu = 1, 2, \dots, n).$$

Wir wählen nun C positiv und so groß, daß $f(t)$ für jedes t positiv ausfällt. Da für ein beliebiges trigonometrisches Polynom $2n$ -ter Ordnung $\varphi(t)$ die Gleichung

$$\frac{1}{2\pi} \int_0^{2\pi} \varphi(t) dt = \frac{1}{2n+1} \sum_{k=1}^{2n+1} \varphi(t_k)$$

besteht, folgt die Behauptung.

2. Weiter stellen wir die folgende

Minimumsaufgabe. *Man betrachte sämtliche nichtnegativen Zahlen*

$$(2) \quad g_1, g_2, \dots, g_{2n+1}$$

und die Zahlen vom absoluten Betrage 1

$$(3) \quad \eta_1 = e^{it_1}, \eta_2 = e^{it_2}, \dots, \eta_{2n+1} = e^{it_{2n+1}},$$

welche den Bedingungen

$$(4) \quad c_\nu = g_1 \eta_1^\nu + g_2 \eta_2^\nu + \dots + g_{2n+1} \eta_{2n+1}^\nu \quad (\nu = 1, 2, \dots, n)$$

genügen, wobei die c_ν gegebene nicht durchweg verschwindende Konstanten sind. Es soll

$$(5) \quad g_1 + g_2 + \dots + g_{2n+1}$$

zu einem Minimum gemacht werden.

Wir haben vorhin gezeigt, daß es Wertsysteme (2), (3) von der zugelassenen Art gibt. Es ist ferner klar, daß das fragliche Minimum existiert. (Man kann ja annehmen, daß die Summen (5), also auch sämtliche g_k , beschränkt sind.) Es sei nun

$$g_1 = \gamma_1, g_2 = \gamma_2, \dots, g_{2n+1} = \gamma_{2n+1},$$

$$\eta_1 = \varepsilon_1 = e^{i\theta_1}, \eta_2 = \varepsilon_2 = e^{i\theta_2}, \dots, \eta_{2n+1} = \varepsilon_{2n+1} = e^{i\theta_{2n+1}}$$

ein Wertsystem, für welches das Minimum angenommen wird. Es können dabei gewisse γ_k verschwinden, jedoch nicht alle, da die c_ν nicht sämtlich Null sind. Es können ferner gewisse ε_k zusammenfallen. Durch Weglassen bzw. Zusammenfassen der entsprechenden Glieder in (4) sieht man jedenfalls folgendes:

Es gibt eine positive ganze Zahl m ($m \leq 2n + 1$), ferner positive Zahlen

$$(6) \quad \gamma_1, \gamma_2, \dots, \gamma_m$$

und voneinander verschiedene Zahlen vom Betrage 1

$$(7) \quad \varepsilon_1 = e^{i\theta_1}, \varepsilon_2 = e^{i\theta_2}, \dots, \varepsilon_m = e^{i\theta_m}$$

derart, daß

1. sämtliche Gleichungen

$$(8) \quad c_\nu = \gamma_1 \varepsilon_1^\nu + \gamma_2 \varepsilon_2^\nu + \dots + \gamma_m \varepsilon_m^\nu \quad (\nu = 1, 2, \dots, n)$$

erfüllt sind;

2. die Summe

$$\gamma_1 + \gamma_2 + \dots + \gamma_m$$

niemals abnehmen kann, wenn man die reellen Zahlen $\gamma_1, \gamma_2, \dots, \gamma_m, \theta_1, \theta_2, \dots, \theta_m$ unter Beachtung der Gleichungen (8) variiert⁶⁾.

3. Für $m \leq n$ ist damit unser Satz bewiesen. Wir zeigen nun, daß der Fall $m > n$ überhaupt nicht eintreten kann. Dann könnte nämlich die Lagrangesche Methode der Multiplikatoren angewendet werden⁷⁾. Es existierten danach $2n + 1$ Konstanten $\lambda_0, \lambda_1, \mu_1, \dots, \lambda_n, \mu_n$, die nicht durchweg gleich Null sind, derart, daß die partiellen Ableitungen von

$$(10) \quad \begin{aligned} & \lambda_0 (\gamma_1 + \gamma_2 + \dots + \gamma_m) \\ & + \sum_{\nu=1}^n \lambda_\nu (\gamma_1 \cos \nu \theta_1 + \gamma_2 \cos \nu \theta_2 + \dots + \gamma_m \cos \nu \theta_m) \\ & + \sum_{\nu=1}^n \mu_\nu (\gamma_1 \sin \nu \theta_1 + \gamma_2 \sin \nu \theta_2 + \dots + \gamma_m \sin \nu \theta_m) \end{aligned}$$

nach den Variablen γ_k, θ_k für das oben erwähnte spezielle Wertsystem sämtlich verschwinden. Dies würde aber heißen, daß

$$\begin{aligned} \lambda_0 + \sum_{\nu=1}^n (\lambda_\nu \cos \nu \theta_k + \mu_\nu \sin \nu \theta_k) &= 0, \\ \sum_{\nu=1}^n \nu (-\lambda_\nu \sin \nu \theta_k + \mu_\nu \cos \nu \theta_k) &= 0 \quad (k = 1, 2, \dots, m) \end{aligned}$$

gilt, oder daß das nicht identisch verschwindende trigonometrische Polynom n -ter Ordnung

$$\lambda_0 + \sum_{\nu=1}^n (\lambda_\nu \cos \nu \theta + \mu_\nu \sin \nu \theta)$$

die Doppelwurzeln $\theta_1, \theta_2, \dots, \theta_m$ besitzt, was wegen $m > n$ unmöglich ist. Damit ist die Darstellbarkeit der c_ν in der Form (1) gezeigt.

⁶⁾ Eine derartige Variation ist freilich nicht notwendigerweise möglich.

⁷⁾ Vgl. R. Courant und D. Hilbert, Methoden der mathematischen Physik I [Berlin:

4. Daß die Zahlen $m, \gamma_k, \varepsilon_k$ eindeutig bestimmt sind, kann auf die bekannte Weise bewiesen werden. Der Vollständigkeit halber sei dies noch ausgeführt.

Man setze

$$c_{-p} = \bar{c}_p \quad (p = 1, 2, \dots, n),$$

$$c_0 = \gamma_1 + \gamma_2 + \dots + \gamma_m;$$

dann ist

$$(11) \quad \sum_{p,q}^n c_{p-q} x_p \bar{x}_q = \sum_{k=1}^m \gamma_k |x_0 + x_1 \varepsilon_k + \dots + x_n \varepsilon_k^n|^2,$$

so daß m der Rang der linksstehenden Hermiteschen Form ist.

Die ε_k genügen der Gleichung

$$[c_{p-q+1} - \varepsilon c_{p-q}] = 0 \quad (p, q = 0, 1, \dots, m-1).$$

Die linksstehende Determinante entsteht nämlich durch zeilenweise Multiplikation aus den beiden folgenden:

$$[\varepsilon_{q+1}^{p+1} - \varepsilon \varepsilon_{q+1}^p], \quad [\gamma_{q+1} \varepsilon_{q+1}^{-p}] \quad (p, q = 0, 1, \dots, m-1),$$

von denen die zweite $\neq 0$ ist, während die erste für $\varepsilon = \varepsilon_k$ ($k = 1, 2, \dots, m$) und nur für diese Werte verschwindet.

Schließlich können die γ_k aus (11) berechnet werden, indem man das Polynom

$$x_0 + x_1 z + \dots + x_n z^n$$

so bestimmt, daß es für $z = \varepsilon_h$ ($h = 1, 2, \dots, m; h \neq k$) verschwindet und für $z = \varepsilon_k$ den Wert 1 erhält (was wegen $m \leq n$ stets möglich ist).

Damit ist der Hauptsatz bewiesen.

§ 2.

Über eine Carathéodory-Fejérsche Aufgabe.

Wir betrachten die Menge \mathcal{U} der im Einheitskreise $r < 1$ regulären harmonischen Funktionen

$$u(r, \theta) = a_0 + 2r(a_1 \cos \theta + b_1 \sin \theta) + 2r^2(a_2 \cos 2\theta + b_2 \sin 2\theta) + \dots,$$

deren $2n + 1$ erste Koeffizienten

$$a_0, a_1, b_1, \dots, a_n, b_n$$

vorgeschriebene Werte haben. Jede derartige Funktion besitzt im Einheitskreise $r < 1$ eine bestimmte untere Grenze μ ⁸⁾. Gesucht wird die obere Grenze (*Maximum*) dieser Zahlen μ , wenn $u(r, \theta)$ sämtliche Funktionen der Menge \mathcal{U} durchläuft.

⁸⁾ Diese Zahl μ ist bekanntlich niemals ein Minimum, es sei denn, daß $u(r, \theta) \equiv \text{konst.}$ ist.

Wir können voraussetzen, daß die Zahlen

$$c_\nu = a_\nu + i b_\nu \quad (\nu = 1, 2, \dots, n)$$

nicht sämtlich Null sind. (Sonst ist offenbar a_0 das gesuchte Maximum, das nur für $u(r, \theta) = a_0$ erreicht wird.) Man wende auf diese Zahlen c_ν das Ergebnis des § 1 an und setze mit den dort ermittelten positiven Konstanten $\gamma_1, \gamma_2, \dots, \gamma_m$

$$(12) \quad c_0 = a_0 = \gamma_0 + \gamma_1 + \gamma_2 + \dots + \gamma_m.$$

Die so bestimmte Zahl γ_0 ist gleich dem gesuchten Maximum. Es wird allein für die Funktion

$$(13) \quad u^*(r, \theta) = \gamma_0 + \sum_{k=1}^m \gamma_k \frac{1-r^2}{1-2r \cos(\theta-\theta_k)+r^2}$$

erreicht.

Beweis. 1. Es gelten die Gleichungen

$$c_\nu - \varepsilon_{\nu 0} \gamma_0 = \sum_{k=1}^m \gamma_k \varepsilon_k^\nu \quad (\nu = 0, 1, 2, \dots, n),$$

wobei $\varepsilon_{\nu 0} = 0$ oder 1 sei, je nachdem $\nu \neq 0$ oder $\nu = 0$ ist. Sie bedeuten, daß für $\nu = 0, \pm 1, \pm 2, \dots, \pm n$

$$\lim_{r \rightarrow 1} \frac{1}{2\pi} \int_0^{2\pi} [u(r, \theta) - \gamma_0] e^{i\nu\theta} d\theta = \sum_{k=1}^m \gamma_k \varepsilon_k^\nu$$

ist. D. h. bei beliebigem x_0, x_1, \dots, x_n

$$(14) \quad \lim_{r \rightarrow 1} \frac{1}{2\pi} \int_0^{2\pi} [u(r, \theta) - \gamma_0] |x_0 + x_1 z + \dots + x_n z^n|^2 d\theta \\ = \sum_{k=1}^m \gamma_k |x_0 + x_1 \varepsilon_k + \dots + x_n \varepsilon_k^n|^2 \quad (z = e^{i\theta}).$$

Wegen $m \leq n$ kann nun das Polynom $x_0 + x_1 z + \dots + x_n z^n$ so bestimmt werden, daß es die einzigen Nullstellen $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_m$ besitzt. Dann ist

$$\lim_{r \rightarrow 0} \frac{1}{2\pi} \int_0^{2\pi} [u(r, \theta) - \gamma_0] |x_0 + x_1 z + \dots + x_n z^n|^2 d\theta = 0 \quad (z = e^{i\theta}),$$

ferner

$$(15) \quad \lim_{r \rightarrow 1} \frac{1}{2\pi} \int_0^{2\pi} [u(r, \theta) - \mu] |x_0 + x_1 z + \dots + x_n z^n|^2 d\theta \\ = (\gamma_0 - \mu)(|x_0|^2 + |x_1|^2 + \dots + |x_n|^2) \quad (z = e^{i\theta}),$$

woraus $\mu \leq \gamma_0$ folgt.

2. Es sei nun $\mu = \gamma_0$. Wir zeigen dann, daß $u(r, \theta) = u^*(r, \theta)$ gilt. Das Polynom $x_0 + x_1 z + \dots + x_n z^n$ sei zunächst wie eben bestimmt. In (15) ist jetzt der Integrand ≥ 0 . Man schließt daraus unmittelbar die Gleichung

$$\lim_{r \rightarrow 1} \int_J [u(r, \theta) - \gamma_0] d\theta = 0;$$

hierbei ist J ein beliebiges abgeschlossenes Intervall, das keinen zu den $\theta_k \bmod 2\pi$ kongruenten Punkt enthält⁹⁾.

Bestimmt man anderseits das Polynom $x_0 + x_1 z + \dots + x_n z^n$ in (14) derart, daß es für $z = \varepsilon_k$ ($k = 1, 2, \dots, m; k \neq 0$) verschwindet und für $z = \varepsilon_0$ den Wert 1 erhält, so gewinnt man das folgende Resultat: Bedeuten J_1, J_2, \dots, J_m abgeschlossene Intervalle, so daß J_k die Zahl θ_k , jedoch keine den anderen $\bmod 2\pi$ kongruenten Zahlen enthält, so ist

$$\lim_{r \rightarrow 1} \frac{1}{2\pi} \int_{J_k} [u(r, \theta) - \gamma_0] d\theta = \gamma_k \quad (k = 1, 2, \dots, m).$$

(Hieraus folgt leicht für eine überall stetige Funktion $f(\theta)$, welche die Periode 2π besitzt,

$$\lim_{r \rightarrow 1} \frac{1}{2\pi} \int_0^{2\pi} [u(r, \theta) - \gamma_0] f(\theta) d\theta = \sum_{k=1}^m \gamma_k f(\theta_k). \quad 10)$$

Es gilt insbesondere

$$\lim_{r \rightarrow 1} \frac{1}{2\pi} \int_0^{2\pi} [u(r, \theta) - \gamma_0] e^{in\theta} d\theta = \sum_{k=1}^m \gamma_k e^{in\theta_k} \quad (n = 0, 1, 2, \dots),$$

woraus $u(r, \theta) = u^*(r, \theta)$ folgt.

§ 3.

Schlußbemerkung.

Vom Standpunkt der eben behandelten Carathéodory-Fejérschen Aufgabe kann der Satz von § 1 und sein dort gegebener Beweis folgendermaßen interpretiert werden.

Der Satz ist offenbar gleichbedeutend mit der Aussage, daß die in § 2 erklärte Funktionenmenge \mathcal{U} genau eine Funktion der Form

$$(13') \quad \gamma_0 + \sum_{k=1}^m \gamma_k \frac{1-r^2}{1-2r \cos(\theta-\theta_k) + r^2}$$

⁹⁾ Vgl. die analogen Überlegungen in meiner Arbeit, Koeffizientenabschätzungen bei ebenen und räumlichen harmonischen Entwicklungen [Math. Annalen 96 (1927), S. 601-632], § 1, 2.

¹⁰⁾ Vgl. a. a. O. ⁹⁾, § 1, 3, 4.

enthält. Hierbei sind die γ_k sämtlich positiv, die Zahlen $e^{i\theta_k}$ voneinander verschieden ($k = 1, 2, \dots, m$) und $m \leq n$.

Der *Beweis* beruht auf folgender Modifikation der Carathéodory-Fejérschen Aufgabe. Man betrachte die Untermenge \mathfrak{U}' von \mathfrak{U} , welche aus allen Funktionen von der Form (13') besteht, jedoch mit $m \leq 2n + 1$, und suche ihre untere Grenze, nämlich γ_0 , möglichst groß zu machen. Da

$$\gamma_0 + \gamma_1 + \gamma_2 + \dots + \gamma_m = a_0$$

vorgeschrieben ist, kommt diese Forderung darauf hinaus, daß $\gamma_1 + \gamma_2 + \dots + \gamma_m$ ein Minimum sei. Dies war aber gerade die Minimumsaufgabe von § 1, 2.

Königsberg, Januar 1927.

(Eingegangen am 19. 1. 1927.)

Eine Illustration zur Riemannschen Vermutung.

Von

Guido Hoheisel in Breslau.

Die Riemannsche Vermutung besagt, daß für jedes $\delta > 0$

$$N\left(\frac{1}{2} + \delta, T\right) = 0$$

ist. Hier bedeutet $N(\sigma_0, T)$ die Anzahl der Nullstellen von $\zeta(s)$ für $\sigma \geq \sigma_0$ und $1 \leq t \leq T$ ($s = \sigma + it$). Pólya hat nun ausgehend von einer Integraldarstellung der Funktion

$$\xi(s) = \frac{s(s-1)}{2} \Gamma\left(\frac{s}{2}\right) \pi^{-\frac{s}{2}} \zeta(s)$$

durch eine geringe Abänderung des Integranden eine Funktion $\xi^*(s)$ erhalten, deren Nullstellen sämtlich auf der Geraden $\sigma = \frac{1}{2}$ lagen. Wir wollen hier durch eine denkbar einfache Konstruktion eine Funktion $\zeta^*(s)$ angeben, die für ein beliebig klein gegebenes festes $\varkappa > 0$ und alle $T \geq 1$ die Beziehung

$$N\left(\frac{1}{2} + \varkappa, T\right) = 0$$

erfüllt, außerdem aber noch viele Eigenschaften mit der ζ -Funktion gemein hat:

1. ζ^* genügt derselben Funktionsgleichung.
2. ζ^* hat ebenfalls nur den Pol $s = 1$ sowie die trivialen Nullstellen der ζ -Funktion, ja „fast alle“ Nullstellen der ζ -Funktion.
3. $\zeta^*(s)$ hat dieselben Fourier-Koeffizienten wie $\zeta(s)$:

$$\lim_{\frac{1}{T}} \int_0^T \zeta^*(\sigma + it) e^{i\lambda t} dt = \lim_{\frac{1}{T}} \int_0^T \zeta(\sigma + it) e^{i\lambda t} dt,$$

gültig für alle λ und nicht zu kleine σ .

4. ζ^* ist generalisiert fastperiodisch im Sinne von Besicovitch; genauer ist (für dieselben σ wie bei 3.)

$$\lim \frac{1}{T} \int_0^T |\zeta(\sigma + it) - \zeta^*(\sigma + it)|^2 dt = 0.$$

5. Ist die Riemannsche Vermutung richtig, so ist

$$\zeta^*(s) \equiv \zeta(s).$$

Könnte man statt 4. zeigen

4a. $\zeta^*(s)$ ist fastperiodisch (für irgendein σ),
so folgte aus 3.:

$$\zeta^*(s) \equiv \zeta(s),$$

d. h. die Riemannsche Vermutung.

$\varkappa > 0$ ist also jetzt fest. Die eventuell vorhandenen Nullstellen von $\zeta(s)$ im Gebiete $\sigma \geq \frac{1}{2} + \varkappa$; $t > 0$ ordnen wir nach wachsenden Koordinaten (jede in ihrer Vielfachheit gezählt):

$$s_{1,1}, s_{2,1}, s_{3,1}, s_{4,1}, \dots,$$

$$s_{\nu,1} - \frac{1}{2} = r_\nu e^{i\left(\frac{\pi}{2} - \delta_\nu\right)}.$$

Aus der Funktionalgleichung und der Konjugiertheit der ζ -Funktion ergeben sich noch die Nullstellen¹⁾

$$s_{\nu,3} = \frac{1}{2} + r_\nu e^{-i\left(\frac{\pi}{2} - \delta_\nu\right)}; \quad s_{\nu,3} = \frac{1}{2} + r_\nu e^{i\left(\frac{\pi}{2} + \delta_\nu\right)};$$

$$s_{\nu,4} = \frac{1}{2} + r_\nu e^{-i\left(\frac{\pi}{2} + \delta_\nu\right)}.$$

Wegen des Carlsonschen Satzes (Arkiv Mat. Astr. Fys. 15)

$$N\left(\frac{1}{2} + \varkappa, T\right) = O(T^{1-4\varkappa^2+0(1)})$$

ist das Produkt

$$g(s) = \prod_{\nu}^{1, \infty} \left(1 - \frac{s - \frac{1}{2}}{s_{\nu,1} - \frac{1}{2}}\right) \dots \left(1 - \frac{s - \frac{1}{2}}{s_{\nu,4} - \frac{1}{2}}\right)$$

konvergent, denn es ist wegen Carlson

$$\sum \frac{1}{r_\nu}$$

konvergent.

¹⁾ Reelle Nullstellen gibt es bekanntlich nicht in $0 < \sigma < +1$.

Definieren wir nun $H(s)$ durch die Gleichung

$$\zeta(s) = H(s) \cdot g(s),$$

definieren wir ferner

$$g^*(s) = \prod_{\nu}^{1, \infty} \left(1 - \frac{s - \frac{1}{2}}{+i r_{\nu}}\right) \left(1 - \frac{s - \frac{1}{2}}{-i r_{\nu}}\right) \left(1 - \frac{s - \frac{1}{2}}{+i r_{\nu}}\right) \left(1 - \frac{s - \frac{1}{2}}{-i r_{\nu}}\right),$$

so behaupten wir jetzt:

$$\zeta^*(s) = H(s) \cdot g^*(s)$$

hat die erwünschten Eigenschaften.

$\zeta^*(s)$ ergibt sich also aus $\zeta(s)$, indem man alle Nullstellen der ζ -Funktion, die nicht im Streifen $\frac{1}{2} - \varkappa < \sigma < \frac{1}{2} + \varkappa$ liegen, unter Erhaltung ihres Absolutbetrages auf die Gerade $\sigma = \frac{1}{2}$ kehrt.

Die Eigenschaften 2 und 5 gehen unmittelbar aus der Konstruktion hervor. Da ferner

$$g(1-s) = g(s); \quad g^*(1-s) = g^*(s)$$

gilt, so ist

$$\begin{aligned} \zeta^*(1-s) &= \zeta(1-s) \cdot \frac{g^*(1-s)}{g(1-s)} = \zeta(1-s) \cdot \frac{g^*(s)}{g(s)} \\ &= \chi(s) \cdot \zeta(s) \cdot \frac{g^*(s)}{g(s)} = \chi(s) \zeta^*(s), \end{aligned}$$

also Eigenschaft 1 erfüllt.

3 und 4 werden sich gleichzeitig ergeben.

Zu ihrer Herleitung betrachten wir den Quotienten

$$\begin{aligned} \frac{\zeta(s)}{\zeta^*(s)} &= \frac{g(s)}{g^*(s)} = \prod_{\nu} \frac{\left(s - \frac{1}{2} - r_{\nu} e^{i\left(\frac{\pi}{2} - \delta_{\nu}\right)}\right) \left(s - \frac{1}{2} - r_{\nu} e^{-i\left(\frac{\pi}{2} + \delta_{\nu}\right)}\right)}{\left(s - \frac{1}{2} - r_{\nu} i\right) \left(s - \frac{1}{2} + r_{\nu} i\right)} \\ &\times \prod_{\nu} \frac{\left(s - \frac{1}{2} - r_{\nu} e^{-i\left(\frac{\pi}{2} - \delta_{\nu}\right)}\right) \left(s - \frac{1}{2} - r_{\nu} e^{i\left(\frac{\pi}{2} + \delta_{\nu}\right)}\right)}{\left(s - \frac{1}{2} + r_{\nu} i\right) \left(s - \frac{1}{2} - r_{\nu} i\right)} \\ &= \prod_{\nu} \frac{\left(r e^{i\left(\frac{\pi}{2} - \delta\right)} - r_{\nu} e^{i\left(\frac{\pi}{2} - \delta_{\nu}\right)}\right) \left(r e^{i\left(\frac{\pi}{2} - \delta\right)} - r_{\nu} e^{-i\left(\frac{\pi}{2} + \delta_{\nu}\right)}\right)}{\left(r e^{i\left(\frac{\pi}{2} - \delta\right)} - r_{\nu} i\right) \left(r e^{i\left(\frac{\pi}{2} - \delta\right)} + r_{\nu} i\right)} \\ &\times \prod_{\nu} \frac{\left(r e^{i\left(\frac{\pi}{2} - \delta\right)} - r_{\nu} e^{-i\left(\frac{\pi}{2} - \delta_{\nu}\right)}\right) \left(r e^{i\left(\frac{\pi}{2} - \delta\right)} - r_{\nu} e^{i\left(\frac{\pi}{2} + \delta_{\nu}\right)}\right)}{\left(r e^{i\left(\frac{\pi}{2} - \delta\right)} + r_{\nu} i\right) \left(r e^{i\left(\frac{\pi}{2} - \delta\right)} - i r_{\nu}\right)}, \end{aligned}$$

indem $s - \frac{1}{2} = r e^{i\left(\frac{\pi}{2} - \delta\right)}$ gesetzt ist.

Das zweite Produkt unterscheidet sich vom ersten nur dadurch, daß $-\delta_\nu$ an Stelle von δ_ν steht. Vor der Hand wird es genügen, das erste Produkt abzuschätzen. Der allgemeine Faktor darin lautet:

$$F_\nu = \frac{r_\nu^2 e^{-2i\delta_\nu} - r^2 e^{-2i\delta}}{r_\nu^2 - r^2 e^{-2i\delta}}.$$

Diese Faktoren zerfallen in drei Klassen:

$$1. \quad \frac{r_\nu}{r} \leq \frac{1}{2}; \quad 2. \quad \frac{1}{2} < \frac{r_\nu}{r} < 2; \quad 3. \quad \frac{r_\nu}{r} \geq 2.$$

Nur die dritte Klasse enthält unendlich viele Elemente. Für die erste Klasse ergibt sich:

$$F_\nu = 1 + \frac{r_\nu^2 e^{2i\delta} (1 - e^{-2i\delta_\nu})}{r^2 - r_\nu^2 e^{2i\delta}} = 1 + G_\nu,$$

$$|G_\nu|^2 = \frac{4r_\nu^4 \sin^2 \delta_\nu}{(r^2 - r_\nu^2)^2 + 4r^2 r_\nu^2 \sin^2 \delta} < \frac{r_\nu^2}{(r^2 - r_\nu^2)^2} < \frac{r^2}{4(r_\nu + r)^2 (r - r_\nu)^2} < \frac{r^2 \cdot 4}{4r^2 \cdot r^2}$$

$$|G_\nu| < \frac{1}{r} = \frac{1}{r^2}.$$

Es gibt $-r$ soll groß sein — nach Carlson höchstens $r^{1-4\kappa^2+0(1)} < r^{1-2\kappa}$
Faktoren der ersten Klasse:

$$\prod_{\left| \frac{r_\nu}{r} \right| \leq \frac{1}{2}} F_\nu < \left(1 + \frac{1}{r}\right)^{r^{1-2\kappa}} = \left(1 + \frac{1}{r^{1-2\kappa}} \cdot \frac{1}{r^{2\kappa}}\right)^{r^{1-2\kappa}} < e^{\frac{1}{r^{2\kappa}}}.$$

Abschätzung der Faktoren der dritten Klasse:

$$e^{2i\delta_\nu} F_\nu = \frac{r_\nu^2 - r^2 e^{-2i(\delta - \delta_\nu)}}{r_\nu^2 - r^2 e^{-2i\delta}} = 1 + \frac{r^2 e^{-2i\delta} (1 - e^{2i\delta_\nu})}{r_\nu^2 - r^2 e^{-2i\delta}} = 1 + H_\nu,$$

$$|H_\nu|^2 = \frac{4r^2 \sin^2 \delta_\nu}{(r_\nu^2 - r^2)^2 + 4r^2 r_\nu^2 \sin^2 \delta} < \frac{4r^2 \sin^2 \delta_\nu \left(\frac{r}{r_\nu}\right)^2}{(r_\nu + r)^2 (r_\nu - r)^2} < \frac{\left(\frac{1}{2}\right)^2}{r_\nu^2 \cdot r^2} = \frac{1}{4r^2 r_\nu^2},$$

$$\left| \prod_{\left| \frac{r_\nu}{r} \right| > 2} F_\nu \right| < \prod \left(1 + \frac{1}{2rr_\nu}\right)$$

$$= \prod \left(1 + \frac{1}{2rr_\nu}\right)^{\frac{1}{r_\nu} \cdot r_\nu} < \prod \left(e^{\frac{1}{2r}}\right)^{\frac{1}{r_\nu}} < \left(e^{\frac{1}{2r}}\right)^{\sum_{\nu=1}^{\infty} \frac{1}{r_\nu}}.$$

Da aber $\sum_{\nu=1}^{\infty} \frac{1}{r_\nu}$ konvergent ist nach Carlson, so folgt

$$\left| \prod_{\left| \frac{r_\nu}{r} \right| > 2} F_\nu \right| < e^{\frac{c}{2r}}.$$

Für die weiteren Abschätzungen sei $s = \frac{1}{2} + re^{i\left(\frac{\pi}{2} - \delta\right)}$ rechts von $\sigma = \frac{1}{2} + \kappa$ angenommen, was ja wegen der Nullstellen von ζ^* nötig ist. Es ist also

$$r \sin \delta > \kappa$$

Abschätzung für Faktoren der Klasse 2:

Wir nehmen davon zunächst nur die Faktoren, für welche

$$|r - r_v| > \log^2 r$$

ist²⁾.

$$F_v = 1 + \frac{r_v^2 e^{2i\delta} (1 - e^{-2i\delta v})}{r^2 - r_v^2 e^{2i\delta}} = 1 + G_v.$$

Nehmen wir jetzt gleich den vom zweiten Produkt für $\frac{\xi}{\zeta^*}$ herrührenden entsprechenden Faktor dazu (δ_v ist in $-\delta_v$ zu verwandeln), so ergibt sich:

$$1 + \frac{r_v^2 e^{2i\delta} (2 - e^{-2i\delta v} - e^{2i\delta v})}{r^2 - r_v^2 e^{2i\delta}} + \frac{r_v^4 e^{4i\delta} (1 - e^{-2i\delta v})(1 - e^{2i\delta v})}{(r^2 - r_v^2 e^{2i\delta})^2} = 1 + L_v + M_v,$$

$$L_v = \frac{4r_v^2 e^{2i\delta} \sin^2 \delta_v}{r^2 - r_v^2 e^{2i\delta}}; \quad |M_v| = |G_v|^2, \quad |L_v| < \frac{1}{|r^2 - r_v^2 e^{2i\delta}|},$$

$$|L_v| < \frac{1}{(r^2 - r_v^2)^2 + 4r_v^2 r^2 \sin^2 \delta} < \frac{1}{r_v^2 \left[(r - r_v)^2 + 4 \left(\sigma - \frac{1}{2} \right)^2 \right]}$$

$$< \frac{2}{r_v^2 (|r - r_v| + 2\sigma - 1)^2},$$

$$|M_v| = |G_v|^2 = \frac{4r_v^4 \sin^2 \delta_v}{(r^2 - r_v^2)^2 + 4r_v^2 r^2 \sin^2 \delta} < \frac{r_v^2}{r_v^2 \left[(r - r_v)^2 + 4 \left(\sigma - \frac{1}{2} \right)^2 \right]}$$

$$< \frac{2}{(|r - r_v| + 2\sigma - 1)^2}.$$

Die Anzahl der Faktoren der zweiten Klasse ist nach Carlson höchstens

$$(2r)^{1-4\kappa^2+0(1)} < 2r^{1-2\kappa}$$

$$\left| \prod_{2 > \frac{r_v}{r} > \frac{1}{2}} F_v \right| < \prod \left(1 + \frac{\sqrt{2}}{r_v (|r - r_v| + 2\sigma - 1)} + \frac{2}{(|r - r_v| + 2\sigma - 1)^2} \right)$$

$$< \prod \left(1 + \frac{2 + \sqrt{2} |r - r_v| + 2\sigma - 1}{r_v (|r - r_v| + 2\sigma - 1)^2} \right) < \prod \left(1 + \frac{6 + \frac{8\sigma}{r}}{(|r - r_v| + 2\sigma - 1)^2} \right).$$

²⁾ Es könnte auch irgendeine andere Funktion sein, falls sie nur stärker als $\log r$ wächst.

Wir zerlegen nun in weitere Produkte, indem wir die Faktoren, für welche $|r - r_y|$ zwischen zwei ganzen Zahlen l und $l+1$ liegt, zusammenfassen. Es gibt gemäß der v. Mangoldt'schen Nullstellenformel höchstens $O(\log r)$ solcher Nullstellen. Für dieses Produkt ist alsdann

$$\left| \prod_{l < |r - r_y| < l+1} F_y \right| < \left(1 + \frac{6 + 8 \frac{\sigma}{r}}{(l + 2\sigma - 1)^2} \right)^{c \cdot \log r}$$

$$\geq \left| \frac{r_y}{r} \right| > \frac{1}{2}$$

$$\left| \prod_{\substack{r_y > \frac{r}{2} \\ |r - r_y| > [\log^2 r]}} F_y \right| < \prod_{l = [\log^2 r]}^{l=r} \left(1 + \frac{6 + 8 \frac{\sigma}{r}}{(l + 2\sigma - 1)^2} \right)^{c \cdot \log r}$$

$$= \prod_{l = [\log^2 r]}^{l=r} \left(1 + \frac{6 + 8 \frac{\sigma}{r}}{(l + 2\sigma - 1)^2} \right)^{(l + 2\sigma - 1)^2 \cdot \frac{c \cdot \log r}{(l + 2\sigma - 1)^2}}$$

$$< \left(e^{6 + \frac{8\sigma}{r}} \right)^{c \cdot \log r \cdot \sum_{l = [\log^2 r]}^{l=r} \frac{1}{(l + 2\sigma - 1)^2}} < e^{\log r \cdot \frac{c}{[\log^2 r] + 2\sigma - 1}} = e^{\frac{c}{\log r}}$$

Sind also in dem Intervall $(r - \log^2 r, r + \log^2 r)$ keine Nullstellen der ζ -Funktion (mit der Abszisse $\geq \frac{1}{2} + \kappa$), so können wir für solche r bei Zusammenfassung der bisherigen Ergebnisse schließen

$$\left| \frac{\zeta}{\zeta^*} \right| < e^{\frac{1}{r^{2k}}} \cdot e^{\frac{c}{2r}} \cdot e^{\frac{c}{\log r}} \rightarrow 1$$

für $r \rightarrow \infty$.

Für eben diese r können wir wörtlich ebenso schließen

$$\left| \frac{\zeta}{\zeta^*} \right| > e^{-\frac{1}{r^{2k}}} \cdot e^{-\frac{c}{2r}} \cdot e^{-\frac{c}{\log r}} \rightarrow 1,$$

indem bei den Abschätzungen in jedem Faktor $1 - \dots$ statt $1 + \dots$ geschrieben wird. Es gilt also für diese r

$$\left| \frac{\zeta}{\zeta^*} \right| \rightarrow 1.$$

Sind dagegen in dem Intervall $(r - \log^2 r, r + \log^2 r)$ noch Nullstellen, so ergibt sich für die dadurch entstehenden Faktoren nur folgende Abschätzung

$$\begin{aligned}
\prod_{l=0}^{[\log^2 r]} \left(1 + \frac{6 + 8 \frac{\sigma}{r}}{(l + 2\sigma - 1)^2} \right)^{e \cdot \log r} &< e^{e \cdot \log r \left(6 + 8 \frac{\sigma}{r} \right) \sum_{l=0}^{[\log^2 r]} \frac{1}{(l + 2\sigma - 1)^2}} \\
&< e^{\left(6 + 8 \frac{\sigma}{r} \right) e \cdot \log r \left(\frac{1}{(2\sigma - 1)^2} + \sum_{l=1}^{\infty} \frac{1}{(l + 2\sigma - 1)(l + 2\sigma - 2)} \right)} \\
&= e^{\left(6 + 8 \frac{\sigma}{r} \right) e \cdot \log r \left(\frac{1}{(2\sigma - 1)^2} + \frac{1}{2\sigma - 1} \right)} = r^{\frac{2\sigma}{(2\sigma - 1)^2} c_1} \cdot e^{c_2 \frac{\sigma^2}{(2\sigma - 1)^2} \cdot \frac{\log r}{r}} \\
&< r^{c_1 \frac{2\sigma}{(2\sigma - 1)^2}} \cdot r^{c_3} < r^{\frac{c_4}{2\sigma - 1}},
\end{aligned}$$

wobei c_1 , c_3 und c_4 auch von σ unabhängig sind.

Es läßt sich also nur schließen

$$\left| \frac{\zeta(s)}{\zeta^*(s)} \right| < r^{\frac{c}{2\sigma - 1}}.$$

Für $\sigma > 3$ läßt sich ebenso schließen

$$\left| \frac{\zeta(s)}{\zeta^*(s)} \right| > r^{-\frac{c}{2\sigma - 1}},$$

indem bei der letzten Abschätzung $1 - \dots$ statt $1 + \dots$ geschrieben wird.

Wir wollen jetzt wieder annehmen, daß das Intervall $(r - \log^2 r, r + \log^2 r)$ nullstellenfrei ist. Es war bewiesen

$$\left| \frac{\zeta}{\zeta^*} \right| \sim 1.$$

Wir wollen jetzt sogar zeigen

$$\frac{\zeta(s)}{\zeta^*(s)} \rightarrow 1.$$

Das Argument von $\frac{\zeta(s)}{\zeta^*(s)}$ ist also zu untersuchen. Nun ist für ein Produkt

$$\begin{aligned}
P &= \prod_r^{\infty} (1 + \alpha_r) \\
\text{tg}(\arg P) &= \sum_{r=1}^{\infty} \frac{\Im(\alpha_r)}{1 + \Re(\alpha_r)}.
\end{aligned}$$

Für die drei Klassen unseres Produktes ergeben sich auch hier wieder verschiedene Abschätzung der Summanden.

In der Klasse 1: $\frac{r_\nu}{r} \leq \frac{1}{2}$ ergibt sich

$$\left| \frac{\Im(G_\nu)}{1 + \Re(G_\nu)} \right| < \frac{|G_\nu|}{1 - |G_\nu|} < \frac{\frac{1}{r}}{1 - \frac{1}{r}}$$

unter Benutzung früherer Abschätzungen.

Die Anzahl der Summanden ist kleiner als $r^{1-2\kappa}$, mithin

$$\sum_{\substack{r_\nu \leq 1 \\ r \leq \frac{1}{2}}} |\operatorname{tg}(\arg F_\nu)| < \frac{2}{r^{2\kappa}}.$$

In der Klasse 3: $\left(\frac{r_\nu}{r} \geq 2\right)$ muß man sofort den aus dem zweiten Produkt für $\frac{\zeta}{\zeta^*}$ entspringenden analogen Summanden hinzunehmen:

$$e^{2i\delta_\nu} F_{\nu,1} = 1 + H_{\nu,1}; \quad e^{-2i\delta_\nu} F_{\nu,2} = 1 + H_{\nu,2},$$

wo $H_{\nu,2}$ aus $H_{\nu,1}$ entsteht durch Ersetzung von δ_ν durch $-\delta_\nu$.

Wegen der Abschätzung

$$|H_\nu| < \frac{1}{2rr_\nu}$$

folgt

$$\sum_{\substack{r_\nu \geq 2 \\ r}} |\operatorname{tg}(\arg F_{\nu,1}) + \operatorname{tg}(\arg F_{\nu,2})| \leq \frac{1}{2r} \sum \frac{1}{r_\nu} < \frac{c'}{r}.$$

Bei den Summanden der zweiten Klasse haben wir die Abschätzung

$$\left| \frac{\Im(L_\nu + M_\nu)}{1 + \Re(L_\nu + M_\nu)} \right| < \frac{|L_\nu| + |M_\nu|}{1 - |L_\nu| - |M_\nu|},$$

sofern $|L_\nu| + |M_\nu| < 1$ ist. Nun ist

$$|L_\nu| < \frac{2}{r_\nu(2\sigma-1)}; \quad |M_\nu| < \frac{2}{(|r-r_\nu|+2\sigma-1)^2}, \quad |r-r_\nu| = l \geq \log^2 r,$$

$$\sum_{\frac{1}{2} < \frac{r_\nu}{r} < 2} |\operatorname{tg}(\arg F_{\nu,1} + \arg F_{\nu,2})| < \sum \frac{2}{\frac{r_\nu(2\sigma-1)}{\frac{1}{2}}} + \sum \frac{2}{\frac{(|r-r_\nu|+2\sigma-1)^2}{\frac{1}{2}}},$$

da offenbar $1 - |L_\nu| - |M_\nu| > \frac{1}{2}$ angenommen werden kann.

Die Anzahl der Summanden ist kleiner als $r^{1-2\kappa}$. Also

$$\sum_{\frac{1}{2} < \frac{r_\nu}{r} < 2} \frac{1}{r_\nu} < \frac{2r^{1-2\kappa}}{r} = \frac{2}{r^{2\kappa}},$$

$$\sum_{\frac{1}{2} < \frac{r_\nu}{r} < 2} \frac{1}{(l+2\sigma-1)^2} \leq \sum_{l=\log^2 r}^r \frac{c \log r}{(l+2\sigma-1)^2} < \frac{c}{\log r}.$$

Insgesamt ergibt sich also: Für alle r , bei denen das Intervall $(r - \log^2 r, r + \log^2 r)$ nullstellenfrei ist, gilt mit $r \rightarrow \infty$

$$\arg \frac{\zeta(s)}{\zeta^*(s)} \rightarrow 0.$$

In der Tat: mit

$$\sum \operatorname{tg} |\alpha_n|$$

ist auch $\sum |\alpha_n|$ konvergent, falls nur $|\alpha_n| < \frac{\pi}{2}$ gewählt ist. Wegen $\operatorname{tg} \alpha_n = \operatorname{tg} |\alpha_n|$ läßt sich also auf die Konvergenz von

$$\sum |\arg F_n|$$

schließen. Da außerdem

$$\sum \operatorname{tg} |\alpha_n| < \sum |\alpha_n| < 2 \sum \operatorname{tg} |\alpha_n|$$

gilt, wofern nur etwa $|\alpha_n| < \frac{\pi}{4}$ ist, so ergibt sich

$$\sum |\arg F_n| \rightarrow 0$$

mit $r \rightarrow \infty$.

Wir können jetzt sehr leicht den Nachweis für die Eigenschaften 3 und 4 der Funktion $\zeta^*(s)$ erbringen.

Es sei $s = \sigma_0 + it$ ($\sigma_0 > \frac{1}{2} + \varkappa$ fest).

Wir können dann in unseren Abschätzungen ruhig r durch t ersetzen, da für große t die Differenz $|t - r|$ beschränkt bleibt. Wir behaupten

$$\lim \frac{1}{T} \int_0^T |\zeta(\sigma_0 + it) - \zeta^*(\sigma_0 + it)|^2 dt \rightarrow 0$$

falls σ_0 nicht zu klein gewählt ist: $\sigma_0 > 2 + \frac{c_1}{2k}$.

Um jede Nullstelle $\sigma + it$ von $\zeta(s)$ mit $\sigma \geq \frac{1}{2} + \varkappa$ denken wir uns das Intervall $(t - \log^2 t, t + \log^2 t)$ gelegt. Die Länge aller solcher Intervalle zwischen 1 und T ist nach Carlson höchstens

$$T^{1-4\varkappa^2+0(1)} \cdot \log^2 T < T^{1-2\varkappa},$$

also verschwindend gering gegenüber T . Für alle t , die nicht in solche Intervalle fallen, gilt offenbar

$$\frac{\zeta^*(s)}{\zeta(s)} = 1 + \varepsilon(t),$$

wobei $\varepsilon(t)$ mit $\frac{1}{t}$ gegen Null geht.

Deute ich durch den Akzent an, daß ich nur über solche t integrieren will, so ist

$$\begin{aligned}
& \int_0^T |\zeta(\sigma_0 + it) - \zeta^*(\sigma_0 + it)|^2 dt \\
&= \int_0^{\sqrt{T}} |\zeta(\sigma_0 + it)|^2 \left| 1 - \frac{\zeta^*(\sigma_0 + it)}{\zeta(\sigma_0 + it)} \right|^2 dt + \int_{\sqrt{T}}^T |\zeta(\sigma_0 + it)|^2 \cdot \varepsilon_1(t) dt \\
&\leq C \cdot \int_0^{\sqrt{T}} |\zeta(\sigma_0 + it)|^2 dt + \varepsilon \int_{\sqrt{T}}^T |\zeta(\sigma_0 + it)|^2 dt,
\end{aligned}$$

wobei $\varepsilon = \text{Max } \varepsilon_1(t)$ in (\sqrt{T}, T) bezeichnet.

Da $\sigma_0 > \frac{1}{2} + \varkappa$ ist, gilt der Mittelwertsatz. Es ist also

$$\begin{aligned}
& \frac{1}{T} \int_0^T |\zeta(\sigma_0 + it) - \zeta^*(\sigma_0 + it)|^2 dt \\
&< \frac{C}{\sqrt{T}} \cdot \frac{1}{\sqrt{T}} \int_0^{\sqrt{T}} |\zeta(\sigma_0 + it)|^2 dt + \varepsilon \cdot \frac{1}{T} \int_0^T |\zeta(\sigma_0 + it)|^2 dt < \frac{C_1}{\sqrt{T}} + \varepsilon C_2.
\end{aligned}$$

Es bleibt noch die Integration über alle die t , welche in ein Intervall fallen. Wir deuten das durch zwei Akzente an. Es sei jetzt $\sigma_0 > 3$. Dann gilt, wie wir schon sahen,

$$\left| \frac{\zeta^*(s)}{\zeta(s)} \right| < t^{\frac{c_1}{2\sigma_0 - 1}}.$$

Es ist demnach

$$\begin{aligned}
\int_0^T |\zeta(\sigma_0 + it) - \zeta^*(\sigma_0 + it)|^2 dt &\leq \int_0^T |\zeta(\sigma_0 + it)|^2 \cdot t^{\frac{2c_1}{2\sigma_0 - 1}} dt \\
&\leq C \cdot T^{\frac{2c_1}{2\sigma_0 - 1}} \int_0^T dt < C \cdot T^{1 - 2\varkappa + \frac{2c_1}{2\sigma_0 - 1}}.
\end{aligned}$$

Wenn also $2\sigma_0 > 1 + \frac{c_1}{\varkappa}$ ist, dann ergibt sich

$$\frac{1}{T} \int_0^T |\zeta(\sigma_0 + it) - \zeta^*(\sigma_0 + it)|^2 dt \rightarrow 0.$$

Also insgesamt:

$$\frac{1}{T} \int_0^T |\zeta(\sigma_0 + it) - \zeta^*(\sigma_0 + it)|^2 dt \rightarrow 0.$$

In ganz ähnlicher Weise läßt sich Eigenschaft 3 herleiten. Man kann diese Eigenschaft aber auch, wenn man will, aus der eben bewiesenen Relation ableiten, so wie es Carlson in einem allgemeineren Falle getan hat (Arkiv 16, n° 18, p. 12).

Eine kleine Anwendung wollen wir noch geben. Es sei

$$N\left(\frac{1}{2} + \kappa, T + 1\right) - N\left(\frac{1}{2} + \kappa, T\right) = o(\log t) = \delta(t) \log t.$$

Dann folgt nach unserer Abschätzung für die Faktoren der zweiten Klasse

$$\left| \prod_{\substack{\frac{1}{2} < \frac{r_\nu}{r} < 2 \\ |r - r_\nu| < \log^2 r}} F_\nu \right| < \prod \left(1 + \frac{6 + 8 \frac{\sigma}{r}}{(l + 2\sigma - 1)^2} \right)^{\delta(r) \log r} < r^{\frac{\sigma}{2\sigma - 1} \cdot \delta(r)}.$$

Es gilt also für alle $s = \sigma + it$ die Abschätzung

$$\left| \frac{\zeta(\sigma + it)}{\zeta^*(\sigma + it)} \right| < t^{0(1)}$$

für $\frac{1}{2} + \kappa < \sigma < C$, wo C beliebig groß sein kann.

Für $\sigma > 1,01$ folgt entsprechend

$$\left| \frac{\zeta^*(\sigma + it)}{\zeta(\sigma + it)} \right| < t^{0(1)},$$

also

$$|\zeta^*(\sigma + it)| < t^{0(1)}$$

für $\sigma > 1,01$. Wegen der Funktionalgleichung ist

$$|\zeta^*(\sigma + it)| < t^\epsilon$$

für $-C < \sigma < -0,01$. Daraus läßt sich, da ja die ganze Funktion $\zeta^*(s) \cdot (1-s)$ ebenso wie $\zeta(s) \cdot (1-s)$ von der Ordnung 1 (Maximaltyp) ist und daher

$$|\zeta^*(s) \cdot (1-s)| < e^{|s|^{1+\epsilon}}$$

für alle s gilt, nach Phragmén-Lindelöf der Schluß ziehen:

$$|\zeta^*(s) \cdot |s|^{-\sigma}|$$

ist beschränkt in $-0,01 < \sigma < 1,01$. Also ist für $\sigma > \frac{1}{2} + \kappa$ nach Carathéodory sogar

$$\log \zeta^*(s) = O(\log t)$$

im Kreise um $2 + it$ mit $\frac{3}{2} - \kappa - \epsilon$, wenn derjenige Zweig verstanden wird, für den in $2 + it$ das Argument von $\zeta^*(s)$ zwischen 0 und 2π liegt. Wenden wir auf diese Funktion den Dreikreisesatz an:

$$r_1 = 0,9; \quad r_3 = \frac{3}{2} - \kappa - \epsilon; \quad r_2 = \frac{3}{2} - \kappa - 2\epsilon,$$

wobei ε eine feste kleine Zahl bedeutet. Dann ist

$$M_2 \leq M_1^{\alpha_1} \cdot M_3^{\alpha_2} \quad (\alpha_1 + \alpha_2 = 1),$$

$$M_2 \leq \delta(t)^{\alpha_1} \cdot c^{\alpha_2} \cdot (\log t)^{\alpha_1 + \alpha_2} = \delta_1(t) \cdot \log t.$$

Also ist für $\sigma > \frac{1}{2} + \varkappa$

$$\log \zeta^*(\sigma + it) = o(\log t).$$

Daraus folgt

$$|\zeta(\sigma + it)| < t^{o(1)}.$$

Also: Falls

$$N\left(\frac{1}{2} + \varkappa, T + 1\right) - N\left(\frac{1}{2} + \varkappa, T\right) = o(\log T)$$

gilt, so ist die Lindelöfsche Vermutung

$$|\zeta(s)| = t^{o(1)}$$

für $\sigma > \varkappa$ richtig.

Es gilt aber, wie man leicht mit Hilfe der Jensenschen Formel zeigen kann, auch die Umkehrung:

Die Lindelöfsche Vermutung führt zu

$$N\left(\frac{1}{2} + \varkappa, T + 1\right) - N\left(\frac{1}{2} + \varkappa, T\right) = o(\log T).$$

Diese notwendige und hinreichende Bedingung für die Lindelöfsche Vermutung hat schon Backlund erkannt [Öfversigt Finska Vetensk. Soc. 61 (1918)].

(Eingegangen am 3. 2. 1927.)

Zur Theorie des arithmetisch-geometrischen Mittels.

Von

Harald Geppert in Gießen.

Den interessantesten und merkwürdigsten Zugang zur Modulfunktion bildet wohl die auf einem algebraischen Iterationsprozeß beruhende Theorie des arithmetisch-geometrischen Mittels (agM). Sie gab auch die Grundlage ab, auf der Gauß als erster zur Modulfunktion vordrang und ihre wesentlichsten Eigenschaften entdeckte¹⁾. In der Folgezeit ist jedoch dieser außerordentlich eleganten Theorie verhältnismäßig wenig Beachtung zuteil geworden, und erst in den letzten Dezennien haben die Gaußschen Gedankengänge durch die Arbeiten der Herren Schlesinger und v. Dávid manche wertvolle Bereicherung erfahren²⁾, durch die es möglich geworden ist, die Theorie der Modulfunktion einheitlich und konsequent lediglich auf Grund des algebraischen Algorithmus aufzubauen. Ein solcher Aufbau ist nicht nur an sich des Interesses wert, sondern auch deswegen von Bedeutung, weil sich einerseits allgemeiner eine weite Klasse automorpher Funktionen durch algebraische Iterationsprozesse gewinnen läßt³⁾ und die Modulfunktion der einfachste Repräsentant dieser Klasse ist, andererseits von den Herren Schapira und v. Dávid⁴⁾ eine äußerst interessante Verallgemeinerung des agM-Algorithmus entwickelt worden ist, die sich auf Iterationsprozesse von höherem als zweitem Grade bezieht. Herr v. Dávid hat es in einer eben erscheinenden Arbeit unternommen, diesen Zugang zur Modulfunktion in systematisch konsequenter Fassung darzustellen⁵⁾; er dringt daselbst bis zur Uniformisierung des Algorithmus und dem Automorphismus der Modulfunktion vor.

¹⁾ Vgl. Gauß 1. 2. 3. 4.; Schlesinger 3. 4. [Die Zahlen beziehen sich auf den Literaturnachweis am Schluß der Arbeit.]

²⁾ Vgl. Schlesinger 1., v. Dávid 1. 2. 3. 4.

³⁾ Vgl. Schlesinger 2. S. 275 ff.

⁴⁾ Vgl. hierzu: v. Dávid, Crelle **135** (1909), S. 62–74; **140** (1911), S. 277–296.

⁵⁾ v. Dávid 4.

Die vorliegende Arbeit soll nun diese v. Dávidsche Arbeit in gewissem Sinne nach der geometrischen Seite hin ergänzen, indem sie das Ziel verfolgt, den Zusammenhang der unendlich vielen möglichen Algorithmen des agM mit ihrer Uniformisierung aufzudecken und diese Verknüpfung in der Ebene der uniformisierenden Variablen geometrisch zu deuten, was uns zu neuen, eleganten Ergebnissen führen wird. Wir beginnen mit einer kurzen zusammenfassenden Darstellung größtenteils bekannter Resultate.

§ 1.

Der Algorithmus und seine Uniformisierung.

Es seien a, b zwei beliebige komplexe Zahlen, die von Null verschieden und nicht entgegengesetzt gleich sind; wir bilden mit ihnen die drei Zahlenfolgen:

$$(1) \quad a_\nu = \frac{1}{2}(a_{\nu-1} + b_{\nu-1}); \quad b_\nu = \sqrt{a_{\nu-1} b_{\nu-1}}; \quad c_\nu = \sqrt{a_\nu^2 - b_\nu^2},$$

$$(\nu = 1, 2, 3, \dots)$$

$$a_0 = a, \quad b_0 = b, \quad c_0 = c,$$

dann konvergieren die Folgen der a_ν, b_ν gegen einen gemeinsamen Grenzwert

$$\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} b_n = \mathfrak{M}(a, b),$$

das arithmetisch-geometrische Mittel (agM) von a, b , und demgemäß strebt die Folge der c_ν gegen Null. Wegen der Homogenität des Algorithmus kommt es bei der Untersuchung des agM nur auf die Verhältnisse

$$\frac{b}{a} = k', \quad \frac{c}{a} = k, \quad k^2 + k'^2 = 1$$

an, und wir bezeichnen daher allgemein:

$$\frac{b_\nu}{a_\nu} = k'_\nu, \quad \frac{c_\nu}{a_\nu} = k_\nu, \quad k_\nu^2 + k'_\nu{}^2 = 1.$$

Der Algorithmus ist entsprechend der Doppeldeutigkeit der auftretenden Quadratwurzeln unendlich vieldeutig und führt zu abzählbar unendlich vielen von Null verschiedenen Werten $\mathfrak{M}(a, b)$. Einer unter diesen unendlich vielen Algorithmen läßt sich als *Normalalgorithmus* auffassen, nämlich derjenige, für den die Wurzelzeichen gemäß der Bedingung

$$(2) \quad \Re\left(\frac{b_\nu}{a_{\nu-1}}\right) \geq 0 \quad (\nu = 1, 2, 3, \dots)$$

oder der damit äquivalenten Forderung

$$(2') \quad \Re\left(\frac{b_\nu}{a_\nu}\right) = \Re(k'_\nu) \geq 0 \quad (\nu = 1, 2, 3, \dots)$$

gewählt werden; er konvergiert zum sog. „einfachsten“ Mittel $M(a, b)$, das das absolut größte aller Mittel $\mathfrak{M}(a, b)$ ist. Mit Ausnahme des Falles, daß $\arg a - \arg b = \pm \pi$ ist, gibt es nur ein einfachstes Mittel, im genannten Falle hingegen zwei zur Geraden ab symmetrisch liegende einfachste Mittel⁶⁾.

Der Algorithmus des agM wird *uniformisiert* durch Einführung der Nullwerte der Jacobischen Thetafunktionen:

$$\vartheta_{00}(\omega) = 1 + 2 \sum_{n=1}^{\infty} q^{n^2}, \quad \vartheta_{01}(\omega) = 1 + 2 \sum_{n=1}^{\infty} (-1)^n q^{n^2},$$

$$\vartheta_{10}(\omega) = 2 \sum_{n=0}^{\infty} q^{(n+\frac{1}{2})^2}, \quad q = e^{\pi i \omega},$$

für die im nachstehenden die folgenden Identitäten Verwendung finden:

$$(3) \quad \vartheta_{00}(\omega) \vartheta_{01}(\omega) = \vartheta_{01}(2\omega)^2, \quad 2\vartheta_{00}(2\omega) \vartheta_{10}(2\omega) = \vartheta_{10}(\omega)^2,$$

$$(3') \quad \vartheta_{00}(\omega)^2 + \vartheta_{01}(\omega)^2 = 2\vartheta_{00}(2\omega)^2, \quad \vartheta_{00}(\omega)^2 - \vartheta_{01}(\omega)^2 = 2\vartheta_{10}(2\omega)^2,$$

$$(4) \quad \vartheta_{00}(\omega) = \sqrt{\frac{i}{\omega}} \vartheta_{00}\left(-\frac{1}{\omega}\right), \quad \vartheta_{01}(\omega) = \sqrt{\frac{i}{\omega}} \vartheta_{10}\left(-\frac{1}{\omega}\right),$$

$$\vartheta_{10}(\omega) = \sqrt{\frac{i}{\omega}} \vartheta_{01}\left(-\frac{1}{\omega}\right),$$

$$(4') \quad \vartheta_{00}(\omega) = \vartheta_{01}(\omega+1), \quad \vartheta_{01}(\omega) = \vartheta_{00}(\omega+1), \quad \vartheta_{10}(\omega) = i^{-\frac{1}{2}} \vartheta_{10}(\omega+1).$$

Es ist wichtig zu bemerken, daß diese Funktionen nicht als deus ex machina in die Theorie eintreten, sondern aus dem Algorithmus (1) heraus zwangsläufig gewonnen werden können⁷⁾. Die angestrebte Uniformisierung gestaltet sich nun folgendermaßen: Zu jedem vorgelegten Wertepaare a, b läßt sich ein in der oberen Halbebene gelegener Wert von ω angeben, so daß die Proportion gilt:

$$(5) \quad a : b : c = \vartheta_{00}(\omega)^2 : \vartheta_{01}(\omega)^2 : \vartheta_{10}(\omega)^2,$$

die wir auch mittels eines Proportionalitätsfaktors $\mu(\omega)$ schreiben können:

$$(5') \quad a = \mu(\omega) \vartheta_{00}(\omega)^2, \quad b = \mu(\omega) \vartheta_{01}(\omega)^2, \quad c = \mu(\omega) \vartheta_{10}(\omega)^2.$$

Aus den Identitäten (3), (3') folgt dann für die übrigen Glieder des Algorithmus (1):

$$(5'') \quad a_\nu = \mu(\omega) \vartheta_{00}(2^\nu \omega)^2, \quad b_\nu = \mu(\omega) \vartheta_{01}(2^\nu \omega)^2, \quad c_\nu = \mu(\omega) \vartheta_{10}(2^\nu \omega)^2,$$

und mithin für den Proportionalitätsfaktor $\mu(\omega)$ die Bestimmung:

$$(6) \quad \mu(\omega) = \lim_{\nu \rightarrow \infty} a_\nu = \lim_{\nu \rightarrow \infty} b_\nu = \mathfrak{M}(a, b).$$

⁶⁾ v. Dávid 3.

⁷⁾ Vgl. Schlesinger 1., S. 354 ff.; v. Dávid 4.

Zu einem festen ω gehört vermittels (5'), (5'') ein *bestimmter* der Algorithmen (1) und ein *bestimmtes* Mittel $\mathfrak{M}(a, b)$. Man erhält die unendlich vielen Algorithmen und die zugehörigen Mittel, wenn man beachtet, daß ω durch die Proportion (5) nicht eindeutig bestimmt ist, sondern daß dieser Proportion durch alle Werte ω' genügt wird, die aus ω durch die Substitutionen mit ganzzahligen Koeffizienten:

$$(7) \quad \omega' = \frac{\alpha\omega + \beta}{\gamma\omega + \delta},$$

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \pmod{4}, \quad \alpha\delta - \beta\gamma = 1$$

hervorgehen.

Diese Verknüpfung der möglichen Algorithmen mit den zugehörigen Werten der uniformisierenden Variablen ω eingehender zu studieren und aufzuklären ist das Ziel vorliegender Abhandlung.

Für positiv reelle a, b, c ist ein Wert von ω gegeben durch den Quotienten⁸⁾

$$(8) \quad \omega = i \frac{M(a, b)}{M(a, c)} = i \frac{M(1, k')}{M(1, k)},$$

wobei $M(a, b)$ und $M(a, c)$ als einfachste Mittel gemäß (2') mit *positiven* Quadratwurzelzeichen zu berechnen sind, und zwar führt der Wert (8) in (5'') eingesetzt zum Mittel $M(a, b)$. Nach dem Prinzip der analytischen Fortsetzung liefert die Bestimmung (8) von ω auch für komplexe a, b, c , wenigstens solange $\Re(k) > 0$, $\Re(k') > 0$ ist, in (5'') eingesetzt das einfachste Mittel $M(a, b)$, und allgemein gehört zum Proportionalitätsfaktor

$$\mu(\omega') = \mathfrak{M}(a, b)$$

die uniformisierende Variable

$$\omega' = i \frac{\mathfrak{M}(a, b)}{\mathfrak{M}(a, c)},$$

wo $\mathfrak{M}(a, b)$ das gleiche Mittel wie in $\mu(\omega')$ bedeutet, $\mathfrak{M}(a, c)$ hingegen durch $\mathfrak{M}(a, b)$ in gewisser, gleich zu besprechender Weise bestimmt ist und daher als mit $\mathfrak{M}(a, b)$ *zusammengehörig* bezeichnet wird⁹⁾. Nach (7) und (8) gilt daher für zwei zusammengehörige Mittel die Beziehung:

$$(9) \quad \frac{\mathfrak{M}(a, b)}{\mathfrak{M}(a, c)} = \frac{\alpha M(a, b) - \beta i M(a, c)}{\gamma i M(a, b) + \delta M(a, c)}; \quad \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \pmod{4};$$

$$\alpha\delta - \beta\gamma = 1.$$

Andererseits bestehen bei der Transformation (7) die identischen Gleichungen

$$\vartheta_{00}(\omega')^2 = (\gamma\omega + \delta)\vartheta_{00}(\omega)^2, \quad \vartheta_{01}(\omega')^2 = (\gamma\omega + \delta)\vartheta_{01}(\omega)^2,$$

$$\vartheta_{10}(\omega')^2 = (\gamma\omega + \delta)\vartheta_{10}(\omega)^2,$$

⁸⁾ Vgl. Gauß 3., S. 197; 4. § 25.

⁹⁾ Vgl. Schlesinger in Gauß 3., S. 281.

und daher gibt der Vergleich von (5') mit den entsprechenden Relationen:

$$(5''') \quad a = \mu(\omega') \vartheta_{00}(\omega')^2 = \mathfrak{M}(a, b) \left(\gamma i \frac{M(a, b)}{M(a, c)} + \delta \right) \vartheta_{00}(\omega)^2 \\ = M(a, b) \vartheta_{00}(\omega)^2; \dots$$

die Beziehungen

$$(10) \quad \frac{1}{\mathfrak{M}(a, b)} = \frac{\gamma i}{M(a, c)} + \frac{\delta}{M(a, b)}, \\ \frac{1}{\mathfrak{M}(a, c)} = \frac{-i\omega'}{\mathfrak{M}(a, b)} = \frac{-i(\alpha\omega + \beta)}{M(a, b)} = \frac{\alpha}{M(a, c)} - \frac{\beta i}{M(a, b)},$$

wobei $\alpha, \beta, \gamma, \delta$ durch die Kongruenz- und Determinantenbedingung (7) aneinander gebunden sind.

Diese wichtigen Gleichungen (10) vermitteln den Zusammenhang der unendlich vielen Mittel $\mathfrak{M}(a, b)$ untereinander und erklären andererseits den Begriff der zusammengehörigen Mittel. Soll nämlich zu einem Mittel $\mathfrak{M}(a, b)$ die uniformisierende Variable, also das zusammengehörige Mittel $\mathfrak{M}(a, c)$ bestimmt werden, so kommt dies darauf hinaus, zu gegebenen ganzzahligen γ, δ die ganzzahligen Größen α, β gemäß der Kongruenzbedingung

$$\alpha \equiv 1, \quad \beta \equiv 0 \pmod{4}$$

und der Gleichung

$$(7') \quad \alpha \delta - \beta \gamma = 1$$

zu errechnen. Durch (7') sind α, β nicht eindeutig bestimmt, sondern mit α, β sind auch

$$\bar{\alpha} = \alpha + 4m\gamma, \quad \bar{\beta} = \beta + 4m\delta \quad (m = \text{ganze Zahl})$$

Lösungen von (7'), d. h. ist $\mathfrak{M}(a, c)$ ein mit $\mathfrak{M}(a, b)$ zusammengehöriges Mittel, so ist auch $\bar{\mathfrak{M}}(a, c)$ ein solches, wobei

$$(11) \quad \frac{1}{\bar{\mathfrak{M}}(a, c)} = \frac{1}{\mathfrak{M}(a, c)} - \frac{4mi}{\mathfrak{M}(a, b)} \quad ^{9a)}$$

ist. Die mit $\mathfrak{M}(a, b)$ zusammengehörigen Werte $\bar{\mathfrak{M}}(a, c)$ liegen also alle auf einem durch den Nullpunkt gehenden Kreise, auf dem sie sich gegen den Nullpunkt hin häufen. Wegen

$$\frac{\bar{\mathfrak{M}}(a, b)}{\bar{\mathfrak{M}}(a, c)} = \frac{\mathfrak{M}(a, b)}{\mathfrak{M}(a, c)} - 4mi$$

ist die Tatsache, daß zu $\mathfrak{M}(a, b)$ unendlich viele durch (11) verbundene Werte $\bar{\mathfrak{M}}(a, c)$ gehören, identisch damit, daß durch $\mathfrak{M}(a, b)$ die zugehörige uniformisierende Variable nur bis auf das Vierfache einer ganzen Zahl

^{9a)} Diese Gleichung findet sich auch im Nachlaß von Gauß; vgl. Gauß 3., S. 219 [1].

festgelegt ist, wie es sein muß, da bei Ersetzung von ω durch $\omega + 4m$ der Algorithmus (5'), (5'') ungeändert bleibt.

Nach dem Gesagten ist zu jedem von Null verschiedenen $\mathfrak{M}(a, b)$, d. h. zu jedem regulären Algorithmus, eine uniformisierende Variable, oder um uns kürzer auszudrücken, ein Kanon¹⁰⁾ ω bestimmbar, nämlich

$$(8') \quad \omega = i \frac{\mathfrak{M}(a, b)}{\mathfrak{M}(a, c)} = i \frac{\mathfrak{M}(1, k')}{\mathfrak{M}(1, k)}.$$

Dann ist der Kanon des mit $\mathfrak{M}(a, b)$ zusammengehörigen Mittels $\mathfrak{M}(a, c)$ der Wert $-\frac{1}{\omega}$, den wir als den adjungierten Kanon bezeichnen. Nach (4) ist nämlich

$$\vartheta_{00} \left(-\frac{1}{\omega}\right)^2 : \vartheta_{01} \left(-\frac{1}{\omega}\right)^2 : \vartheta_{10} \left(-\frac{1}{\omega}\right)^2 = \vartheta_{00}(\omega)^2 : \vartheta_{10}(\omega)^2 : \vartheta_{01}(\omega)^2 = a : c : b,$$

$$a = \mu \left(-\frac{1}{\omega}\right) \vartheta_{00} \left(-\frac{1}{\omega}\right)^2 = \mu \left(-\frac{1}{\omega}\right) \frac{\omega}{i} \vartheta_{00}(\omega)^2 = \mathfrak{M}(a, b) \vartheta_{00}(\omega)^2,$$

also

$$\mu \left(-\frac{1}{\omega}\right) = \mathfrak{M}(a, c).$$

Natürlich ist auch $-\frac{1}{\omega} + 4m$ ein zu ω adjungierter Kanon.

§ 2.

Der Fundamentalbereich. Der Bereich des einfachsten Mittels.

Einem bestimmten Wert von ω entspricht vermittels (5'') ein bestimmter agM-Algorithmus. Es ist daher zu erwarten, daß den möglichen Algorithmen Teilbereiche der ω -Ebene zugeordnet sind, in denen sie geometrisch gedeutet werden können. Zu jedem Wertetripel a, b, c läßt sich ein Kanon ω bestimmen, der in dem Fundamentalbereich Φ der Kongruenzgruppe Γ^4

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \pmod{4}$$

gelegten ist. Dieser läßt sich aus dem vierzipfeligen Fundamentalbereich F mit den Ecken in $-1, 0, +1, \infty i$ der zweistufigen Kongruenzgruppe Γ^2

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \pmod{2}$$

mit Leichtigkeit ableiten. Denn Γ^4 ist eine Untergruppe von Γ^2 und die zugehörigen Nebengruppen sind

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \pmod{4},$$

¹⁰⁾ Diesen Ausdruck hat Gauß geprägt, vgl. Gauß 1., S. 477f.

sie können aus Γ^4 dadurch erzeugt werden, daß man Γ^4 mit einer der Substitutionen

$$\Sigma: \begin{pmatrix} 1 & 0 \\ \pm 2 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & \pm 2 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} -3 & \mp 2 \\ \pm 2 & 1 \end{pmatrix}$$

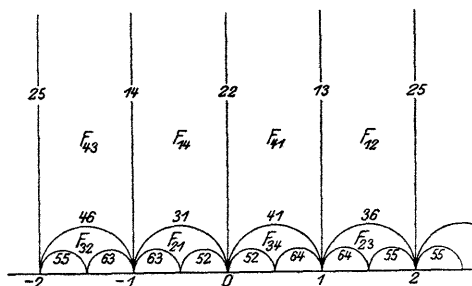


Fig. 1.

von rechts oder links her komponiert. Zerlegen wir also F durch die imaginäre Achse in die beiden Hälften F_{14} und F_{41} (vgl. Fig. 1), so erhält man eine mögliche Form des Fundamentalbereiches Φ von Γ^4 , indem man auf F_{14} die Substitutionen

$$\begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} -3 & -2 \\ 2 & 1 \end{pmatrix},$$

und auf F_{41} die Substitutionen

$$\begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} -3 & 2 \\ -2 & 1 \end{pmatrix}$$

von Σ anwendet und die so entstehenden Bereiche zusammenfügt. Dadurch geht der in der Figur 1 dargestellte Bereich Φ hervor, nämlich ein Kreisbogenzähneck mit den Ecken in ± 2 , $\pm \frac{3}{2}$, ± 1 , $\pm \frac{1}{2}$, 0 , $i\infty$. Φ ist der *Fundamentalbereich* der Gruppe Γ^4 .

Wir werden im nächsten Paragraphen den agM-Algorithmus in Φ untersuchen und wollen hier zunächst die folgende Frage beantworten: Wo muß ω liegen, damit der zugehörige Algorithmus (5'') das *einfachste* Mittel liefert? Ist ω ein zu $M(a, b)$ führender Kanon, so gilt für jedes andere Mittel $\mathfrak{M}(a, b)$ nach (5''')

$$\mathfrak{M}(a, b) = \frac{M(a, b)}{\gamma\omega + \delta}.$$

Nun ist $M(a, b)$ das absolut größte aller Mittel, also

$$|\mathfrak{M}(a, b)| = \frac{|M(a, b)|}{|\gamma\omega + \delta|} \leq |M(a, b)|, \quad \text{d. h.} \quad |\gamma\omega + \delta| \geq 1,$$

ω muß also so liegen, daß für alle der Kongruenzbedingung genügenden γ, δ die Größe $|\gamma\omega + \delta| \geq 1$ ist, d. h. außerhalb der um $-\frac{\delta}{\gamma}$, 0 als Mittelpunkte mit dem Radius $\frac{1}{|\gamma|}$ beschriebenen Kreise. Für $\gamma = \pm 4$ ergibt dies die Kreise mit dem Radius $\frac{1}{4}$ um die Punkte $\pm \frac{1}{4}$; $\pm \frac{3}{4}$; $\pm \frac{5}{4}$; $\pm \frac{7}{4}$, ... Die Kreise für absolut größere γ sind in diesen enthalten, wir finden somit als zulässigen Bereich für ω den Kreisbogenbereich mit den Spitzen in 0 , $\pm \frac{n}{2}$ ($n = 1, 2, 3, \dots$), d. h. den Fundamentalbereich Φ und die

aus ihm durch Verschiebung um das Vierfache ganzer Zahlen hervorgehenden äquivalenten Bereiche, so daß etwa das Bild einer spanischen Wand mit den genannten Ecken entsteht (s. Fig. 2). Daß ω in diesem Gebiete — wir wollen es E nennen — liegt, ist eine *notwendige* Bedingung dafür, daß der Algorithmus (5'') das einfachste

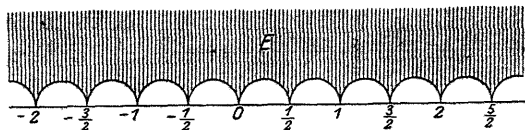


Fig. 2.

Mittel $M(a, b)$ liefert. Daß diese Bedingung auch *hinreichend* ist, werden wir im nächsten Paragraphen beweisen.

Wir sind nun auch imstande, die *Lagenbedingung* anzugeben, die ω erfüllen muß, wenn der adjungierte Kanon $-\frac{1}{\omega} + m$ das einfachste Mittel $M(a, c)$ liefern soll. Spiegelt man nämlich E am Einheitskreis und dann an der imaginären Achse, so erhält man den Bereich \bar{E} , in dem alle ω liegen, die durch die Transformation $-\frac{1}{\omega} + 4m$ nach E geworfen werden. Dieses Gebiet \bar{E} wird begrenzt von den Parallelen durch ± 2 zur imaginären Achse und den Kreisen mit den Mittelpunkten $\pm \frac{3}{2}, \pm \frac{5}{6}, \pm \frac{7}{12}, \dots, \pm \frac{2n+1}{n(n+1)}$ und den Radien $\frac{1}{2}, \frac{1}{6}, \frac{1}{12}, \dots, \frac{1}{n(n+1)}$;

es entsteht so die Figur 3. Liegt ω in \bar{E} , so liegt der adjungierte Kanon in E und liefert das einfachste Mittel $M(a, c)$.

Es genügt offenbar für das folgende, die Verhältnisse bei $M(a, b)$ allein zu studieren, da man durch passende Transformation zu $M(a, c)$ übergehen kann.

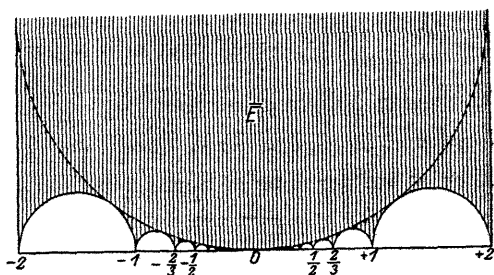


Fig. 3.

§ 3.

Der Algorithmus im Fundamentalbereich.

Wir gehen nun dazu über, das Verhalten des Algorithmus (5'') im Bereich Φ zu studieren; dazu zerfallen wir Φ nach der in der Figur 1 angedeuteten Weise in acht Teilgebiete $F_{14}, F_{41}, \dots, F_{28}$.¹¹⁾ Jedes derselben bilden wir durch die Gleichungen

¹¹⁾ Die Doppelindizes zeigen an, auf welchen Quadranten der k' - bzw. k -Ebene sich die entsprechenden Gebiete abbilden. Desgleichen charakterisieren die den Seiten beigefügten Doppelzahlen, auf welche Strecken der k' - (k -) Ebene (siehe Fig. 4) sich die betreffenden Seiten abbilden.

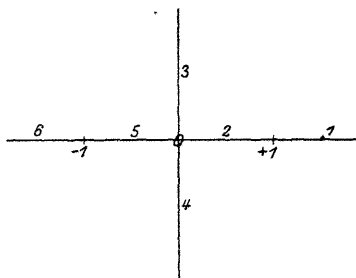


Fig. 4.

$$(12) \quad k' = \left[\frac{\vartheta_{01}(\omega)}{\vartheta_{00}(\omega)} \right]^2, \quad k = \left[\frac{\vartheta_{10}(\omega)}{\vartheta_{00}(\omega)} \right]^2$$

auf die k - bzw. k' -Ebene ab und verfolgen in ihm den Algorithmus (5'').

Untersuchen wir zunächst F_{14} und F_{41} und bezeichne $\bar{\omega}$ den stets positiven Imaginärteil von ω , so ist auf der linken Begrenzung von F_{14} und der rechten von F_{41} :

$$\omega = \mp 1 + \bar{\omega}i,$$

$$\vartheta_{00}(\omega) = \vartheta_{01}(i\bar{\omega}), \quad \vartheta_{01}(\omega) = \vartheta_{00}(i\bar{\omega}), \quad \vartheta_{10}(\omega) = e^{\mp \frac{\pi i}{4}} \vartheta_{10}(i\bar{\omega}).$$

Somit wird auf der linken Begrenzung von F_{14} bzw. der rechten von F_{41} :

$$k' = \left[\frac{\vartheta_{00}(i\bar{\omega})}{\vartheta_{01}(i\bar{\omega})} \right]^2 \geq 1 \text{ reell und durchläuft alle Werte von } 1 \text{ bis } +\infty,$$

$$k = e^{\mp \frac{\pi i}{2}} \left[\frac{\vartheta_{10}(i\bar{\omega})}{\vartheta_{01}(i\bar{\omega})} \right]^2 \text{ rein imaginär und durchläuft alle Werte von } 0 \text{ bis } \mp i\infty;$$

auf der imaginären Achse aber, die F_{14} und F_{41} trennt, ist

$$k' = \left[\frac{\vartheta_{01}(i\bar{\omega})}{\vartheta_{00}(i\bar{\omega})} \right]^2 \leq 1 \text{ reell und erfüllt das Intervall von } 0 \text{ bis } +1,$$

$$k = \left[\frac{\vartheta_{10}(i\bar{\omega})}{\vartheta_{00}(i\bar{\omega})} \right]^2 \leq 1 \text{ reell und erfüllt das Intervall von } 0 \text{ bis } +1.$$

Die beiden halbkreisförmigen Begrenzungen von F_{14} , F_{41} gehen aus den Geraden $\mp 1 \dots i\infty$ bzw. durch die Transformation $\omega \rightarrow -\frac{1}{\omega}$ hervor, daher ist auf der unteren Begrenzung von F_{14} bzw. F_{41} nach den Formeln (4)

$$k' = e^{\pm \frac{\pi i}{2}} \left[\frac{\vartheta_{10}(i\bar{\omega})}{\vartheta_{01}(i\bar{\omega})} \right]^2 \text{ rein imaginär und durchläuft alle Werte von } 0 \text{ bis } \pm i\infty,$$

$$k = \left[\frac{\vartheta_{00}(i\bar{\omega})}{\vartheta_{01}(i\bar{\omega})} \right]^2 \geq 1 \text{ reell und durchläuft alle Werte von } +1 \text{ bis } +\infty.$$

Zusammenfassend entspricht also dem Gebiete F_{14} (F_{41}) der erste (vierte) Quadrant der k' - und der vierte (erste) Quadrant der k -Ebene, so daß daselbst $\Re(k) \geq 0$, $\Re(k') \geq 0$ ist.

Soll nun der in F_{14} und F_{41} geltende Algorithmus untersucht werden, so kommt dies darauf hinaus, die in der Reihe (5'') auftretenden Vorzeichen und ihr Verhältnis zum Algorithmus des einfachsten Mittels zu verfolgen. In F_{14} und F_{41} gilt folgende Eigenschaft der ϑ -Quotienten:

Ist

$$\Re \left(\frac{b_m}{a_{m-1}} \right) = \Re \left(\left[\frac{\vartheta_{01}(2^m \omega)}{\vartheta_{00}(2^{m-1} \omega)} \right]^2 \right) \geq 0,$$

so ist auch

$$(13) \quad \Re \left(\frac{b_{n+1}}{a_n} \right) = \Re \left(\left[\frac{\vartheta_{01}(2^{n+1}\omega)}{\vartheta_{00}(2^n\omega)} \right]^2 \right) \geq 0 \quad \text{für } n = m, m+1, m+2, \dots,$$

d. h. es gilt von der m -ten Stelle ab der Algorithmus des einfachsten Mittels.

Nach den Formeln (3), (3') ist nämlich

$$\left[\frac{\vartheta_{01}(2^{n+1}\omega)}{\vartheta_{00}(2^n\omega)} \right]^2 = \frac{\vartheta_{01}(2^n\omega)}{\vartheta_{00}(2^n\omega)},$$

andererseits gilt

$$\frac{\vartheta_{01}(2^m\omega)}{\vartheta_{00}(2^m\omega)} = \sqrt{\frac{2\vartheta_{00}(2^{m-1}\omega)\vartheta_{01}(2^{m-1}\omega)}{\vartheta_{00}(2^{m-1}\omega)^2 + \vartheta_{01}(2^{m-1}\omega)^2}} = \sqrt{\frac{2\frac{\vartheta_{01}(2^{m-1}\omega)}{\vartheta_{00}(2^{m-1}\omega)}}{1 + \left[\frac{\vartheta_{01}(2^{m-1}\omega)}{\vartheta_{00}(2^{m-1}\omega)} \right]^2}},$$

und dabei ist für die Quadratwurzel derjenige Zweig zu nehmen, der für rein imaginäres ω , d. h. reelles $q < 1$, positiv reell ist. Setzt man also

$$\frac{\vartheta_{01}(2^{m-1}\omega)}{\vartheta_{00}(2^{m-1}\omega)} = z,$$

so ist

$$\frac{\vartheta_{01}(2^m\omega)}{\vartheta_{00}(2^m\omega)} = + \sqrt{\frac{2z}{1+z^2}}.$$

Nach Voraussetzung ist nun

$$|\arg z| \leq \frac{\pi}{2},$$

$$|\arg(1+z^2)| \leq 2 \arg z,$$

folglich

$$\arg z \geq \arg \left(\frac{2z}{1+z^2} \right) = \arg(2z) - \arg(1+z^2) \geq -\arg z, \quad \text{wenn } \arg z \geq 0,$$

$$-\arg z \geq \arg \left(\frac{2z}{1+z^2} \right) = \arg(2z) - \arg(1+z^2) \geq \arg z, \quad \text{wenn } \arg z \leq 0,$$

also

$$\left| \arg \left(\frac{2z}{1+z^2} \right) \right| \leq \frac{\pi}{2}, \quad \left| \arg \frac{\vartheta_{01}(2^m\omega)}{\vartheta_{00}(2^m\omega)} \right| \leq \frac{\pi}{2}.$$

In der Tat ist damit die erste der Beziehungen (13) bewiesen; allgemein findet man auf gleichem Wege, daß

$$\left| \arg \frac{\vartheta_{01}(2^{m+\kappa}\omega)}{\vartheta_{00}(2^{m+\kappa}\omega)} \right| \leq \frac{\pi}{2^{\kappa+2}} \quad (\kappa = 0, 1, 2, \dots),$$

womit die Behauptung (13) erledigt ist.

Da nun, wie oben gezeigt, in F_{14}, F_{41} die Beziehung $\Re(k') \geq 0$ gilt, so folgt aus dem eben Bewiesenen, daß allgemein daselbst

$$\Re(k'_r) \geq 0 \quad (r = 1, 2, 3, \dots),$$

ist, d. h. in F_{14} und F_{41} liefert (5'') den Algorithmus des einfachsten Mittels $M(a, b)$. Ferner führt die Transformation $\omega \rightarrow -\frac{1}{\omega}$ den Bereich $F_{14} + F_{41}$ in sich selbst über, daher liefert auch der adjungierte Kanon $-\frac{1}{\omega}$ das einfachste Mittel $M(a, c)$, und daselbst sind $M(a, b)$ und $M(a, c)$ zusammengehörige Mittel.

Wir kommen nun zur Untersuchung von F_{43} , F_{12} , die aus F_{41} bzw. F_{14} durch die Transformation $\omega' = \omega \mp 2$ entstehen, so daß [nach (4')] daselbst folgt

$$\begin{aligned} \vartheta_{00}(\omega')^2 &= \vartheta_{00}(\omega)^2, & \vartheta_{01}(\omega')^2 &= \vartheta_{01}(\omega)^2, & \vartheta_{10}(\omega')^2 &= -\vartheta_{10}(\omega)^2, \\ \vartheta_{00}(2^n \omega')^2 &= \vartheta_{00}(2^n \omega)^2, & \vartheta_{01}(2^n \omega')^2 &= \vartheta_{01}(2^n \omega)^2, \\ \vartheta_{10}(2^n \omega')^2 &= \vartheta_{10}(2^n \omega)^2 & (n = 1, 2, 3, \dots) \end{aligned}$$

Hieraus ergibt sich, daß sich F_{43} auf den vierten Quadranten der k' - und den dritten der k -Ebene abbildet, während F_{12} dem ersten bzw. zweiten Quadranten dieser Ebenen entspricht; und weiter erkennt man aus den angeführten Beziehungen, daß auch in F_{43} und F_{12} der Algorithmus des einfachsten Mittels gilt. Der adjungierte Modul dieser beiden Gebiete liegt in F_{34} bzw. F_{21} , und wir werden gleich sehen, daß zu diesen beiden Gebieten gleichfalls das einfachste Mittel gehört, so daß also in F_{43} und F_{14} die Mittel $M(a, b)$ und $M(a, c)$ zusammengehörig sind.

Zusammenfassend bildet sich — und das ist für das Folgende wichtig — der obere Teil des Fundamentalbereiches, d. h. die Gebiete F_{43} , F_{14} , F_{41} , F_{12} , durch (12) auf die rechte k' -Halbebene und die volle k -Ebene ab.

Wir untersuchen nun F_{21} , das aus F_{41} durch die Transformation $\omega' = \frac{\omega}{-2\omega + 1} = \frac{-1}{-\frac{1}{\omega} + 2}$ entsteht. Aus (4), (4') folgt daher

$$\begin{aligned} \vartheta_{00}(\omega') &= \sqrt{\frac{i}{\omega} - 2i} \vartheta_{00}\left(-\frac{1}{\omega} + 2\right) = \sqrt{\frac{i}{\omega} - 2i} \vartheta_{00}\left(-\frac{1}{\omega}\right) \\ &= \sqrt{1 - 2\omega} \vartheta_{00}(\omega), \\ \vartheta_{01}(\omega') &= \sqrt{\frac{i}{\omega} - 2i} \vartheta_{10}\left(-\frac{1}{\omega} + 2\right) = i \sqrt{\frac{i}{\omega} - 2i} \vartheta_{10}\left(-\frac{1}{\omega}\right) \\ &= i \sqrt{1 - 2\omega} \vartheta_{01}(\omega), \\ \vartheta_{10}(\omega') &= \sqrt{\frac{i}{\omega} - 2i} \vartheta_{01}\left(-\frac{1}{\omega} + 2\right) = \sqrt{\frac{i}{\omega} - 2i} \vartheta_{01}\left(-\frac{1}{\omega}\right) \\ &= \sqrt{1 - 2\omega} \vartheta_{10}(\omega), \end{aligned}$$

somit

$$\vartheta_{00}(\omega')^2 : \vartheta_{01}(\omega')^2 : \vartheta_{10}(\omega')^2 = \vartheta_{00}(\omega)^2 : -\vartheta_{01}(\omega)^2 : \vartheta_{10}(\omega)^2,$$

d. h. es entspricht F_{21} der zweite Quadrant der k' - und der erste der k -Ebene. Um den Algorithmus (5'') zu verfolgen, beachten wir, daß, wenn ω in F_{21} liegt, dann 2ω in dem Kreisbogendreieck mit den Ecken $0, -1, -2$ liegt, das ganz in $F_{43} + F_{14}$ enthalten ist, so daß nach obigem

$$\Re(k'_1) = \Re\left(\left[\frac{\vartheta_{01}(2\omega)}{\vartheta_{00}(2\omega)}\right]^2\right) \geq 0$$

ist. Da zu $F_{43} + F_{14}$ der Algorithmus des einfachsten Mittels gehört, so ist also erst recht $\Re(k'_r) \geq 0$ ($r = 2, 3, \dots$), d. h. wenn ω in F_{21} liegt, gibt (5'') den Algorithmus des einfachsten Mittels. Man kann dasselbe auch direkt den Transformationsformeln entnehmen, denn es wird z. B.

$$\vartheta_{00}(2\omega') = \sqrt{1-2\omega} \vartheta_{10}(2\omega), \quad \vartheta_{01}(2\omega') = i^{\frac{1}{2}} \sqrt{1-2\omega} \vartheta_{01}(2\omega),$$

also

$$\left[\frac{\vartheta_{01}(2\omega')}{\vartheta_{00}(\omega')}\right]^2 = i \left[\frac{\vartheta_{01}(2\omega)}{\vartheta_{00}(\omega)}\right]^2 = i \frac{\vartheta_{01}(\omega)}{\vartheta_{00}(\omega)},$$

wo ω in F_{41} zu liegen kommt. Dieser Ausdruck hat aber daselbst einen positiven Realteil, da k' im vierten Quadranten, also $\frac{\vartheta_{01}(\omega)}{\vartheta_{00}(\omega)}$ in dem von den Geraden mit dem Argument 0 und $-\frac{\pi}{4}$ begrenzten Winkelraum gelegen ist.

Aus (10) entnimmt man für diesen Bereich die Formel

$$(14) \quad \frac{1}{M(a, b)} = \frac{-2i}{M(a, c)} + \frac{1}{M(a, -b)}; \quad M(a, b) = \frac{M(a, -b) M(a, c)}{M(a, c) - 2i M(a, -b)},$$

gültig, wenn $k' = \frac{b}{a}$ im zweiten Quadranten, $k = \frac{c}{a}$ im ersten Quadranten liegt.

Den Bereich F_{34} erhält man auf ähnliche Weise aus F_{14} durch den Übergang

$$\omega' = \frac{\omega}{2\omega+1} = \frac{1}{\frac{1}{\omega}+2},$$

so daß

$$\vartheta_{00}(\omega')^2 : \vartheta_{01}(\omega')^2 : \vartheta_{10}(\omega')^2 = \vartheta_{00}(\omega)^2 : -\vartheta_{01}(\omega)^2 : \vartheta_{10}(\omega)^2,$$

wird; also entspricht F_{34} dem dritten Quadranten der k' - und dem vierten der k -Ebene. $2\omega'$ liegt ganz in $F_{41} + F_{13}$, so daß allgemein

$$\Re\left(\left[\frac{\vartheta_{01}(2^n \omega')}{\vartheta_{00}(2^n \omega')}\right]^2\right) \geq 0 \quad (n = 1, 2, \dots),$$

d. h. in F_{34} gilt der Algorithmus des einfachsten Mittels. Die Anwendung von (10) ergibt die Formel

$$(14') \quad \frac{1}{M(a, b)} = \frac{2i}{M(a, c)} + \frac{1}{M(a, -b)}; \quad M(a, b) = \frac{M(a, -b) M(a, c)}{M(a, c) + 2i M(a, -b)}, \quad 12)$$

¹²⁾ Diese Formel findet sich auch in Gauß' Nachlaß (vgl. 1., S. 478; 4., § 33), jedoch ohne die Angabe der Gültigkeitsbedingung und der ergänzenden Formel (14).

gültig, wenn $k' = \frac{b}{a}$ im dritten, $k = \frac{c}{a}$ im vierten Quadranten liegen. Liegt ω in F_{21} oder F_{34} , so liegt der adjungierte Kanon in F_{12} bzw. F_{43} , liefert also das einfachste Mittel $M(a, c)$ und in $F_{21} + F_{34}$ sind demnach $M(a, b)$ und $M(a, c)$ zusammengehörige Mittel.

Es sind noch die beiden Zipfel F_{32} und F_{23} zu behandeln, die aus F_{34} bzw. F_{21} durch die Substitution $\omega'' = \omega' \mp 2$, also aus F_{14} bzw. F_{41} durch den Übergang

$$\omega'' = \frac{-3\omega - 2}{2\omega + 1} \quad \text{bzw.} \quad \omega'' = \frac{-3\omega + 2}{-2\omega + 1}$$

hervorgehen. Daher wird:

$$\begin{aligned} \vartheta_{00}(\omega'')^2 : \vartheta_{01}(\omega'')^2 : \vartheta_{10}(\omega'')^2 &= \vartheta_{00}(\omega')^2 : \vartheta_{01}(\omega')^2 : -\vartheta_{10}(\omega')^2 \\ &= \vartheta_{00}(\omega)^2 : -\vartheta_{01}(\omega)^2 : -\vartheta_{10}(\omega)^2, \end{aligned}$$

d. h. es bilden sich F_{32} (F_{23}) auf den dritten (zweiten) Quadranten der k' - und den zweiten (dritten) Quadranten der k -Ebene ab. Da die Größen $2^n \omega''$ ($n = 1, 2, 3, \dots$) wieder in den Bereich zu liegen kommen, für den k' einen positiven Realteil hat, so gilt allgemein

$$\Re \left(\left[\frac{\vartheta_{01}(2^n \omega'')}{\vartheta_{00}(2^n \omega'')} \right]^2 \right) \geq 0 \quad (n = 1, 2, \dots),$$

d. h. in F_{32} und F_{23} gibt (5'') das einfachste Mittel $M(a, b)$. Jedoch nicht so beim adjungierten Kanon! Es ist nämlich

$$\begin{aligned} \vartheta_{00}\left(-\frac{1}{\omega''}\right) &= \sqrt{\frac{\omega''}{i}} \vartheta_{00}(\omega' \mp 2) = \sqrt{\frac{\omega''}{i}} \vartheta_{00}\left(\frac{1}{\omega \pm 2}\right) \\ &= \sqrt{\frac{\omega''}{i}} \sqrt{\frac{i}{\omega} \pm 2i} \vartheta_{00}\left(-\frac{1}{\omega}\right) = \sqrt{\omega''(1 \pm 2\omega)} \frac{1}{i} \vartheta_{00}(\omega), \\ \vartheta_{01}\left(-\frac{2}{\omega''}\right) &= \sqrt{\frac{\omega''}{2i}} \vartheta_{00}\left(\frac{\omega'}{2} \mp 1\right) = i^{\mp \frac{1}{2}} \sqrt{\frac{\omega''}{2i}} \vartheta_{10}\left(\frac{\omega'}{2}\right) \\ &= i^{\mp \frac{1}{2}} \sqrt{\frac{\omega''}{2i}} \sqrt{\frac{2i}{\omega} \pm 4i} \vartheta_{01}\left(-\frac{2}{\omega}\right) = i^{\mp \frac{1}{2}} \sqrt{\omega''(1 \pm 2\omega)} \frac{1}{2i} \vartheta_{10}\left(\frac{\omega}{2}\right), \end{aligned}$$

also der Quotient

$$\left[\frac{\vartheta_{01}\left(-\frac{2}{\omega''}\right)}{\vartheta_{00}\left(-\frac{1}{\omega''}\right)} \right]^2 = i^{\mp 1} \frac{1}{2} \left[\frac{\vartheta_{10}\left(\frac{\omega}{2}\right)}{\vartheta_{00}(\omega)} \right]^2 = i^{\mp 1} \frac{\vartheta_{10}(\omega)}{\vartheta_{00}(\omega)} = i^{\mp 1} \sqrt{k},$$

wo das obere bzw. untere Vorzeichen zu nehmen ist, je nachdem ω in F_{14} oder F_{41} , also k im vierten oder ersten Quadranten liegt. Sonach ist in jedem Falle

$$\Re \left(\left[\frac{\vartheta_{01}\left(-\frac{2}{\omega''}\right)}{\vartheta_{00}\left(-\frac{1}{\omega''}\right)} \right]^2 \right) \leq 0,$$

d. h. der Algorithmus (5'') des adjungierten Kanons ist in F_{32} und F_{23} nicht der des einfachsten Mittels, daselbst sind also $M(a, b)$ und $M(a, c)$ nicht zusammengehörige Mittel¹³⁾.

Vergleichen wir jetzt unsere Resultate mit denen des vorigen Paragraphen, so sehen wir, daß in E und nur dort überall der Algorithmus des einfachsten Mittels gilt, und zwar liegt ω für positives $\Re\left(\frac{b}{a}\right)$ in dem oberen Bereich $F_{43}, F_{14}, F_{41}, F_{12}$ und den um $4n$ verschobenen Streifen, für negatives $\Re\left(\frac{b}{a}\right)$ in $F_{32}, F_{21}, F_{34}, F_{23}$ und den verschobenen Sicheln.

§ 4.

Die Deutung des Algorithmus in der ω -Ebene.

Die möglichen Algorithmen (1) klassifizieren sich in *reguläre* und *irreguläre*. Als *regulär* bezeichnen wir einen Algorithmus, der von irgend-einer Stelle ν ab der Forderung (2) genügt, als *irregulär* hingegen einen solchen, bei dem kein solcher Index ν angegeben werden kann, von dem ab der Algorithmus des einfachsten Mittels gilt, bei dem also die Quadratwurzelbestimmung ständig oszilliert. Die irregulären Algorithmen haben als Grenzwert Null¹⁴⁾, nur die regulären führen zu von Null verschiedenen $\mathfrak{M}(a, b)$. Wir wollen einen regulären Algorithmus, der erst von der $\nu + 1$ -ten Stelle — und von keiner früheren — ab ständig der Forderung (2) genügt, als einen *Algorithmus ν -ter Stufe* bezeichnen, so daß hiernach $M(a, b)$ zum Algorithmus von höchstens 0-ter Stufe gehört. Diesen letzteren haben wir eben in der ω -Ebene lokalisiert, es entspricht ihm nämlich ein ω aus dem Bereich $\Phi \pm 4n = E$ (vgl. Fig. 2), und umgekehrt liefert jedes ω dieses Bereiches durch (5'') einen Algorithmus höchstens 0-ter Stufe. Wir lokalisieren zunächst allgemein die Algorithmen ν -ter Stufe.

Dazu bemerken wir, daß der Ausführung eines Schrittes, d. h. dem Übergang von $\nu - 1$ zu ν in (1) im ϑ -Algorithmus (5'') der Übergang zum doppelten Kanon entspricht. Soll also ω den Kanon eines Algorithmus erster Stufe darstellen, so muß 2ω — und erst dieses — ganz im Bereich E des einfachsten Mittels liegen, mithin liegt dann ω in den Kreisbogendreiecken mit den Ecken: $\frac{n}{2}, \frac{2n+1}{4}, \frac{n+1}{2}$ ($n=0, \pm 1, \pm 2, \pm 3, \dots$), also in Sicheln, die E umlagern. Die Gesamtheit dieser Sicheln bildet das Gebiet E^1 der Algorithmen erster Stufe. Ist ω der Kanon eines Algorithmus zweiter Stufe, so muß 4ω — aber noch nicht 2ω — in E zu

¹³⁾ Dies ist schon aus dem Vergleich der Figuren 2 und 3 ersichtlich.

¹⁴⁾ Vgl. v. Dávid 4.

liegen kommen, d. h. ω muß in den Kreisbogendreiecken mit den Ecken $\frac{n}{4}, \frac{2n+1}{8}, \frac{n+1}{4}$ ($n = 0, \pm 1, \pm 2, \dots$) liegen, die in ihrer Gesamtheit ein das Gebiet E^1 umlagerndes Sichelgebiet E^2 ergeben. Man erkennt sofort allgemein, daß die uniformisierende Variable der Algorithmen ν -ter Stufe in den Kreisbogendreiecken mit den Ecken $\frac{n}{2^\nu}, \frac{2n+1}{2^{\nu+1}}, \frac{n+1}{2^\nu}$ ($n = 0, \pm 1, \pm 2, \dots$) liegen muß, deren Gesamtheit ein $E^{\nu-1}$ umlagerndes Sichelgebiet E^ν bildet. In dieser Weise wird die ω -Halbebene entsprechend den zugehörigen Algorithmen parzelliert (vgl. Fig. 5). Man erkennt hieraus so-

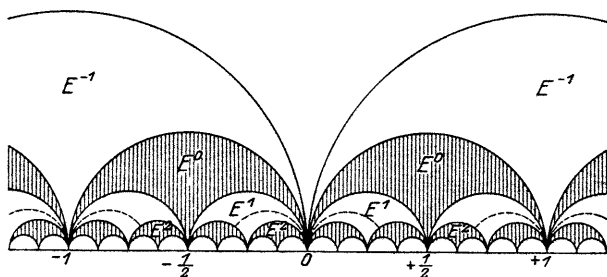


Fig. 5.

fort, daß es zu jedem ω der oberen Halbebene, das nicht auf der reellen Achse liegt, mit Notwendigkeit einen endlichen Index ν geben muß, von dem ab (5'') dem Algorithmus des einfachsten Mittels folgt, d. h. daß also nur die regulären Algorithmen durch die Jacobischen Thetafunktionen uniformisiert werden können. Übrigens werden durch unser Parzellierungsverfahren auch diejenigen Punkte der reellen ω -Achse erreicht, die eine endliche dyadische Entwicklung gestatten.

Wir haben ferner zu zeigen, daß in dieser Weise alle regulären Algorithmen erfaßt werden; zu gegebenen Anfangswerten a, b gibt es 2^ν verschiedene Algorithmen ν -ter Stufe, da bei den ersten ν Schritten die Wahl des Wurzelvorzeichens beliebig bleibt. Es kommt daher der verlangte Nachweis darauf hinaus, zu zeigen, daß allgemein E^ν äquivalent ist mit der Summe $E + E^1 + E^2 + \dots + E^{\nu-1}$, d. h. mit dem ganzen darüberliegenden Teil der ω -Ebene. Wir zeigen dies zunächst für E^1 . Die Substitutionen

$$(15) \quad \omega' = \frac{-(4n-1)\omega + 4n^2}{-4\omega + (4n+1)},$$

$$(15') \quad \omega' = \frac{(4n+1)\omega - 4n^2}{4\omega - (4n-1)}$$

gehören zur Kongruenzgruppe Γ^4 und leisten die folgenden Abbildungen:

$$(15) \text{ Dreieck } n, n + \frac{1}{2}, i\infty \text{ von } E \rightarrow \text{Sichel } n, n - \frac{1}{2}, n - \frac{1}{4} \text{ von } E^1,$$

$$(15') \text{ Dreieck } n, n - \frac{1}{2}, i\infty \text{ von } E \rightarrow \text{Sichel } n, n + \frac{1}{2}, n + \frac{1}{4} \text{ von } E^1,$$

so daß durch (15), (15') eine Äquivalenzbeziehung zwischen ganz E und ganz E^1 vermittelt wird; die Zuordnung von Streifen und Sicheln erfolgt über Kreuz. Die Algorithmen erster Stufe zerfallen in zwei Arten, je nachdem sie beim nullten Schritt der Forderung (2') genügen oder nicht (beim ersten Schritt genügen sie ihr bestimmt nicht!). Um diese beiden Möglichkeiten zu lokalisieren, haben wir die Abbildung des Oberteils von E , d. h. von $F_{43}, F_{14}, F_{41}, F_{12}, \dots$, durch (15), (15') zu untersuchen. Nun bildet sich ab:

$$(15) \text{ Dreieck } n, n + \frac{1}{2} + \frac{i}{2}, i\infty \text{ von } E \rightarrow \text{Sichel } n, n - \frac{3}{10} + \frac{i}{10}, n - \frac{1}{4},$$

$$(15') \text{ Dreieck } n, n - \frac{1}{2} + \frac{i}{2}, i\infty \text{ von } E \rightarrow \text{Sichel } n, n + \frac{3}{10} + \frac{i}{10}, n + \frac{1}{4}.$$

Durch die Punkte $n \pm \frac{3}{10} + \frac{i}{10}$ wird also E^1 in zwei Gebiete geteilt, die den beiden möglichen Arten von Algorithmen erster Stufe entsprechen; sie sind in der Figur 5 gekennzeichnet.

Um nun E^2 zu untersuchen, nehmen wir die vier Substitutionen von Γ^4 :

$$(16) \quad \omega' = \frac{-(8n-1)\omega + 8n^2}{-8\omega + (8n+1)},$$

$$(16') \quad \omega' = \frac{-(8n+1)\omega + 8n^2}{-8\omega + (8n-1)},$$

$$(16'') \quad \omega' = \frac{(5-8n)\omega + 8(n^2+n-1)}{-8\omega + (8n+13)},$$

$$(16''') \quad \omega' = \frac{(8n+5)\omega - 8(n^2-n-1)}{8\omega + (13-8n)}.$$

Diese vermitteln die Äquivalenzbeziehung zwischen $E + E^1$ einerseits, E^2 andererseits; denn es bildet sich ab:

$$(16) \text{ Dreieck } n, n + \frac{1}{4}, i\infty \rightarrow \text{Sichel } n, n - \frac{1}{4}, n - \frac{1}{8},$$

$$(16') \text{ Dreieck } n, n - \frac{1}{4}, i\infty \rightarrow \text{Sichel } n, n + \frac{1}{4}, n + \frac{1}{8},$$

$$(16'') \text{ Dreieck } n + \frac{3}{2}, n + \frac{7}{4}, i\infty \rightarrow \text{Sichel } n - \frac{1}{2}, n - \frac{3}{4}, n - \frac{5}{8},$$

$$(16''') \text{ Dreieck } n - \frac{3}{2}, n - \frac{7}{4}, i\infty \rightarrow \text{Sichel } n + \frac{1}{2}, n + \frac{3}{4}, n + \frac{5}{8}.$$

Die Algorithmen zweiter Stufe gliedern sich in vier Arten, je nachdem

beim nullten oder ersten Schritt die Forderung (2') verletzt oder erfüllt wird. Man lokalisiert diese vier Arten, indem man das Ober- und Untergebiet von E und E^1 durch (16)...(16''') auf E^2 abbildet.

Wir zeigen jetzt allgemein, daß E^r äquivalent ist mit $E + E^1 + \dots + E^{r-1}$. Wir bilden nämlich zunächst im Falle, daß $r = 2m - 1$ ungerade ist, die Substitutionen:

$$(17) \quad \omega' = \frac{\omega(1 - 4n + 16n^2 - 64n^3 \pm \dots + (-1)^m 4^m n^m) + (-1)^m 4n^{m+1}}{4^m \omega + (4n + 1)} + \text{ganze Zahl},$$

$$(17') \quad \omega' = \frac{\omega(1 - 4n + 16n^2 - 64n^3 \pm \dots + (-1)^m 4^m n^m) - (-1)^m 4n^{m+1}}{-4^m \omega + (4n + 1)} + \text{ganze Zahl};$$

beide gehören der Kongruenzgruppe Γ^4 an und besitzen die Determinante 1, und es bildet ab:

(17) den Streifen $-\frac{n}{4^{m-1}}$, $-\frac{n+\frac{1}{2}}{4^{m-1}}$, ∞i auf die Sichel der Breite $\frac{1}{2^r}$ mit den Ecken:

$$\left. \begin{aligned} &(-1)^m \left\{ n^m - \frac{1}{4} n^{m-1} + \frac{1}{16} n^{m-2} \mp \dots + \frac{(-1)^{m-1}}{4^{m-1}} n \right\} \\ &(-1)^m \left\{ n^m - \frac{1}{4} n^{m-1} + \frac{1}{16} n^{m-2} \mp \dots + \frac{(-1)^{m-1}}{4^{m-1}} n + \frac{2(-1)^m}{4^m} \right\} \\ &(-1)^m \left\{ n^m - \frac{1}{4} n^{m-1} + \frac{1}{16} n^{m-2} \mp \dots + \frac{(-1)^{m-1}}{4^{m-1}} n + \frac{(-1)^m}{4^m} \right\} \end{aligned} \right\} + \text{ganze Zahl},$$

(17') den Streifen $\frac{n}{4^{m-1}}$, $\frac{n+\frac{1}{2}}{4^{m-1}}$, ∞i auf die Sichel der Breite $\frac{1}{2^r}$ mit den Ecken:

$$\left. \begin{aligned} &(-1)^{m-1} \left\{ n^m - \frac{1}{4} n^{m-1} + \frac{1}{16} n^{m-2} \mp \dots + \frac{(-1)^{m-1}}{4^{m-1}} n \right\} \\ &(-1)^{m-1} \left\{ n^m - \frac{1}{4} n^{m-1} + \frac{1}{16} n^{m-2} \mp \dots + \frac{(-1)^{m-1}}{4^{m-1}} n + \frac{2(-1)^m}{4^m} \right\} \\ &(-1)^{m-1} \left\{ n^m - \frac{1}{4} n^{m-1} + \frac{1}{16} n^{m-2} \mp \dots + \frac{(-1)^{m-1}}{4^{m-1}} n + \frac{(-1)^m}{4^m} \right\} \end{aligned} \right\} + \text{ganze Zahl}.$$

Die Gesamtheit der abgebildeten Streifen bildet den über E^r liegenden Bereich $E + E^1 + \dots + E^{r-1}$; es ist noch nachzuweisen, daß die Gesamtheit der Sichel den Bereich E^r erschöpft, d. h. die Gesamtheit der mittleren Sichelcken:

$$\begin{aligned} \pm \frac{1}{4^m} \{1 - 4n + 16n^2 \mp \dots + (-1)^m 4^m n^m\} \\ = \pm \frac{1}{4^m} \cdot \frac{1 + (-1)^m 4^{m+1} n^{m+1}}{1 + 4n} \end{aligned}$$

muß bis auf ganze Zahlen identisch sein mit der Gesamtheit der Zahlen

$$\frac{l}{2^v} + \frac{1}{4^m} = \frac{2l+1}{4^m} \quad (l = 0, \pm 1, \pm 2, \dots),$$

wo gerades l dem oberen, ungerades l dem unteren Vorzeichen entspricht, d. h. es muß zu jeder ganzen Zahl l ein n geben, so daß

$$-n + 4n^2 - 16n^3 \pm \dots + (-1)^{m-1} 4^{m-2} n^{m-1} \equiv l \pmod{4^{m-1}}$$

ist. Dies ist aber selbstverständlich, denn läßt man n alle Zahlen von 1 bis 4^{m-1} durchlaufen, so erschöpft die linke Seite alle Restklassen mod 4^{m-1} , da andernfalls aus

$$(n' - n) \{-1 + 4(n + n') - 16(n^2 + nn' + n'^2) + \dots\} \equiv 0 \pmod{4^{m-1}}$$

die Beziehung $n' - n \equiv 0 \pmod{4^{m-1}}$ folgen müßte.

Der Fall eines ungeraden v ist damit erledigt; für gerade $v = 2m$ treten an Stelle von (17), (17') die Substitutionen

$$\left. \begin{aligned} \omega' &= \frac{\omega(1 - 4n + 16n^2 \mp \dots + (-1)^{m+1} 4^{m+1} n^{m+1}) - (-1)^m 8n^{m+2}}{2^{2m+1} \omega} && + (4n + 1) \\ \omega' &= \frac{\omega(1 - 4n + 16n^2 \mp \dots + (-1)^{m+1} 4^{m+1} n^{m+1}) + (-1)^m 8n^{m+2}}{-2^{2m+1} \omega} && + (4n + 1) \end{aligned} \right\} + \text{ganze Zahl,}$$

und der weitere Beweis verläuft dem obigen parallel.

Man kann nun die gefundene Parzellierung auch auf das Gebiet E des einfachsten Mittels ausdehnen, wenn man die rückwärtige Fortsetzung der Algorithmen (1) bzw. (5'') verfolgt. Die Fortsetzung von (1) nach rückwärts wird gegeben durch die Gleichungen

$$(1') \quad a_{-(v+1)} = a_{-v} + c_{-v}, \quad b_{-(v+1)} = a_{-v} - c_{-v}; \quad c_{-v} = \sqrt{a_{-v}^2 - b_{-v}^2}$$

$$(v = 0, 1, 2, \dots)$$

und ist daher ebenso vieldeutig wie die Progression (1) selbst. Auch (5'') läßt sich aber nach rückwärts verlängern, indem man nur v negative Zahlen durchlaufen läßt. Diese Algorithmen gliedern sich wieder in solche (-1) -ter, (-2) -ter, ... Stufe, indem zur $(-n)$ -ten Stufe diejenigen gehören, für die von der $(-n+1)$ -ten Stelle und erst von dieser ab $\Re\left(\frac{b_v}{a_v}\right) \geq 0$, ($v > -n$) ist; wir lokalisieren die Algorithmen $(-n)$ -ter Stufe in der ω -Ebene, indem der zugehörige Kanon im Gebiete $E^{(-n)}$ liegt, das

von den Kreisbogendreiecken mit den Ecken $m2^n$, $(m + \frac{1}{2})2^n$, $(m + 1)2^n$ ($m = 0, \pm 1, \pm 2, \dots$) gebildet wird. Dadurch zerfällt E in die Summe der Gebiete $E^0 + E^{-1} + E^{-2} + \dots$, und die ganze ω -Ebene ist in gleichmäßiger Weise parzelliert (vgl. Fig. 5). Man sieht hieraus, daß für kein endliches ω der Algorithmus (5'') nach vorn und hinten der Forderung (2') genügen kann, es sei denn, daß ω auf der imaginären Achse liegt, also a, b positiv reell sind; für jedes andere ω gibt es mit Notwendigkeit einen (positiven oder negativen) Index ν , für den (2') durchbrochen wird.

Literaturnachweis.

Ludwig von Dávid:

1. Theorie des Gaußschen verallgemeinerten und speziellen arithmetisch-geometrischen Mittels. Math.-naturw. Berichte aus Ungarn 25 (1907), S. 153–171.
2. Sur une application des fonctions modulaires à la théorie de la moyenne arithmético-géométrique. Math.-naturw. Berichte aus Ungarn 27 (1909), S. 164–171.
3. Zur Gaußschen Theorie der Modulfunktion. Rendiconti di Palermo 35 (1913), S. 82–89.
4. Arithmetisch-geometrisches Mittel und Modulfunktion. Erscheint im Journal für reine und angew. Math. 159 (1928).

Carl Friedrich Gauß:

1. Werke 3 (1866), S. 361–401; 461–478; herausgegeben von E. Schering.
2. Werke 8 (1900), S. 99–105; herausgegeben von R. Fricke.
3. Werke 10, 1 (1917), S. 172–289; herausgegeben von L. Schlesinger.
4. Anziehung eines elliptischen Ringes und Nachlaß zur Theorie des arithmetisch-geometrischen Mittels und der Modulfunktion, herausgegeben von Harald Geppert, Ostwalds Klassiker 225 (1927).

Ludwig Schlesinger:

1. Über die Gaußsche Theorie des arithmetisch-geometrischen Mittels und ihre Beziehungen zur Theorie der elliptischen Modulfunktion. Sitzungsberichte der Preuß. Akad. d. Wiss. 28 (1898), S. 346–360.
2. Handbuch der Theorie der linearen Differentialgleichungen 2, 2. Leipzig (1898).
3. C. F. Gauß: Fragmente zur Theorie des arithmetisch-geometrischen Mittels aus den Jahren 1797–1799. Materialien für eine wissenschaftliche Biographie von Gauß, Heft 2 (1912).
4. Über Gauß' Arbeiten zur Funktionentheorie. Materialien für eine wissenschaftliche Biographie von Gauß, Heft 3 (1912).

Gießen, Mathematisches Seminar, den 24. Februar 1927.

Über den Bereich absoluter Konvergenz von Potenzreihen mehrerer Veränderlichen.

(Aus einem Schreiben an Herrn F. Hartogs.)

Von

Heinrich Tietze in München.

Für den Satz¹⁾, daß dem Bereich der absoluten Konvergenz einer Potenzreihe (Laurent-Reihe)

$$(1) \quad \sum_{-\infty}^{+\infty} a_{r_1 \dots r_n} x_1^{r_1} \dots x_n^{r_n}$$

(alle $x_i \neq 0$ angenommen) ein *konvexer* Bereich im (ξ_1, \dots, ξ_n) -Raum entspricht (wobei $\xi_i = \log |x_i|$), haben Sie kürzlich den schönen, bis jetzt leider noch nicht veröffentlichten Beweis aus der unter Ihrer Anregung entstandenen Dissertation des Herrn Pfeufer²⁾ vorgetragen, der sich dabei Laurentscher Reihen mit nicht ganzzahligen Exponenten bedient. Die im folgenden mitgeteilte Beweisführung für den fraglichen Satz benützt einen Hilfssatz über Reihen nicht-negativer Glieder.

Hilfssatz. Wenn $0 < \vartheta < 1$ ist und die Reihen $\sum b_r$, $\sum c_r$ konvergieren ($b_r \geq 0$, $c_r \geq 0$), so konvergiert auch $\sum b_r^\vartheta c_r^{1-\vartheta}$; dies folgt nämlich aus der Konvergenz von $\sum (b_r + c_r)$ und aus $b_r^\vartheta c_r^{1-\vartheta} \leq m_r^\vartheta m_r^{1-\vartheta} = m_r \leq b_r + c_r$, wo $m_r = \text{Max}(b_r, c_r)$. (Analog konvergiert zugleich mit $\sum b_r', \dots, \sum b_r^{(k)}$ auch $\sum (b_r')^{\vartheta_r'} \dots (b_r^{(k)})^{\vartheta_r^{(k)}}$ für positive $\vartheta_r', \dots, \vartheta_r^{(k)}$ mit $\vartheta_r' + \dots + \vartheta_r^{(k)} = 1$.)

¹⁾ Bisherige Literatur (Fabry, Lemaire, Hartogs, Faber [für $n > 2$]) s. Enc. II, C. 4 (Bieberbach), S. 519 ff.

²⁾ Christian Pfeufer, *Konvexe Bereiche und Laurentsche Reihen*, Dissertation München 1924 (ungedruckt).

Sind nun ξ'_1, \dots, ξ'_n und ξ''_1, \dots, ξ''_n zwei Systeme reeller Zahlen ξ_1, \dots, ξ_n , für welche $\sum |a_{v_1 \dots v_n} x_1^{v_1} \dots x_n^{v_n}| = \sum |a_{v_1 \dots v_n}| e^{\sum v_i \xi_i}$ konvergiert, so konvergiert dem Hilfssatz gemäß für $0 < \vartheta < 1$ auch

$$\sum (|a_{v_1 \dots v_n}| e^{\sum v_i \xi'_i})^\vartheta \cdot (|a_{v_1 \dots v_n}| e^{\sum v_i \xi''_i})^{1-\vartheta} = \sum |a_{v_1 \dots v_n}| e^{\sum v_i (\vartheta \xi'_i + (1-\vartheta) \xi''_i)}.$$

Mit zwei Punkten (ξ'_i) , (ξ''_i) gehört also stets auch jeder Punkt $(\vartheta \xi'_i + (1-\vartheta) \xi''_i)$ ihrer Verbindungsstrecke zum Bereich \mathfrak{B} der absoluten Konvergenz: \mathfrak{B} ist konvex³⁾.

Der Beweis ist anwendbar auch auf Punkte (ξ'_i) , (ξ''_i) auf dem Rand von \mathfrak{B} (falls es solche Randpunkte von \mathfrak{B} gibt, die zu \mathfrak{B} gehören). Ist also etwa \mathfrak{R}_p ein nicht ins Innere von \mathfrak{B} dringender p -dimensionaler linearer Raum ($1 \leq p \leq n-1$), so müssen die auf \mathfrak{R}_p liegenden Randpunkte von \mathfrak{B} selbst eine konvexe Punktmenge bilden.

München, den 4. Sept. 1927.

³⁾ Ist (1) speziell eine gewöhnliche Potenzreihe (alle $v_i \geq 0$), so tritt die Zusatzbedingung hinzu, daß zugleich mit jedem Punkt (ξ'_1, \dots, ξ'_n) auch alle Punkte $\xi_1 \leq \xi'_1, \dots, \xi_n \leq \xi'_n$ zu \mathfrak{B} gehören.

(Eingegangen am 18. 9. 1927.)

Su un problema di Abel.

Von

Leonida Tonelli in Bologna.

Abel, in due Note del 1823 e del 1826¹⁾, considerò il seguente problema:

« Determinare una curva C , posta in un piano verticale, in modo che un grave, abbandonato con velocità iniziale nulla da un suo punto qualunque M e obbligato a percorrerla, arrivi al suo punto più basso O dopo un tempo che sia una funzione data $f(z)$ dell' altezza z del punto M sopra O . »

Considerato, nel piano della curva C , un sistema cartesiano ortogonale di assi, con l'origine nel punto O , l'asse delle x orizzontale, l'asse delle y rivolto verso l'alto, e scritta l'equazione della curva nella forma

$$(1) \quad x = \Phi(y)^2,$$

i principî della meccanica dànno immediatamente

$$\int_0^z \frac{\sqrt{1 + \Phi'^2(y)}}{\sqrt{2g(z-y)}} dy = f(z)^3,$$

equazione che, posto

$$\sqrt{1 + \Phi'^2(y)} = \varphi(y),$$

¹⁾ V. « Œuvres complètes » (Christiania, 1831) pp. 11 e 97.

²⁾ È evidente che la curva non può essere incontrata in più di un punto da una parallela all'asse delle x , perchè, altrimenti, esisterebbero su di essa o dei minimi distinti da O — e da essi il grave, abbandonato con velocità iniziale nulla, non potrebbe muoversi e non potrebbe quindi giungere in O — o dei tratti orizzontali — ed anche dai punti interni a questi tratti il mobile non potrebbe muoversi con velocità nulla.

³⁾ Qui e nel seguito, g rappresenta l'accelerazione dovuta alla gravità.

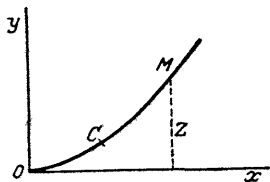


Fig. 1.

si scrive

$$(2) \quad \int_0^z \frac{\varphi(y) dy}{\sqrt{z-y}} = \sqrt{2g} f(z),$$

e per determinare la funzione incognita $\Phi(y)$, basta determinare la $\varphi(y)$ mediante questa equazione. Abel risolse l'equazione (2) (e con essa una equazione più generale) mediante la formola

$$(3) \quad s(z) = \sqrt{\frac{2g}{\pi}} \int_0^z \frac{f(y) dy}{\sqrt{z-y}},$$

dove $s(z)$ indica la lunghezza dell'arco OM della C , vale a dire, dove è

$$s(z) = \int_0^z \sqrt{1 + \Phi'^2(y)} dy = \int_0^z \varphi(y) dy.$$

Del problema studiato da Abel, che rappresenta, come è ben noto, il primo esempio di inversione di integrali definiti o, se si vuole, di equazioni integrali, si occuparono, in seguito, molti analisti, fra i quali citeremo il Sonine, il Volterra, il Goursat, il Burgatti, ecc. Per altro, in nessuno dei lavori pubblicati su tale problema, si trova un'analisi precisa delle condizioni di validità della soluzione.

A questo proposito devono farsi due osservazioni. La prima è che, mentre i ragionamenti che si fanno per giungere alla risoluzione dell'equazione (2) nulla lasciano a desiderare quando la funzione incognita $\varphi(y)$ è limitata e la funzione data $f(z)$ è continua con derivata prima limitata, nel caso opposto, invece, tali ragionamenti vanno opportunamente giustificati; in particolare, vanno del tutto giustificate le inversioni nell'ordine delle integrazioni di certi integrali multipli le cui funzioni integrande non sono limitate.

La seconda osservazione da farsi è che, ottenuta la soluzione dell'equazione (2), *resta ancora da vedere se il problema di Abel sia o no risolubile*. Ed infatti, poichè la funzione $\varphi(y)$ rappresenta la derivata della lunghezza dell'arco $s(y)$, ne viene che, affinchè la soluzione della (2) risolva il problema di Abel, è necessario (e sufficiente) che essa verifichi la disuguaglianza $\varphi(y) \geq 1$. E può ben accadere che la soluzione della (2) non soddisfi a questa condizione, ciò che appunto avviene se è, per esempio,

$$f(z) = \sqrt{\frac{z}{g}}.$$

Quando si tiene conto della condizione ora indicata, ci si accorge subito che, *anche nei casi più semplici, compreso quello della curva tautocrona,*

o è la funzione $\varphi(y)$ che risulta illimitata oppure è illimitata la derivata della $f(z)$, e quindi si è proprio in quelle circostanze in cui occorre giustificare l'inversione delle integrazioni cui dianzi accennavamo.

Nel presente lavoro, ci proponiamo di esaminare con tutta precisione le condizioni di risolubilità del problema di Abel, e di dare la piena giustificazione dei ragionamenti necessari per giungere alla sua soluzione.

1. Lemma I.

Indichiamo con T il triangolo del piano (x, y) avente per vertici i punti $(0, 0)$, $(1, 0)$ e $(1, 1)$, e dimostriamo il seguente lemma:

Se: 1.^o $A(x, y)$ è una funzione limitata e integrabile⁴⁾ in T ; 2.^o $\varphi(y)$ è una funzione integrabile nell'intervallo $(0, 1)$; 3.^o α è un numero tale che $0 < \alpha < 1$; allora:

a) esiste finito, per quasi tutti gli x di $(0, 1)$, l'integrale

$$(4) \quad \int_0^x \frac{A(x, y)}{(x-y)^\alpha} \varphi(y) dy;$$

b) esistono, pure finiti, per tutti gli z di $(0, 1)$ gli integrali

$$(5) \quad \int_0^z \frac{dx}{(z-x)^{1-\alpha}} \int_0^x \frac{A(x, y)}{(x-y)^\alpha} \varphi(y) dy, \quad \int_0^z \varphi(y) dy \int_y^z \frac{A(x, y)}{(z-x)^{1-\alpha}(x-y)^\alpha} dx;$$

c) vale sempre l'uguaglianza

$$(6) \quad \int_0^z \frac{dx}{(z-x)^{1-\alpha}} \int_0^x \frac{A(x, y)}{(x-y)^\alpha} \varphi(y) dy = \int_0^z \varphi(y) dy \int_y^z \frac{A(x, y)}{(z-x)^{1-\alpha}(x-y)^\alpha} dx.$$

Dimostriamo, innanzi tutto, che, per ogni z di $(0, 1)$, la funzione di (x, y)

$$\frac{A(x, y) \varphi(y)}{(z-x)^{1-\alpha}(x-y)^\alpha}$$

è integrabile nel triangolo T_z di vertici $(0, 0)$, $(z, 0)$ e (z, z) . Ed infatti, se N è un numero tale che sia, in tutto T , $|A(x, y)| \leq N$, abbiamo, in T_z ,

$$(7) \quad \left| \frac{A(x, y) \varphi(y)}{(z-x)^{1-\alpha}(x-y)^\alpha} \right| \leq \frac{N |\varphi(y)|}{(z-x)^{1-\alpha}(x-y)^\alpha}.$$

Ora, essendo, come è noto, per $0 \leq y < z$

$$(8) \quad \int_y^z \frac{dx}{(z-x)^{1-\alpha}(x-y)^\alpha} = \int_0^1 \frac{dt}{(1-t)^{1-\alpha} t^\alpha} = \frac{\pi}{\operatorname{sen} \pi \alpha},$$

⁴⁾ Intenderemo sempre, nel seguito, l'integrabilità nel senso del Lebesgue.

esiste finito (per l'integrabilità della $\varphi(y)$) l'integrale ripetuto

$$\int_0^z dy \int_y^z \frac{|\varphi(y)|}{(z-x)^{1-\alpha}(x-y)^\alpha} dx = \frac{\pi}{\operatorname{sen} \pi \alpha} \int_0^z |\varphi(y)| dy$$

e quindi, per un teorema da me dato altrove⁵⁾, la funzione

$$\frac{|\varphi(y)|}{(z-x)^{1-\alpha}(x-y)^\alpha}$$

è integrabile nel triangolo T_z . In virtù della (7) è allora integrabile in T_z anche la funzione

$$\frac{A(x, y) \varphi(y)}{(z-x)^{1-\alpha}(x-y)^\alpha}.$$

Applicando il teorema di Fubini sugli integrali multipli⁶⁾, possiamo dunque affermare: a) che questa funzione è linearmente integrabile, in T_z , rispetto alla y , per quasi tutti i valori di x , e quindi che esiste finito l'integrale (4) per quasi tutti gli x di $(0, z)$, od anche di $(0, 1)$, essendo z un qualunque valore di $(0, 1)$; b) che esistono finiti entrambi gli integrali (5); c) che vale l'uguaglianza (6).

2. Considerazioni generali.

Consideriamo l'uguaglianza

$$(9) \quad \int_0^x \frac{A(x, y)}{(x-y)^\alpha} \varphi(y) dy = f(x),$$

con $0 < \alpha < 1$, $A(x, y)$ funzione limitata e integrabile in T , $\varphi(y)$ e $f(x)$ funzioni integrabili in $(0, 1)$.

Supponiamo che l'uguaglianza (9) valga per quasi tutti gli x di $(0, 1)$. Per il lemma del n.º precedente, esiste finito l'integrale

$$\int_0^z \frac{dx}{(z-x)^{1-\alpha}} \int_0^x \frac{A(x, y)}{(x-y)^\alpha} \varphi(y) dy,$$

per tutti i valori di z di $(0, 1)$. Abbiamo perciò, per tutti questi valori di z ,

$$(10) \quad \int_0^z \frac{dx}{(z-x)^{1-\alpha}} \int_0^x \frac{A(x, y)}{(x-y)^\alpha} \varphi(y) dy = \int_0^z \frac{f(x)}{(z-x)^{1-\alpha}} dx.$$

⁵⁾ *Sull'integrazione per parti* (Rend. R. Accad. dei Lincei (5) 18 (1909), p. 246).

⁶⁾ *Sugli integrali multipli* (Rend. R. Accad. Lincei (5) 16 (1907), p. 608).

Inoltre, sempre per il nostro lemma, è possibile invertire l'ordine delle integrazioni nel primo membro di questa uguaglianza, in modo che può scriversi

$$(11) \quad \int_0^z \varphi(y) dy \int_y^z \frac{A(x, y)}{(z-x)^{1-\alpha} (x-y)^\alpha} dx = \int_0^z \frac{f(x)}{(z-x)^{1-\alpha}} dx,$$

e qui la funzione

$$K(z, y) \equiv \int_y^z \frac{A(x, y)}{(z-x)^{1-\alpha} (x-y)^\alpha} dx$$

risulta limitata, come lo è la $A(x, y)$, perchè, se è, in tutto T ,

$$|A(x, y)| \leq N,$$

è pure

$$|K(z, y)| \leq N \int_y^z \frac{dx}{(z-x)^{1-\alpha} (x-y)^\alpha} = \frac{\pi N}{\operatorname{sen} \pi \alpha}.$$

Viceversa, se supponiamo valida la (11) per tutti gli z di $(0, 1)$, la (9) risulta necessariamente verificata per quasi tutti gli x di $(0, 1)$. Ed infatti, posto

$$D(x) \equiv \int_0^x \frac{A(x, y)}{(x-y)^\alpha} \varphi(y) dy - f(x),$$

per il lemma del n.° 1, questa funzione risulta finita per quasi tutti gli x di $(0, 1)$. Sempre in virtù dello stesso lemma, la (11) può scriversi nella forma (10) e quindi nella forma

$$(12) \quad \int_0^z \frac{D(x)}{(z-x)^{1-\alpha}} dx = 0.$$

E siccome la funzione $\frac{D(x)}{(z-x)^{1-\alpha}}$ risulta con ciò integrabile in $(0, z)$, ivi risulta integrabile anche la $D(x)$, potendo scriversi

$$D(x) = \frac{D(x)}{(z-x)^{1-\alpha}} \cdot (z-x)^{1-\alpha},$$

dove $(z-x)^{1-\alpha}$ è funzione finita e continua della x in tutto $(0, z)$. E per essere z un qualsiasi valore di $(0, 1)$, la $D(x)$ risulta integrabile in tutto $(0, 1)$. Dalla (12) otteniamo poi, per ogni t di $(0, 1)$,

$$\int_0^t \frac{dz}{(t-z)^\alpha} \int_0^z \frac{D(x)}{(z-x)^{1-\alpha}} dx = 0,$$

e, per il lemma già usato, è possibile di invertire l'ordine delle integrazioni e scrivere

$$\int_0^t D(x) dx \int_x^t \frac{dz}{(t-z)^\alpha (z-x)^{1-\alpha}} = 0,$$

donde si ha, per la (8),

$$\int_0^t D(x) dx = 0,$$

e quindi $D(x) = 0$ quasi dappertutto in $(0, 1)$, vale a dire si ha la validità della (9) quasi dappertutto.

Possiamo dunque enunciare il seguente risultato:

Nelle ipotesi poste per α , $A(x, y)$, $\varphi(y)$ e $f(x)$, condizione necessaria e sufficiente affinchè la (9) valga per quasi tutti gli x di $(0, 1)$, è che valga la (11) per tutti gli z di $(0, 1)$.

3. Osservazione.

Dall'osservazione già fatta che la funzione $K(z, y)$, che figura sotto il segno di integrale nel primo membro della (11), è limitata, segue che tale primo membro è una funzione della z che tende a zero per $z \rightarrow +0$. Pertanto, se vale la (11), anche il secondo membro di questa equazione deve tendere a zero per $z \rightarrow +0$. Possiamo affermare perciò che, *se non vale la uguaglianza*

$$(13) \quad \lim_{z \rightarrow +0} \int_0^z \frac{f(x)}{(z-x)^{1-\alpha}} dx = 0,$$

la (9) non può essere verificata per quasi tutti gli x di $(0, 1)$, e quindi a fortiori neppure per tutti gli x di $(0, 1)$.

Così, in particolare, se è $f(x) = \frac{1}{x^n}$, con $\alpha \leq n < 1$, la (9) non può essere verificata, vale a dire la (9), riguardata come un'equazione nella $\varphi(y)$, non ammette soluzione integrabile, perchè, avendosi, nel caso attuale,

$$\int_0^z \frac{f(x)}{(z-x)^{1-\alpha}} dx = \int_0^z \frac{dx}{x^n (z-x)^{1-\alpha}} \geq \int_0^z \frac{dx}{x^\alpha (z-x)^{1-\alpha}} = \frac{\pi}{\operatorname{sen} \pi \alpha},$$

la (13) non risulta verificata.

4. L'equazione di Abel.

Nei due lavori già citati, Abel non si limitò a risolvere l'equazione (2), ma risolse l'equazione più generale

$$(14) \quad \int_0^x \frac{\varphi(y)}{(x-y)^\alpha} dy = f(x),$$

con $0 < \alpha < 1$, e dove $\varphi(y)$ è la funzione incognita e la $f(x)$ è una funzione data. Questa equazione è un caso particolare della (9), e l'equazione (11) corrispondente si riduce, in virtù della (8), alla

$$(15) \quad \int_0^z \varphi(y) dy = \frac{\text{sen } \pi \alpha}{\pi} \int_0^z \frac{f(x)}{(z-x)^{1-\alpha}} dx.$$

La formula risolvante data da Abel è precisamente questa che ora abbiamo scritta e di cui la (3) è un caso particolare.

Applicando quanto si è stabilito nel n.° 2, possiamo dire che:

Se è $0 < \alpha < 1$, e $f(x)$ è una funzione integrabile in $(0, 1)$, condizione necessaria e sufficiente affinché una funzione $\varphi(y)$, integrabile in $(0, 1)$, verifichi la (14) per quasi tutti gli x di $(0, 1)$, è che essa verifichi la (15) per tutti gli z di $(0, 1)$.

Il primo membro della (15) è una funzione assolutamente continua della z , in tutto $(0, 1)$, la quale si annulla per $z = 0$; dunque, affinché esista una funzione $\varphi(y)$ che verifichi la (15) in tutto $(0, 1)$, è necessario che l'integrale che figura nel secondo membro della (15) medesima sia una funzione assolutamente continua della z in tutto $(0, 1)$, che si annulli per $z = 0$. Questa condizione è anche sufficiente. Ed infatti, se essa è verificata, si soddisfa alla (15) prendendo $\varphi(y)$ uguale alla derivata del secondo membro dell'equazione stessa, là dove questa derivata esiste finita, e in modo arbitrario altrove. Possiamo perciò affermare che:

Condizione necessaria e sufficiente affinché esista una funzione integrabile $\varphi(y)$ che verifichi la (14) per quasi tutti gli x di $(0, 1)$, è che l'integrale

$$(16) \quad \int_0^z \frac{f(x)}{(z-x)^{1-\alpha}} dx$$

sia una funzione della z assolutamente continua in tutto $(0, 1)$ e nulla per $z = 0$.

Osserviamo che, se la $f(x)$ è una funzione limitata, l'integrale (16), per $z \rightarrow +0$ tende certamente a zero. Ed infatti, se è sempre $|f(x)| \leq M$, è, per $z > 0$,

$$\left| \int_0^z \frac{f(x)}{(z-x)^{1-\alpha}} dx \right| \leq M \int_0^z \frac{dx}{(z-x)^{1-\alpha}} = \frac{M}{\alpha} z^\alpha.$$

Se invece la $f(x)$ è illimitata, l'integrale (16) può non tendere a zero per $z \rightarrow +0$, come precisamente avviene, secondo quanto si è detto nel n.° 3, se è $f(x) = \frac{1}{x^\alpha}$ con $\alpha \leq n < 1$. In questo caso, pertanto, non

può esistere nessuna funzione $\varphi(y)$ che soddisfi alla (14) per tutti gli x di $(0, 1)$, od anche solo per quasi tutti questi valori.

Se prescindiamo dalle funzioni nulle quasi dappertutto, la (15), se ammette una soluzione $\varphi(y)$, ne ammette una sola. Pertanto, *prescindendo dalle funzioni nulle quasi dappertutto, non può esistere più di una funzione integrabile che verifichi la (14) quasi dappertutto in $(0, 1)$.*

5. Lemma II.

Nel n.° precedente, abbiamo trovato la condizione necessaria e sufficiente affinché esista una funzione $\varphi(y)$ che renda soddisfatta la equazione (14) quasi dappertutto. Occorre ora di giungere a condizioni che assicurino l'esistenza di una $\varphi(y)$ che verifichi la (14) per tutti gli x di $(0, 1)$. A questo scopo, dobbiamo premettere il seguente lemma.

Se $\psi(x)$ è una funzione assolutamente continua in $(0, 1)$, la funzione

$$(17) \quad \Phi(z) = \int_0^z \frac{\psi(y)}{(z-y)^\alpha} dy,$$

dove supponiamo $0 < \alpha < 1$, è anch'essa assolutamente continua in $(0, 1)$, e la sua derivata è data quasi dappertutto da

$$(18) \quad \Phi'(z) = \frac{\psi(0)}{z^\alpha} + \int_0^z \frac{\psi'(y)}{(z-y)^\alpha} dy.$$

Ed infatti, integrando per parti il secondo membro di (17), si ha

$$\Phi(z) = \frac{\psi(0)}{1-\alpha} z^{1-\alpha} + \frac{1}{1-\alpha} \int_0^z \psi'(y) (z-y)^{1-\alpha} dy.$$

Ora è, per $0 \leq y < x$,

$$\frac{(z-y)^{1-\alpha}}{1-\alpha} = \int_y^z \frac{dx}{(x-y)^\alpha};$$

è perciò, sostituendo,

$$\Phi(z) = \frac{\psi(0)}{1-\alpha} z^{1-\alpha} + \int_0^z \psi'(y) dy \int_y^z \frac{dx}{(x-y)^\alpha}.$$

Ma per il lemma I, quando in esso si faccia $\varphi(y) = \psi'(y)$ e $A(x, y) = (z-x)^{1-\alpha}$, si ha che, nell'integrale ripetuto ora scritto, si può invertire l'ordine delle integrazioni, e scrivere così

$$(19) \quad \Phi(z) = \frac{\psi(0)}{1-\alpha} z^{1-\alpha} + \int_0^z dx \int_0^x \frac{\psi'(y)}{(x-y)^\alpha} dy.$$

Il primo termine del secondo membro di questa uguaglianza è una funzione assolutamente continua della x in tutto $(0, 1)$, perchè è

$$\frac{z^{1-\alpha}}{1-\alpha} = \int_0^z \frac{dz}{z^\alpha}.$$

Il secondo termine dello stesso secondo membro è pure una funzione assolutamente continua di x in tutto $(0, 1)$ essendo esso una funzione integrale. Dunque la $\Phi(z)$ è essa pure assolutamente continua in tutto $(0, 1)$. Inoltre, la sua derivata è data quasi dappertutto dalla somma delle derivate dei due termini del secondo membro di (19). E ciò prova la (18).

6. L'equazione di Abel, con secondo membro assolutamente continuo.

Supponiamo che la funzione $f(x)$, del secondo membro di (14), sia assolutamente continua in tutto $(0, 1)$. Allora l'integrale (16) risulta senz'altro funzione assolutamente continua della z , in tutto $(0, 1)$, in virtù del lemma del n.º precedente. Inoltre, per essere la $f(x)$ necessariamente limitata, per l'ipotesi qui fatta, l'integrale (16), per un'osservazione fatta nel n.º 4, tende a zero al tendere a zero di z ; dai risultati del n.º 4 segue, pertanto, che *esiste una ed una sola funzione $\varphi(y)$ (a prescindere dalle funzioni nulle quasi dappertutto), integrabile in $(0, 1)$ e soddisfacente alla (14) per quasi tutti i valori di x* . Questa funzione $\varphi(y)$ deve verificare la (15) e quindi è data, quasi dappertutto, dalla derivata del secondo membro di tale uguaglianza, derivata che si ottiene per mezzo del lemma del n.º precedente. È dunque quasi dappertutto

$$(20) \quad \varphi(y) = \frac{\text{sen } \pi \alpha}{\pi} \left\{ \frac{f(0)}{y^{1-\alpha}} + \int_0^y \frac{f'(z)}{(y-z)^{1-\alpha}} dz \right\}.$$

Possiamo ora dimostrare che la funzione $\varphi(y)$, definita da questa uguaglianza, verifica la (14), non soltanto per quasi tutti gli x , ma per tutti gli x di $(0, 1)$.

Occorre, innanzi tutto, provare che l'integrale del primo membro della (14), quando in esso si ponga la funzione $\varphi(y)$ data dalla (20), è finito per ogni valore di x di $(0, 1)$. Ora, per la (20), si ha

$$(21) \quad \int_0^x \frac{\varphi(y)}{(x-y)^\alpha} dy = \frac{\text{sen } \pi \alpha}{\pi} f(0) \int_0^x \frac{dy}{y^{1-\alpha} (x-y)^\alpha} + \frac{\text{sen } \pi \alpha}{\alpha} \int_0^x \frac{dy}{(x-y)^\alpha} \int_0^y \frac{f'(z)}{(y-z)^{1-\alpha}} dz.$$

L'integrale del primo termine del secondo membro di questa uguaglianza è finito per tutti i valori positivi di x , perchè è sempre uguale a $\frac{\pi}{\text{sen } \pi \alpha}$; pertanto, il termine considerato è costantemente uguale a $f(0)$. Anche l'integrale ripetuto dell'ultimo termine della (21) esiste sempre finito, come risulta dal lemma I, quando in esso si faccia $A(x, y) = 1$ e si sostituiscano $\varphi(y)$, x, y, z e α rispettivamente con $f'(z)$, y, z, x e $1 - \alpha$. Dunque l'integrale del primo membro di (21) esiste sempre finito; inoltre, è, per quanto si è detto,

$$\int_0^x \frac{\varphi(y)}{(x-y)^\alpha} dy = f(0) + \frac{\text{sen } \pi \alpha}{\alpha} \int_0^x \frac{dy}{(x-y)^\alpha} \int_0^y \frac{f'(z)}{(y-z)^{1-\alpha}} dz,$$

e, sempre per il lemma I,

$$\int_0^x \frac{\varphi(y)}{(x-y)^\alpha} dy = f(0) + \frac{\text{sen } \pi \alpha}{\pi} \int_0^x f'(z) dz \int_z^x \frac{dy}{(x-y)^\alpha (y-z)^{1-\alpha}}.$$

Tenendo presente la (8), si ha, infine,

$$\int_0^x \frac{\varphi(y)}{(x-y)^\alpha} dy = f(0) + \int_0^x f'(z) dz = f(x).$$

Così risulta provato che la (14) vale per tutti gli x di $(0, 1)$, quando la $\varphi(y)$ sia la funzione data dalla (20). Possiamo dunque concludere che:

Se la funzione $f(x)$ è assolutamente continua in tutto $(0, 1)$, e se è $0 < \alpha < 1$, l'equazione di Abel (14), quando in essa si consideri come incognita la $\varphi(y)$, ammette una ed una sola soluzione integrabile (se si prescinde dalle funzioni nulle quasi dappertutto), e tale soluzione è data dalla formula

$$\varphi(y) = \frac{\text{sen } \pi \alpha}{\pi} \left\{ \frac{f(0)}{y^{1-\alpha}} + \int_0^y \frac{f'(x)}{(y-x)^{1-\alpha}} dx \right\}.$$

In particolare, se consideriamo l'equazione (2) del problema di Abel, abbiamo che:

Se la funzione $f(z)$ è, in un certo intervallo $(0, \delta)$, assolutamente continua, l'equazione (2) ammette, nell'intervallo $(0, \delta)$, una ed una sola soluzione $\varphi(y)$, nel campo delle funzioni integrabili secondo Lebesgue (e quando si prescinde dalle funzioni nulle quasi dappertutto), e tale soluzione è data dalla formula

$$(22) \quad \varphi(y) = \frac{\sqrt{2g}}{\pi} \left\{ \frac{f(0)}{y^{\frac{1}{2}}} + \int_0^y \frac{f'(z)}{(y-z)^{\frac{1}{2}}} dz \right\}.$$

7. Il problema di Abel.

Supposta la $f(z)$ assolutamente continua nell'intervallo $(0, \delta)$, abbiamo ottenuto, con la formula (22), l'unica soluzione esistente della equazione (2). Ma tale soluzione può non essere soluzione del problema di Abel enunciato nell'introduzione. Ed infatti, avendo posto $\varphi(y) = \sqrt{1 + \Phi'^2(y)}$, ne viene che la $\varphi(y)$, per essere soluzione del problema di Abel, deve soddisfare ovunque, oppure quasi dappertutto se la Φ' non esiste sempre finita, alla disuguaglianza

$$(23) \quad \varphi(y) \geq 1;$$

e questa condizione può non essere soddisfatta dalla (22), come avviene per esempio se è $f(z) = \sqrt{\frac{z}{g}}$. In tale caso, infatti, la (22) dà

$$\varphi(y) = \frac{1}{\pi \sqrt{2}} \int_0^y \frac{dz}{\sqrt{z(y-z)}} = \frac{1}{\sqrt{2}}.$$

La condizione (23) è poi anche sufficiente affinché la soluzione dell'equazione (2) sia soluzione del problema di Abel. Ed infatti, se vale la (23), avendosi

$$(24) \quad \Phi'(y) = \sqrt{\varphi^2(y) - 1},$$

il secondo membro di questa uguaglianza è sempre reale e si ha

$$|\Phi'(y)| < \varphi(y),$$

e la $\Phi'(y)$ risulta integrabile come la $\varphi(y)$. Integrando allora la $\Phi'(y)$, si ha la $\Phi(y)$ e quindi l'equazione della curva cercata.

Qui osserviamo che, dalla funzione $\varphi(y)$, definita dalla (22) e supposta soddisfacente alla (23), si possono dedurre infinite funzioni $\Phi(y)$. Ed infatti, essendo la $\Phi'(y)$ definita dalla (24), della $\Phi'(y)$ resta determinato soltanto il valore assoluto, il segno restando arbitrario. Noi intenderemo, nel seguito, di scegliere sempre il segno positivo, e così potremo parlare di funzione $\Phi(y)$ *univocamente determinata dalla $\varphi(y)$* ; e pure in questo senso potremo parlare di unicità della soluzione del problema di Abel.

Ritornando alla condizione (23), rileviamo che essa dà, in virtù della stessa equazione (2),

$$(25) \quad f(z) \geq \frac{1}{\sqrt{2g}} \int_0^z \frac{dy}{\sqrt{z-y}} = \sqrt{\frac{2z}{g}},$$

perciò *il problema di Abel non ha soluzione se non è, in tutto un intervallo $(0, \delta)$, $f(z) \geq \sqrt{\frac{2z}{g}}$.*

Nel caso che sia esattamente $f(z) = \sqrt{\frac{2z}{g}}$, in tutto l'intervallo $(0, \delta)$, la (22) dà costantemente $\varphi(y) = 1$, onde $\Phi(y) = 0$, e la curva di discesa è la verticale.

Per discutere la risolubilità del problema di Abel, distingueremo il caso in cui è $f(0) \neq 0$ da quello in cui è $f(0) = 0$.

a) Ferma l'ipotesi che la $f(z)$ sia assolutamente continua in $(0, \delta)$, supponiamo dapprima che sia $f(0) \neq 0$ e quindi, dato il significato della funzione $f(z)$, $f(0) > 0$. Supponiamo, inoltre, che, in vicinanza del punto $z = 0$, sia sempre (o quasi dappertutto) $f'(z) \geq 0$. Allora la (23) risulta verificata in tutto un intervallo $(0, \bar{\delta})$. Ed infatti, l'ultimo termine della formula (22) è, per la $f'(z) \geq 0$, sempre ≥ 0 per tutti gli y prossimi a $y = 0$, mentre il termine $\frac{\sqrt{2g}}{\pi} f(0) y^{-\frac{1}{2}}$ tende all'infinito, per $y \rightarrow +0$. Dunque, in tutto un intervallo $(0, \bar{\delta})$, vale la (23).

Le ipotesi ora fatte sono verificate, per esempio, se è $f(z) = c$ (cost. > 0). In questo caso, che è quello della *curva tautocrona*, la (22) dà $\varphi(y) = \frac{c}{\pi} \sqrt{\frac{2g}{y}}$, donde

$$s(y) = \int_0^y \varphi(y) dy = \frac{2c}{\pi} \sqrt{y},$$

che è l'equazione caratteristica della *cicloide*.

Per il caso in cui la $f'(z)$ non sia sempre ≥ 0 , in vicinanza del punto $z = 0$, possiamo dire che la conclusione a cui siamo giunti più sopra vale ancora purchè la $f'(z)$ risulti limitata o, più generalmente, purchè $|f'(z)|^{2+\sigma}$, per un $\sigma > 0$, risulti integrabile in un intervallo $(0, \delta_1)$. Ed infatti, in tale caso, si ha, in tutto $(0, \delta_1)$, per la disuguaglianza di Schwarz generalizzata,

$$\begin{aligned} \left| \int_0^y \frac{f'(z)}{(y-z)^{\frac{1}{2}}} dz \right| &\leq \left[\int_0^y |f'(z)|^{2+\sigma} dz \right]^{\frac{1}{2+\sigma}} \left[\int_0^y (y-z)^{-\frac{2+\sigma}{2+2\sigma}} dz \right]^{\frac{1+\sigma}{2+\sigma}} \\ &\leq \left[\frac{2(1+\sigma)}{\sigma} \right]^{\frac{1+\sigma}{2+\sigma}} \delta_1^{\frac{\sigma}{2(2+\sigma)}} \left[\int_0^{\delta_1} |f'(z)|^{2+\sigma} dz \right]^{\frac{1}{2+\sigma}}, \end{aligned}$$

donde segue che l'ultimo termine della (22) resta, in modulo, inferiore ad un numero fisso, in tutto l'intervallo $(0, \delta_1)$, ciò che basta per giustificare la nostra affermazione.

La (23) risulta verificata in tutto un intervallo $(0, \bar{\delta})$ anche nel caso in cui, in vicinanza di $z = 0$, sia $f'(z) \geq -kz^{-\frac{1}{2}}$, con k costante positiva,

perchè, in tale caso, è sempre, in vicinanza del punto $z = 0$,

$$\int_0^y \frac{f'(z)}{(y-z)^{\frac{1}{2}}} dz \geq -k \int_0^y \frac{dz}{\sqrt{z(y-z)}} = -k\pi,$$

e dalla (22) segue ancora $\varphi(y) \rightarrow +\infty$, per $y \rightarrow +0$.

b) Supponiamo ora che sia $f(0) = 0$, sempre ferma restando l'ipotesi che la $f(z)$ sia assolutamente continua in $(0, \delta)$. Per la $f(0) = 0$, la (22) si riduce alla

$$(26) \quad \varphi(y) = \frac{\sqrt{2g}}{\pi} \int_0^y \frac{f'(z)}{\sqrt{y-z}} dz.$$

Per quanto abbiamo già detto, se in tutto un intervallo $(0, \bar{\delta})$ non è sempre $f(z) \geq \sqrt{\frac{2z}{g}}$, la (23) non può essere soddisfatta. Dalla (26) risulta poi subito che la (23) non è soddisfatta se la derivata $f'(z)$ si mantiene limitata in vicinanza del punto $z = 0$, perchè, in tale caso, il secondo membro della (26) tende a zero per $y \rightarrow +0$. Lo stesso dicasi se la $f'(z)$ diventa infinita per $z \rightarrow +0$ di ordine $\gamma < \frac{1}{2}$, o, più generalmente, se è, in vicinanza di $z = 0$, $|f'(z)| < \frac{1}{\sqrt{2gz}}$, perchè, in tale ipotesi, risulta, in vicinanza di $z = 0$,

$$\left| \int_0^y \frac{f'(z)}{(y-z)^{\frac{1}{2}}} dz \right| < \frac{1}{\sqrt{2g}} \int_0^y \frac{dz}{\sqrt{z(y-z)}} = \frac{\pi}{\sqrt{2g}},$$

e perciò $\varphi(y) < 1$.

Se, invece, si ha $f'(z) \geq \frac{1}{\sqrt{2gz}}$ in tutto un intervallo $(0, \bar{\delta})$, in tale intervallo è

$$\int_0^y \frac{f'(z)}{(y-z)^{\frac{1}{2}}} dz \geq \frac{1}{\sqrt{2g}} \int_0^y \frac{dz}{\sqrt{z(y-z)}} = \frac{\pi}{\sqrt{2g}},$$

e perciò $\varphi(y) \geq 1$, vale a dire la (23) risulta soddisfatta in tutto $(0, \bar{\delta})$.

8. Conclusione.

Possiamo riassumere i risultati della discussione fatta con la seguente proposizione:

A. Sia $f(z)$ una funzione assolutamente continua in un intervallo $(0, \delta)$, con $\delta > 0$.

Il problema di Abel, se ha soluzione, ne ha una sola (nel senso indicato nel n.º 7, e non considerando le funzioni non integrabili secondo Lebesgue), ed essa è data dalla formula

$$(27) \quad x = \Phi(y) = \int_0^y \sqrt{\varphi^2(y) - 1} dy,$$

con

$$\varphi(y) = \frac{\sqrt{2g}}{\pi} \left\{ \frac{f(0)}{y^{\frac{1}{2}}} + \int_0^y \frac{f'(z)}{(y-z)^{\frac{1}{2}}} dz \right\}.$$

Se è $f(0) > 0$, la soluzione esiste in un certo intervallo $(0, \bar{\delta})$ ($0 < \bar{\delta} \leq \delta$), purchè, inoltre, sia, in tutto (o in quasi tutto) l'intervallo $(0, \delta)$:

a) $f'(z) \geq 0$;

b) oppure, la $f'(z)$ risulti limitata;

c) oppure, $|f'(z)|^{2+\sigma}$ risulti integrabile per un $\sigma > 0$ (ciò che avviene, in particolare, se la $f'(z)$ diventa ∞ , per $z \rightarrow +0$, di ordine $\gamma < \frac{1}{2}$);

d) oppure sia $f'(z) \geq \frac{-k}{\sqrt{z}}$, con k costante.

Se è $f(0) = 0$, la soluzione esiste in un certo intervallo $(0, \bar{\delta})$, purchè:

e) sia sempre, in $(0, \delta)$, $f'(z) \geq \frac{1}{\sqrt{2gz}}$ (ciò che avviene, in particolare, se la $f'(z)$ diventa ∞ , per $z \rightarrow +0$, di ordine $\gamma > \frac{1}{2}$).

La soluzione non esiste in un intervallo $(0, \bar{\delta})$ se, in tale intervallo, non è sempre $f(z) \geq \sqrt{\frac{2z}{g}}$.

La soluzione non esiste se, essendo $f(0) = 0$, in vicinanza del punto $z = 0$ la $f'(z)$ si mantiene limitata oppure è $|f'(z)| < \frac{1}{\sqrt{2gz}}$, e, in particolare, se $f'(z)$ diventa infinita, per $z \rightarrow +0$, di ordine $\gamma < \frac{1}{2}$.

Va rilevato che, se è $f(z) = \sqrt{\frac{2z}{g}}$, si ha $\Phi(y) = 0$ e la curva di discesa è la verticale; se è, invece, $f(z) = cz^n$, con c costante (> 0) e $n > \frac{1}{2}$, il problema di Abel non ha soluzione.

B. Supponendo sempre la $f(z)$ assolutamente continua, se è $f(0) > 0$ e vale una qualunque delle condizioni a), b), c), d) del teorema A, oppure se è $f(0) = 0$ e vale la condizione e) del teorema A, la soluzione $x = \Phi(y)$ del problema di Abel ha, in un certo intervallo $(0, \delta')$, la derivata (là dove questa derivata esiste) sempre maggiore di 0 (e quindi la tangente alla curva non è mai verticale). È solo eccettuato il caso in cui sia $f(z) = \sqrt{\frac{2z}{g}}$ in tutto un intervallo $(0, \delta')$,

Nel caso eccettuato è, in tutto $(0, \delta')$, $\Phi'(y) = 0$ e perciò $\Phi(y) = 0$.

La proposizione B risulta senz'altro da quanto abbiamo detto nel n.° precedente.

C. Supponiamo che la $f(z)$ sia continua in tutto un intervallo $(0, \delta)$, e che, per ogni z tale che $0 < z \leq \delta$, la $f'(z)$ esista finita e continua. Allora, se è $f(0) > 0$ e vale una qualunque delle condizioni a), b), c), d) del teorema A, oppure se è $f(0) = 0$ e la $f'(z)$ diventa infinita, per $z \rightarrow +0$, di ordine $\gamma > \frac{1}{2}$, la $f(z)$ risulta assolutamente continua in tutto $(0, \delta)$, e la soluzione $x = \Phi(y)$ del problema di Abel è tale che, nell'intervallo in cui essa vale, la derivata $\Phi'(y)$ esiste sempre, è finita e continua per $y > 0$, tende a $+\infty$ per $y \rightarrow +0$ ed ha il valore $+\infty$ per $y = 0$.

Se poi è $f(0) = 0$, con $f'(z) \geq \frac{1}{\sqrt{2gz}}$ per $0 < z \leq \delta$, la $f(z)$ risulta ancora assolutamente continua in tutto $(0, \delta)$, e la soluzione $x = \Phi(y)$ del problema di Abel è tale che, nell'intervallo in cui essa vale, la $\Phi'(y)$ esiste finita e continua per $y > 0$. Se, in particolare, è $f(z) = k\sqrt{z}$ in $(0, \delta)$, con $k \geq \sqrt{\frac{2}{g}}$, risulta

$$\Phi(y) = y \sqrt{g \frac{k^2}{2} - 1},$$

e la derivata $\Phi'(y)$ esiste ed è continua anche per $y = 0$.

Per dimostrare questa proposizione, si osservi che, per le ipotesi fatte, è

$$(28) \quad f(\delta) - f(z) = \int_z^\delta f'(z) dz,$$

per ogni z tale che $0 < z \leq \delta$, ed è anche

$$\lim_{z \rightarrow +0} \int_z^\delta f'(z) dz = f(\delta) - f(0).$$

Di qui e da una qualunque delle condizioni a), b), c), d), e) del teorema A, segue l'integrabilità nel senso del Lebesgue (e quindi l'assoluta integrabilità) della $f'(z)$ in tutto $(0, \delta)$,⁷⁾ e perciò la validità della (28) anche per $z = 0$, ciò che prova l'assoluta continuità della $f(z)$ in tutto $(0, \delta)$.

Osserviamo poi che, per $y > 0$, in tutto l'intervallo in cui la soluzione del problema di Abel esiste, può scriversi

$$(29) \quad \varphi(y) = \frac{\sqrt{2g}}{\pi} \left\{ \frac{f(0)}{y^{\frac{1}{2}}} + \int_0^{\sqrt{y}} \frac{f'(z)}{(y-z)^{\frac{1}{2}}} dz + \int_{\sqrt{y}}^{\delta} \dots \right\},$$

⁷⁾ Ciò è evidente per le condizioni a), b), c), e); per la condizione d) basta osservare che è (essendo $k > 0$) $|f'(z)| \leq f'(z) + \frac{2k}{\sqrt{z}}$, e che, perciò, è pure, per $0 < z \leq \delta$,

$$\int_z^\delta |f'(z)| dz \leq f(\delta) - f(z) + 4k(\sqrt{\delta} - \sqrt{z}).$$

con $0 < \bar{y} < y$. Ora l'integrale

$$(30) \quad \int_0^{\bar{y}} \frac{f'(z)}{(y-z)^{\frac{1}{2}}} dz$$

è una funzione continua della y per tutti gli $y > \bar{y}$: ed infatti, se è $y' > \bar{y}$, si ha

$$\left| \int_0^{\bar{y}} \frac{f'(z)}{(y-z)^{\frac{1}{2}}} dz - \int_0^{\bar{y}} \frac{f'(z)}{(y'-z)^{\frac{1}{2}}} dz \right| \leq \int_0^{\bar{y}} |f'(z)| \left| \frac{1}{(y-z)^{\frac{1}{2}}} - \frac{1}{(y'-z)^{\frac{1}{2}}} \right| dz.$$

Ma, preso un $\varepsilon > 0$ e tenuto fisso y , possiamo determinare un $\eta > 0$ e $< y - \bar{y}$, in modo che dalla $|y' - y| < \eta$ segua

$$\left| \frac{1}{(y-z)^{\frac{1}{2}}} - \frac{1}{(y'-z)^{\frac{1}{2}}} \right| < \varepsilon,$$

per tutti gli z di $(0, \bar{y})$. Abbiamo allora, per tutti gli y' tali che $|y' - y| < \eta$,

$$\left| \int_0^{\bar{y}} \frac{f'(z)}{(y-z)^{\frac{1}{2}}} dz - \int_0^{\bar{y}} \frac{f'(z)}{(y'-z)^{\frac{1}{2}}} dz \right| \leq \varepsilon \int_0^{\bar{y}} f'(z) dz,$$

e questo prova la continuità di (30) nel punto y considerato, e quindi per ogni $y > \bar{y}$.

Per vedere come si comporta l'ultimo termine della (29), indichiamo con N un numero tale che, in (\bar{y}, δ) , sia sempre $|f'(z)| < N$. Allora, per tutti gli y' tali che $|y' - y| < y - \bar{y}$, è

$$\left| \int_{\bar{y}}^{y'} \frac{f'(z)}{(y'-z)^{\frac{1}{2}}} dz \right| < N \int_{\bar{y}}^{y'} \frac{dz}{(y'-z)^{\frac{1}{2}}} < N(y - \bar{y})^{\frac{1}{2}} \int_0^1 \frac{dt}{(1-t)^{\frac{1}{2}}};$$

dunque, per tutti gli y prossimi ad un determinato valore maggiore di 0, l'ultimo termine della (29) può farsi piccolo come vuolsi, prendendo opportunamente \bar{y} .

Dalla (29) risulta pertanto la continuità di $\varphi(y)$ per ogni $y > 0$. Dalla continuità della $\varphi(y)$ segue poi immediatamente la continuità della $\Phi'(y)$ in virtù della (27).

Se è $f(0) > 0$ e vale una qualunque delle condizioni a), b), c), d), del teorema A, si è già veduto nel n.º 7 che la $\varphi(y)$ tende a $+\infty$ per $y \rightarrow +0$. Dunque, nelle ipotesi dette, è anche $\Phi'(y) \rightarrow +\infty$, per $y \rightarrow +0$. Ed è pure, perciò, $\Phi'(0) = +\infty$.

Se, invece, è $f(0) = 0$ e la $f'(z)$ diventa infinita, per $z = 0$, di ordine $\gamma > \frac{1}{2}$, prendendo un numero ν tale che $\frac{1}{2} < \nu < \gamma$, si ha, per tutti

gli z positivi e minori di un certo \bar{z} , $f'(z) > z^{-r}$, e quindi, per tutti gli y positivi e minori di \bar{z} ,

$$\varphi(y) > \frac{\sqrt{2g}}{\pi} \int_0^y \frac{dz}{z^r(y-z)^{\frac{1}{2}}} > \frac{\sqrt{2g}}{\pi} y^{\frac{1}{2}-r} \int_0^y \frac{dz}{\sqrt{z(y-z)}} = \sqrt{2g} y^{\frac{1}{2}-r},$$

donde, per $y \rightarrow +0$, $\varphi(y) \rightarrow +\infty$, e di conseguenza anche $\Phi'(y) \rightarrow +\infty$.
E si ha ancora necessariamente $\Phi'(0) = +\infty$.

Osservazione. La proposizione C vale anche se, invece di supporre che la $f'(z)$ esista finita e continua per ogni z tale che $0 < z \leq \delta$, si suppone che la $f'(z)$ esista finita per ciascuno degli z ora indicati e sia limitata in ogni intervallo (δ_1, δ) tale che $0 < \delta_1 < \delta$.

Bologna, 14 Marzo 1927.

(Eingegangen am 17. 3. 1927.)

Das Anfangswertproblem einer beliebigen nichtlinearen hyperbolischen Differentialgleichung beliebiger Ordnung in zwei Variablen. Existenz, Eindeutigkeit und Abhängigkeitsbereich der Lösung.

Von

Kurt Friedrichs in Aachen und Hans Lewy in Göttingen.

Soll eine partielle Differentialgleichung, die man für die zeitliche Änderung irgendeines physikalischen Zustandes abgeleitet hat, wirklich das Gesetz dieser Zustandsänderung darstellen, so wird man von ihr verlangen, daß sie folgende mathematischen Eigenschaften besitzt. Ist zu irgendeiner Zeit irgendein Zustand vorgegeben, so besitzt die Differentialgleichung eine Lösung, die den Zustand auch für die folgende Zeit darstellt. Und zwar darf es nur eine Lösung geben, denn sonst würde die Differentialgleichung den Zustand nicht vollständig bestimmen. Soll ferner eine räumlich begrenzte Änderung eines vorgegebenen Zustandes sich nur mit endlicher Geschwindigkeit auswirken, so darf der Wert der Lösung in einem Punkte nicht von sämtlichen Anfangswerten abhängen, sondern nur von einem räumlich begrenzten Teil. Von einer solchen Eigenschaft kann man natürlich nur dann sprechen, wenn man auch nicht analytische Anfangswerte zuläßt.

Die linearen hyperbolischen Differentialgleichungen zweiter Ordnung erfüllen bekanntlich diese Forderungen. Für beliebige hyperbolische Differentialgleichungen zweiter Ordnung in zwei Veränderlichen wurde dies Verhalten von H. Lewy¹⁾ nachgewiesen. Für lineare Differentialgleichungen höherer Ordnung mit konstanten Koeffizienten hat G. Herglotz²⁾ explizite Darstellungen der Lösung entwickelt, in denen die gekennzeichnete Art der Abhängigkeit von den Anfangswerten zum Ausdruck kommt. In

¹⁾ H. Lewy, Über das Anfangswertproblem einer hyperbolischen nichtlinearen Differentialgleichung ..., Math. Annalen 98.

²⁾ G. Herglotz, Sächsische Akademieberichte 1926, S. 287.

unserer Arbeit beweisen wir, daß unsere Forderungen für jede hyperbolische Differentialgleichung beliebiger Ordnung in zwei unabhängigen Veränderlichen erfüllt sind.

Das Anfangswertproblem einer Differentialgleichung n -ter Ordnung $F=0$ für eine Funktion $u(x, y)$ der beiden Variablen x, y formulieren wir folgendermaßen. Auf einer Kurve³⁾ in der (x, y) -Ebene seien die Werte der Funktion u und ihrer Ableitungen bis zur n -ten Ordnung derartig vorgegeben, daß sie der gegebenen Differentialgleichung genügen. Längs dieser Kurve sollen für die vorgegebenen Anfangswerte $u, u_x, u_y, u_{xx}, \dots, u_{y^n} = \frac{\partial^n u}{\partial y^n}$ eine Reihe von Streifenbedingungen erfüllt sein, d. h.

$$du = u_x dx + u_y dy,$$

$$du_x = u_{xx} dx + u_{xy} dy,$$

$$\dots \dots \dots$$

$$du_{y^{n-1}} = u_{y^{n-1}x} dx + u_{y^n} dy.$$

Um das Anfangswertproblem zu lösen, leiten wir ein — für eine Lösung u von $F=0$ und ihre Ableitungen u_x, u_y, \dots sicher erfülltes — System von Differentialgleichungen ab, in dem die Größen $x, y, u, u_x, \dots, u_{y^n}$ als unbekannte Funktionen aufgefaßt werden und ihre ersten Ableitungen nach einer „charakteristischen Richtung“ auftreten. Dieses Gleichungssystem nennen wir das System der „charakteristischen Gleichungen“. Wir ersetzen nun unser Anfangswertproblem für $F=0$ durch ein Anfangswertproblem für ein System von charakteristischen Differentialgleichungen. Alle solchen Gleichungen, die sich für eine charakteristische Richtung ergeben, würden nicht ausreichen, um diese Funktionen völlig zu bestimmen. Da es aber, wie wir fordern, n verschiedene reelle charakteristische Richtungen gibt, so können wir die Gesamtheit der charakteristischen Gleichungen, die in allen Richtungen gelten, als ein System partieller Differentialgleichungen auffassen, welches jedoch jetzt mehr Differentialgleichungen als gesuchte Funktionen enthält⁴⁾. Aus dieser Gesamtheit suchen wir ein System von ebensoviel Gleichungen wie Funktionen heraus, das bei vorgegebenen Anfangswerten gerade eine Lösung zuläßt. Wir weisen die Existenz dieser Lösung mit Hilfe eines Differenzenverfahrens nach und zeigen, daß die Lösung des ausgewählten Systems der charakteristischen Gleichungen auch die Lösung des ursprünglichen Anfangswertproblems für $F=0$ liefert.

³⁾ Die Tangente dieser Kurve muß stetig und darf nirgends charakteristisch sein vgl. § 2).

⁴⁾ Vgl. H. Lewy, loc. cit.

Aus unserem Konvergenzbeweis ergibt sich unmittelbar das Resultat, daß der Wert der Lösung in einem Punkte von den Vorgaben nur auf einem Teil der Anfangskurve abhängt; d. h. daß ihr Wert sich nicht ändert, wenn wir die Vorgaben außerhalb dieses Teils der Anfangskurve ändern. Den *Abhängigkeitsbereich* der Lösung erhalten wir folgendermaßen: Wir ziehen von dem betreffenden Punkte aus die charakteristischen Linien, bis sie die Anfangskurve treffen. Das von den beiden *äußersten* Schnittpunkten begrenzte Stück gibt das Abhängigkeitsgebiet.

1. Teil.

§ 1.

Die charakteristischen Differentialgleichungen⁵⁾.

Wir gehen aus von einer Differentialgleichung n -ter Ordnung

$$(1) \quad F(x, y, u, u_x, u_y, u_{xx}, \dots) = 0,$$

in der sämtliche Ableitungen der Funktion $u(x, y)$ bis zur n -ten Ordnung vorkommen können. Für die höchsten Ableitungen führen wir folgende Abkürzungen ein:

$$\begin{aligned} u_x^n &= [0], \\ u_x^{n-1}y &= [1], \\ \dots & \dots \\ u_y^n &= [n]. \end{aligned}$$

Ferner verwenden wir die Abkürzungen:

$$\begin{aligned} \frac{\partial F}{\partial [0]} &= F_0, \\ \dots & \dots \\ \frac{\partial F}{\partial [n]} &= F_n. \end{aligned}$$

Verstehen wir nun unter $u(x, y)$ eine Lösung der Differentialgleichung (1), so muß für sie auch die durch totale Differentiation nach x entstehende Gleichung

$$(2) \quad \frac{dF}{dx} = F_0 [0]_x + F_1 [1]_x \dots + F_n [n]_x + [F]_x = 0$$

erfüllt sein. Dabei bedeutet

$$[F]_x = \frac{\partial F}{\partial x} + \frac{\partial F}{\partial u} u_x + \dots + \frac{\partial F}{\partial u_{y^{n-1}x}} u_{y^{n-1}x}.$$

⁵⁾ Die im folgenden verwandte Methode läßt sich unmittelbar zur Aufstellung der Charakteristikentheorie auch für Differentialgleichungen in mehr unabhängigen Variablen und für Systeme von Differentialgleichungen benutzen.

Wir betrachten nun irgendeine Richtung in der (x, y) -Ebene und bezeichnen die Ableitungen der Ausdrücke x, y, u, u_x, \dots nach einem Parameter in dieser Richtung mit x', y', u', u'_x, \dots . Dann gelten die folgenden Streifenrelationen

$$\begin{aligned}[0]' &= [0]_x x' + [0]_y y', \\ [1]' &= [1]_x x' + [1]_y y', \\ &\dots \\ [n-1]' &= [n-1]_x x' + [n-1]_y y',\end{aligned}$$

die wir wegen

$$[0]_y = [1]_x, [1]_y = [2]_x, \dots$$

auch in der Form

$$(3) \quad \begin{aligned}x'[0]_x + y'[1]_x - [0]' &= 0, \\ x'[1]_x + y'[2]_x - [1]' &= 0, \\ &\dots \\ x'[n-1]_x + y'[n]_x - [n-1]' &= 0\end{aligned}$$

schreiben können. Die n Gleichungen (3) bilden zusammen mit Gleichung (2) ein System linearer Gleichungen für die $n+1$ Größen $[0]_x, [1]_x, \dots, [n]_x$. Im allgemeinen wird die Determinante dieser Gleichungen von null verschieden sein. *Verschwindet aber die Determinante, so verschwindet auch, da das Gleichungssystem ja eine Lösung hat, jede $n+1$ reihige Determinante der Matrix*

$$\begin{vmatrix} F_0 & F_1 & F_2 & \dots & F_{n-1} & F_n & -[F]_x \\ x' & y' & 0 & \dots & 0 & 0 & [0]' \\ 0 & x' & y' & \dots & 0 & 0 & [1]' \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & x' & y' & [n-1]' \end{vmatrix}.$$

Das Verschwinden der Determinante

$$(4) \quad A_n \equiv \begin{vmatrix} F_0 & F_1 & \dots & F_{n-1} & F_n \\ x' & y' & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & x' & y' \end{vmatrix} = 0$$

stellt eine homogene Bedingung für x', y' , d. h. eine Bedingung für die Richtung dar, nach der die durch den Strich angedeuteten Ableitungen zu bilden sind. Das Verschwinden der übrigen $(n+1)$ -reihigen Determinanten, z. B. der durch Auslassung der vorletzten Spalte entstehenden

$$(5) \quad \begin{vmatrix} F_0 & F_1 & \dots & F_{n-1} & -[F]_x \\ x' & y' & \dots & 0 & [0]' \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & x' & [n-1]' \end{vmatrix} = 0$$

stellt dann eine Bedingung für die Ableitungen der Größen $[0], [1], \dots, [n-1]$ nach dieser Richtung dar. Die Gleichung (4) $\Delta_n = 0$ nennen wir die *charakteristische Hauptgleichung* und jede Richtung $y':x'$ der (x, y) Ebene, die ihr genügt, eine *charakteristische Richtung*. Die Gleichung $\Delta_n = 0$ ist vom n -ten Grade in $\rho = y':x'$. Wir verlangen nun von unserer Differentialgleichung $F=0$ und ihrer Lösung u , daß die zugehörige charakteristische Hauptgleichung $\Delta_n = 0$ *n und nur n verschiedene reelle Wurzeln* besitzt^{a)}. Diese Voraussetzung ist nichts anderes als unsere Forderung, daß die Gleichung $F=0$ hyperbolisch ist.

Die Gleichung (5) repräsentiert demgemäß n einzelne Gleichungen, je nachdem auf welche der n charakteristischen Richtungen sich die durch den Strich angedeutete Ableitung bezieht.

Geht man anstatt von der Gleichung (2), die aus $\frac{dF}{dx} = 0$ entstanden ist, von der Gleichung $\frac{dF}{dy} = 0$ aus, und fügt die n Streifenrelationen

$$\begin{aligned} x'[0]_y + y'[1]_y - [1]' &= 0, \\ \dots & \\ x'[n-1]_y + y'[n]_y - [n]' &= 0 \end{aligned}$$

hinzu, so erhält man ein Gleichungssystem für die $n+1$ Unbekannten $[0]_y, [1]_y, \dots, [n]_y$. Seine Matrix lautet:

$$\begin{vmatrix} F_0 & F_1 & \dots & F_{n-1} & F_n & -[F]_y \\ x' & y' & \dots & 0 & 0 & [1]' \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & x' & y' & [n]' \end{vmatrix}.$$

Die Determinante der $n+1$ Gleichungen ist wieder Δ_n . Kennzeichnet der Strich wieder die Ableitung nach einer charakteristischen Richtung, so verschwindet Δ_n , und wir schließen dann wie früher auf das gleichzeitige Bestehen z. B. der Relation

$$(6) \quad \begin{vmatrix} F_0 & F_1 & \dots & F_{n-1} & -[F]_y \\ x' & y' & \dots & 0 & [1]' \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & x' & [n]' \end{vmatrix} = 0,$$

^{a)} Wir setzen voraus, daß für keine charakteristische Richtung $x'=0$ oder $y'=0$ ist, was darauf hinauskommt, daß F_0 und F_n nicht gleich null sind. Da die Größen F_0, F_1, \dots, F_n nicht alle gleich null sein können, läßt sich dies für Umgebungen jedes Punktes immer dadurch erreichen, daß man an Stelle von x, y neue unabhängige Veränderliche ξ, η einführt und $F=0$ als Differentialgleichung für u als Funktion von ξ und η auffaßt; denn dabei gehen die charakteristischen Richtungen in sich über.

die n einzelne Gleichungen repräsentiert, je nachdem auf welche der n charakteristischen Richtungen sich der Strich bezieht.

Die Determinante Δ_n können wir in die Form

$$(7) \quad \frac{\Delta_n}{(x')^n} = \Delta_n(\varrho) = F_0 \varrho^n - F_1 \varrho^{n-1} + \dots + (-1)^n F_n$$

setzen, während wir an Stelle der Gleichungen (5) und (6) auch

$$(8) \quad \Delta_{n-1}(\varrho) [n-1]' - \Delta_{n-2}(\varrho) [n-2]' + \dots + (-1)^{n-1} \Delta_0(\varrho) [0]' - (-1)^n [F]_x x' = 0,$$

$$(9) \quad \Delta_{n-1}(\varrho) [n]' - \Delta_{n-2}(\varrho) [n-1]' + \dots + (-1)^{n-1} \Delta_0(\varrho) [1]' - (-1)^n [F]_y x' = 0$$

schreiben können. Dabei bedeuten $\Delta_0(\varrho)$, $\Delta_1(\varrho)$, ... die Determinanten

$$(10) \quad \begin{aligned} \Delta_0(\varrho) &= F_0, \\ \Delta_1(\varrho) &= \begin{vmatrix} F_0 & F_1 \\ 1 & \varrho \end{vmatrix}, \\ \Delta_2(\varrho) &= \begin{vmatrix} F_0 & F_1 & F_2 \\ 1 & \varrho & 0 \\ 0 & 1 & \varrho \end{vmatrix}, \\ &\dots \end{aligned}$$

Wir merken noch für später die Relationen an:

$$(10a) \quad \Delta_{k+1}(\varrho) - \varrho \Delta_k(\varrho) = (-1)^{k+1} F_{k+1}.$$

§ 2.

Die Bedeutung der Bedingung, daß die Anfangskurve nicht charakteristisch ist.

Es ist nach dem Vorgehenden plausibel, welche Bedingung man für die Anfangskurve $x(t)$, $y(t)$ stellen wird, wenn man in dem in der Einleitung formulierten Anfangswertproblem für $F=0$ beliebige Werte von u und seinen Ableitungen bis zur n -ten Ordnung vorgeben will, die nur der Gleichung $F=0$ und den Streifenbedingungen längs der Anfangskurve zu genügen brauchen. Wir werden nämlich verlangen, daß die Anfangskurve nirgends charakteristisch ist, d. h. daß auf ihr der durch Gleichung (4) definierte Ausdruck Δ_n nirgends verschwindet. Würde nämlich Δ_n an einer Stelle der Anfangskurve gleich null sein, so müßten dort auch noch die Gleichungen (5) und (6) erfüllt sein, wenn es eine Lösung von $F=0$ mit diesen Anfangswerten gibt. Das würde aber im allgemeinen eine neue Bedingung für die Anfangswerte darstellen.

§ 3.

Das Anfangswertproblem der charakteristischen Differentialgleichungen.

Eine Lösungsfläche $u(x, y)$ unserer Differentialgleichung $F = 0$ wird von n Scharen von charakteristischen Linien überdeckt. An Stelle der Linien $x = \text{konst.}$ und $y = \text{konst.}$ führen wir die beiden „äußersten“ dieser Charakteristikenscharen als Koordinatenlinien $\alpha = \text{konst.}$ und $\omega = \text{konst.}$ ein. Hierbei bezieht sich der Begriff der beiden äußersten Scharen auf die Anfangslinie, und zwar in folgendem Sinne. Denken wir uns in einem Punkte P der Anfangslinie eine Richtung aus der Richtung der Anfangslinie — die selbst nicht charakteristisch ist — um zwei Rechte gedreht, so wird sie dabei einmal mit sämtlichen n charakteristischen Richtungen zusammenfallen. Die erste und die letzte, mit der sie zusammenfällt, nennen wir die beiden äußersten charakteristischen Richtungen. Lassen wir den Punkt P auf der Anfangslinie wandern, so ändert sich jede charakteristische Richtung stetig²⁾, und es gehen äußerste Richtungen in äußerste Richtungen über.

Wir wählen die Parameter α und ω so, daß in einer Ebene mit den rechtwinkligen Koordinaten α, ω der Geraden $\alpha + \omega = 0$ die Anfangskurve entspricht. Wir fassen nun die Größen $x, y, u, u_x, \dots, [n]$ als voneinander unabhängige Funktionen der Parameter α und ω auf; d. h. wir fordern nicht, daß $u_x = \frac{\partial u}{\partial x}, \dots, [n] = \frac{\partial^n u}{\partial y^n}$ ist. Dann stellen die charakteristischen Gleichungen (4), (5) und (6) und die Streifengleichungen

$$\begin{aligned} u' &= u_x x' + u_y y' \\ &\dots \dots \dots \\ u'_{y^{n-1}} &= [n-1] x' + [n] y' \end{aligned}$$

ein System von Differentialgleichungen erster Ordnung für diese Funktionen dar. Wir können nämlich den Strich einmal als Ableitung nach dem Parameter α , das andere Mal als Ableitung nach ω deuten; der Strich kann aber auch die Ableitungen nach den $n-2$ anderen charakteristischen Richtungen andeuten; diese setzen sich aus den Ableitungen nach α und ω linear zusammen.

Wir können jetzt das Anfangswertproblem für die Differentialgleichung $F = 0$ ersetzen durch das Anfangswertproblem des Systems der charakteristischen Differentialgleichungen erster Ordnung für die Funktionen $x, y, u, \dots, [n]$ der Variablen α und ω ; natürlich mit denselben Anfangsbedingungen auf der Geraden $\alpha + \omega = 0$. Die Anzahl der charakteristischen Gleichungen und Streifenbedingungen ist aber größer als die Anzahl der

²⁾ Falls die Anfangskurve eine stetige Tangente besitzt und die Größen F_0, F_1, \dots, F_n längs der Anfangskurve stetig sind.

gesuchten Funktionen, so daß wir uns veranlaßt sehen, ein System von ebensoviel Gleichungen als gesuchten Funktionen auszuwählen. Wir können das so einrichten, daß das Gleichungssystem in den Ableitungen linear wird. Dies geschieht auf folgende Weise.

Wir denken uns die charakteristische Hauptgleichung (4) in ihre Linearfaktoren zerlegt:

$$A_n = F_0 (y' - \varrho_\alpha x') (y' - \varrho_\beta x') \dots (y' - \varrho_\omega x') = 0.$$

Die Größen $\varrho_\alpha, \varrho_\beta, \dots, \varrho_\omega$ sind also bekannte Funktionen der gesuchten Funktionen $x, y, \dots, [n]$. Die Indizes sind dabei so gewählt, daß der Linearfaktor $y' - \varrho_\alpha x'$ bzw. $y' - \varrho_\omega x'$ verschwindet, wenn für y' und x' die Ableitungen nach der Richtung α bzw. ω eingesetzt werden. Wir bezeichnen mit dem oberen rechten Index α oder ω die Ableitungen nach dem Parameter α bzw. ω und wählen die beiden Gleichungen

$$(11) \quad \boxed{\begin{array}{l} y^\alpha - \varrho_\alpha x^\alpha = 0 \\ y^\omega - \varrho_\omega x^\omega = 0 \end{array}}$$

für unser zu bildendes Gleichungssystem aus.

Die $n - 2$ anderen Linearfaktoren von A_n dienen uns dazu, Ableitungen nach den $n - 2$ anderen charakteristischen Richtungen — die wir entsprechend durch Indizes β, γ, \dots kennzeichnen wollen — durch die Ableitungen nach α und ω linear auszudrücken. Ist nämlich $f(\alpha, \omega)$ irgendeine Funktion von α und ω , so ist z. B. $f^\beta = f^\alpha \alpha^\beta + f^\omega \omega^\beta$. Das Verhältnis $\alpha^\beta : \omega^\beta$ berechnen wir dabei aus der Gleichung

$$y^\beta - \varrho_\beta x^\beta = (y^\alpha - \varrho_\beta x^\alpha) \alpha^\beta + (y^\omega - \varrho_\beta x^\omega) \omega^\beta = 0.$$

Ferner wählen wir die n Gleichungen aus, die aus der Gleichung (8) entstehen, indem wir ϱ und den Strich auf die n verschiedenen charakteristischen Richtungen $\alpha, \beta, \dots, \omega$ beziehen und demgemäß schreiben

$$(12) \quad \boxed{\begin{array}{l} \Delta_{n-1}(\varrho_\alpha) [n-1]^\alpha - \Delta_{n-2}(\varrho_\alpha) [n-2]^\alpha + \dots - (-1)^n [F]_x x^\alpha = 0 \\ \Delta_{n-1}(\varrho_\beta) [n-1]^\beta - \Delta_{n-2}(\varrho_\beta) [n-2]^\beta + \dots - (-1)^n [F]_x x^\beta = 0 \\ \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \\ \Delta_{n-1}(\varrho_\omega) [n-1]^\omega - \Delta_{n-2}(\varrho_\omega) [n-2]^\omega + \dots - (-1)^n [F]_x x^\omega = 0 \end{array}}$$

Weiter benutzen wir die Gleichung (9) in einem einzigen Exemplar, nämlich dem auf α bezüglichen:

$$(13) \quad \boxed{\Delta_{n-1}(\varrho_\alpha) [n]^\alpha - \Delta_{n-2}(\varrho_\alpha) [n-1]^\alpha + \dots - (-1)^n [F]_y x^\alpha = 0}$$

Endlich benutzen wir noch sämtliche Streifengleichungen in der Richtung α für die Größen $u, u_x, u_y, \dots, u_{x^{n-1}}, u_{x^{n-2}y}, \dots, u_{y^{n-1}}$, die — es sei noch einmal betont — als voneinander unabhängige Funktionen (vgl. S. 206) von α und ω fungieren:

$$(14) \quad \begin{array}{l} \alpha U \equiv u^\alpha - u_x x^\alpha - u_y y^\alpha = 0 \\ \vdots \\ \alpha U_{y^{n-1}} \equiv u_{y^{n-1}}^\alpha - [n-1] x^\alpha - [n] y^\alpha = 0 \end{array}$$

Hier ist $u^\alpha = \frac{\partial u}{\partial \alpha}, \dots, u_{y^{n-1}}^\alpha = \frac{\partial u_{y^{n-1}}}{\partial \alpha}$ gesetzt.

Dieses lineare partielle Differentialgleichungssystem (11), (12), (13), (14) mit ebensoviel Gleichungen wie gesuchten Funktionen hat bei dreimal beschränkt differenzierbar auf der Anfangslinie $\alpha + \omega = 0$ vorgegebenen Anfangswerten in einer gewissen Umgebung dieser Linie eine eindeutige mit stetigen Ableitungen bis zur zweiten Ordnung versehene Lösung: $x, y, u, \dots, [n]$ als Funktionen von α, ω . Unser Gleichungssystem ist nämlich ein Spezialfall eines allgemeineren solchen Systems, von dem wir in Teil II die Existenz der Lösung beweisen, unter der Voraussetzung, daß seine *Determinante* (vgl. § 4) auf der Anfangslinie *nicht verschwindet*.

Es wird sich schließlich zeigen, daß die Lösung des Anfangswertproblems der charakteristischen Differentialgleichungen die gesuchte Lösung der Gleichung $F = 0$ liefert.

§ 4.

Die Determinante verschwindet nicht.

Die Determinante unseres Gleichungssystems erklären wir folgendermaßen: Wir setzen in unsere Differentialgleichungen für die verschiedenen Ableitungen von x , nämlich $x^\alpha, x^\beta, \dots, x^\omega$, immer denselben Buchstaben ξ , für $y^\alpha, y^\beta, \dots, y^\omega$ denselben Buchstaben η , und entsprechend für die Ableitungen aller unserer gesuchten Funktionen. Wir bekommen so ein lineares homogenes Gleichungssystem für die „Unbekannten“ ξ, η, \dots ; seine Determinante bezeichnen wir als die Determinante unseres Differentialgleichungssystems. Sie wird von null verschieden sein, sobald wir wissen, daß dieses lineare Gleichungssystem *nur die eine* Auflösung $\xi = 0, \eta = 0, \dots$ besitzt.

Um dies einzusehen, beachten wir, daß aus den beiden aus (11) entstehenden Gleichungen

$$\eta - \varrho_\alpha \xi = 0 \quad \text{und} \quad \eta - \varrho_\omega \xi = 0$$

$\eta = \xi = 0$ folgt, da ϱ_α nicht gleich ϱ_ω ist. Benutzen wir dieses Resultat,

so sagen die Streifenbedingungen (14) nichts anderes aus, als daß die dort auftretenden den Größen $u^\alpha, \dots, u_{y^{n-1}}^\alpha$ entsprechenden Unbekannten null sind. Wir haben nur noch das Verschwinden derjenigen Unbekannten nachzuweisen, die den Ableitungen von $[0], [1], \dots$ entsprechen. Dazu schreiben wir uns die Determinante der Gleichungen (12) auf:

$$(15) \quad \begin{vmatrix} \Delta_{n-1}(\varrho_\alpha) & \Delta_{n-2}(\varrho_\alpha) & \dots & \Delta_0(\varrho_\alpha) \\ \Delta_{n-1}(\varrho_\beta) & \Delta_{n-2}(\varrho_\beta) & \dots & \Delta_0(\varrho_\beta) \\ \dots & \dots & \dots & \dots \\ \Delta_{n-1}(\varrho_\omega) & \Delta_{n-2}(\varrho_\omega) & \dots & \Delta_0(\varrho_\omega) \end{vmatrix}$$

$\Delta_k(\varrho)$ ist ein Polynom k -ten Grades in ϱ , und der Koeffizient der höchsten Potenz ist F_0 (vgl. (10)). Die Determinante (15) läßt sich also auf die Form bringen:

$$F_0^n \cdot \begin{vmatrix} \varrho_\alpha^{n-1} & \varrho_\alpha^{n-2} & \dots & \varrho_\alpha & 1 \\ \varrho_\beta^{n-1} & \varrho_\beta^{n-2} & \dots & \varrho_\beta & 1 \\ \dots & \dots & \dots & \dots & \dots \\ \varrho_\omega^{n-1} & \varrho_\omega^{n-2} & \dots & \varrho_\omega & 1 \end{vmatrix}$$

Sie ist von null verschieden, da die n Größen $\varrho_\alpha, \varrho_\beta, \dots, \varrho_\omega$ voneinander verschieden sind und $F_0 \neq 0$ vorausgesetzt wurde. (Vgl. Anm. *) S. 204.) Hieraus folgt, daß die den Ableitungen von $[0], [1], \dots$ entsprechenden Unbekannten verschwinden mit Ausnahme der letzten, die der Ableitung von $[n]$ entspricht. Diese kommt in der Gleichung (13) mit dem Koeffizienten $\Delta_{n-1}(\varrho_\alpha)$ vor. Es ist aber $\varrho_\alpha \cdot \Delta_{n-1}(\varrho_\alpha) + (-1)^n F_n = \Delta_n(\varrho_\alpha) = 0$ (vgl. (10a) und (4)), und unsere letzte Unbekannte verschwindet also auch, da wir $F_n \neq 0$ vorausgesetzt haben.

Damit ist also bewiesen, daß die Determinante des ausgewählten Gleichungssystems auf der Anfangslinie von null verschieden ist.

§ 5.

Die Lösung des charakteristischen Systems liefert die Lösung der Gleichung $F=0$.

Um nachzuweisen, daß die Lösungen des Anfangswertproblems unserer charakteristischen Differentialgleichungen auch die Lösungen des Anfangswertproblems von $F=0$ liefern, haben wir zwei Dinge zu zeigen.

1. Die Gleichung $F=0$ ist überall für unsere Funktionen x, y, u, u_x, \dots erfüllt.

2. Die Funktion $u_{y^k x^k}$ ist wirklich die entsprechende Ableitung von

u nach x und y ⁸⁾; dazu genügt es nachzuweisen, daß unsere Streifenrelationen, die in Richtung α erfüllt sind (vgl. (14)), auch nach der ω -Richtung gelten:

$$\begin{aligned} {}^\omega U &\equiv u^\omega - u_x x^\omega - u_y y^\omega = 0 \\ &\dots \dots \dots \dots \dots \dots \dots \\ {}^\omega U_{x^{n-1}} &= 0 \quad \dots \quad {}^\omega U_{y^{n-1}} = 0. \end{aligned}$$

Um $F = 0$ nachzuweisen, addieren wir zur Gleichung (12)_a die mit ϱ_α multiplizierte Gleichung (13). Beachten wir die Relationen (vgl. (10a))

$$(16) \quad \Delta_{k+1}(\varrho_\alpha) - \varrho_\alpha \Delta_k(\varrho_\alpha) = (-1)^{k+1} F_{k+1}$$

und $\Delta_n(\varrho_\alpha) = 0$, so erhalten wir

$$\begin{aligned} 0 &= F_n [n]^\alpha + F_{n-1} [n-1]^\alpha + \dots + F_0 [0]^\alpha + [F]_x x^\alpha + [F]_y y^\alpha \\ &\equiv F_n [n]^\alpha + F_{n-1} [n-1]^\alpha + \dots + F_0 [0]^\alpha \\ &\quad + F_x x^\alpha + F_y y^\alpha + F_u u_x x^\alpha + F_u u_y y^\alpha + \dots \end{aligned}$$

Der letzte Ausdruck ist aber, wie man unter Beachtung der Streifenrelationen (14) erkennt, die Ableitung von F nach dem Parameter α . Da aber F für die Anfangswerte verschwindet, verschwindet es demnach für unsere Lösung überhaupt.

Um die Gültigkeit der Streifenrelationen nach der ω -Richtung abzuleiten, setzen wir voraus, daß alle gesuchten Funktionen $x, y, \dots, [n]$ stetige gemischte zweite Ableitungen nach α und ω besitzen. Unter dieser Voraussetzung gelten — wenn wir für v irgendeine unserer Funktionen einsetzen — folgende identischen Relationen⁹⁾:

$$(17) \quad \begin{aligned} {}^\alpha V^\omega - {}^\omega V^\alpha &\equiv v_x^\alpha x^\omega + v_y^\alpha y^\omega - v_x^\omega x^\alpha - v_y^\omega y^\alpha, \\ {}^\alpha V^\omega - {}^\omega V^\alpha &\equiv {}^\alpha V_x x^\omega + {}^\alpha V_y y^\omega - {}^\omega V_x x^\alpha - {}^\omega V_y y^\alpha, \end{aligned}$$

$$(18) \quad {}^\alpha V^\omega - {}^\omega V^\alpha \equiv \left(\frac{\partial}{\partial y} v_x - \frac{\partial}{\partial x} v_y \right) (y^\alpha x^\omega - x^\alpha y^\omega).$$

In (18) sind v_x und v_y als Funktionen von x und y gedacht (vgl. Anm. 8)).

Für die Größen ${}^\omega U, {}^\omega U_x, \dots, {}^\omega U_{y^{n-1}}$, deren Verschwinden wir nachweisen wollen, können wir nun ein gewöhnliches Differentialgleichungssystem aufstellen. Zu dem Zweck betrachten wir die Gleichung (12)_a, die wir wieder in der alten Form (5)

⁸⁾ Man kann an Stelle von α und ω die Größen x, y als unabhängige Variable einführen, da $\frac{\partial(x, y)}{\partial(\alpha, \omega)} = x^\alpha x^\omega (\varrho_\omega - \varrho_\alpha) \neq 0$ ist.

⁹⁾ ${}^\alpha V^\omega$ bedeutet $\frac{\partial}{\partial \omega} ({}^\alpha V)$; entsprechend später ${}^\alpha U_x^\omega = \frac{\partial}{\partial \omega} ({}^\alpha U_x), \dots$

$$(19) \quad \begin{vmatrix} F_0 & F_1 & \dots & F_{n-1} & -[F]_x \\ x^\alpha & y^\alpha & \dots & 0 & [0]^\alpha \\ \cdot & \cdot & \dots & \cdot & \cdot \\ 0 & 0 & \dots & x^\alpha & [n-1]^\alpha \end{vmatrix} = 0$$

aufgeschrieben denken. Von der letzten Spalte subtrahieren wir die nullte, erste, ..., $(n-1)$ -te Spalte, nachdem wir sie bzw. mit $[0]_x$, $[1]_x$, ..., $[n-1]_x$ multipliziert haben. Außerdem subtrahieren wir die in der letzten Spalte mit $[n]_x$ multiplizierte Determinante (4):

$$(20) \quad \begin{vmatrix} F_0 & F_1 & \dots & F_{n-1} & F_n \\ x^\alpha & y^\alpha & \dots & 0 & 0 \\ \cdot & \cdot & \dots & \cdot & \cdot \\ 0 & 0 & \dots & x^\alpha & y^\alpha \end{vmatrix}.$$

Dadurch wird der Wert von (19) nicht verändert, weil die Determinante (20) verschwindet.

Wir erhalten so aus (19), indem wir $[0]^\alpha = [0]_x x^\alpha + [0]_y y^\alpha, \dots$ ausschreiben,

$$(21) \quad \begin{vmatrix} F_0 & \dots & F_{n-1} & -(F_x + F_u u_x + \dots + F_0 [0]_x + \dots + F_n [n]_x) \\ x^\alpha & \dots & 0 & ([0]_y - [1]_x) y^\alpha \\ \cdot & \dots & \cdot & \cdot \\ 0 & \dots & x^\alpha & ([n-1]_y - [n]_x) y^\alpha \end{vmatrix} = 0.$$

Das oberste Glied der letzten Spalte unterscheidet sich um

$$F_u \left(u_x - \frac{\partial u}{\partial x} \right) + \dots + F_{u_{y^{n-1}x}} \left(u_{y^{n-1}x} - \frac{\partial}{\partial x} u_{y^{n-1}} \right)$$

von dem Ausdruck $-\frac{dF}{dx}$, der nach der oben bewiesenen Gleichung $F=0$ verschwindet.

Die hier auftretenden Größen $u_x - \frac{\partial u}{\partial x}, \dots, u_{y^{n-1}x} - \frac{\partial}{\partial x} u_{y^{n-1}}$ lassen sich durch die Größen ${}^oU, \dots, {}^oU_{y^{n-1}}$ ausdrücken. Es ist nämlich identisch (vgl. (14) und S. 210)

$${}^oU y^\omega - {}^oU y^\alpha = \left(u_x - \frac{\partial u}{\partial x} \right) (x^\omega y^\alpha - y^\omega x^\alpha)$$

.....

$${}^oU_{y^{n-1}} y^\omega - {}^oU_{y^{n-1}} y^\alpha = \left(u_{y^{n-1}x} - \frac{\partial}{\partial x} u_{y^{n-1}} \right) (x^\omega y^\alpha - y^\omega x^\alpha),$$

und die Größen ${}^oU, \dots, {}^oU_{x^{n-1}}$ verschwinden (vgl. (14)).

Die Größen $\frac{\partial}{\partial y}[0] - \frac{\partial}{\partial x}[1], \dots, \frac{\partial}{\partial y}[n-1] - \frac{\partial}{\partial x}[n]$ drücken wir vermöge der Gleichung (18) durch ${}^{\omega}U_{x^{n-1}}, \dots, {}^{\omega}U_{y^{n-1}}$ aus, indem wir dort bzw. $v = u_{x^{n-1}}, \dots, v = u_{y^{n-1}}$ setzen und ${}^{\alpha}U_{x^{n-1}} = 0, \dots, {}^{\alpha}U_{y^{n-1}} = 0$ beachten.

Machen wir in der Gleichung (21) die eben angegebenen Einsetzungen, und multiplizieren wir sie noch mit $\frac{1}{y^\alpha} (x^\omega y^\alpha - y^\omega x^\alpha)$, so erhalten wir die Gleichung

$$\begin{vmatrix} F_0 & \dots & F_{n-1} & F_u {}^\omega U + F_{u_x} {}^\omega U_x + \dots + F_{u_{y^{n-1}}} {}^\omega U_{y^{n-1}} \\ x^\alpha & \dots & 0 & - {}^\omega U_{x^{n-1}} \\ \dots & \dots & \dots & \dots \\ 0 & \dots & x^\alpha & - {}^\omega U_{y^{n-1}} \end{vmatrix} = 0,$$

die wir auch in der Form - vgl. die Ableitung der Gleichung (12)_α aus (5) -

$$(22)_\alpha \quad \Delta_{n-1}(\varrho_\alpha) {}^\omega U_{y^{n-1}}^\alpha - \Delta_{n-2}(\varrho_\alpha) {}^\omega U_{y^{n-2}x}^\alpha + \dots - (-1)^n (F_u {}^\omega U + \dots) x^\alpha = 0$$

schreiben können. Gehen wir an Stelle von (12)_α von (12)_β, ..., (12)_ω aus, so erhalten wir auf dieselbe Weise die Gleichungen

$$(22)_\beta \quad \Delta_{n-1}(\varrho_\beta) {}^\omega U_{y^{n-1}}^\beta - \Delta_{n-2}(\varrho_\beta) {}^\omega U_{y^{n-2}x}^\beta + \dots - (-1)^n (F_u {}^\omega U + \dots) x^\beta = 0$$

$$(22)_\omega \quad \Delta_{n-1}(\varrho_\omega) {}^\omega U_{y^{n-1}}^\omega - \Delta_{n-2}(\varrho_\omega) {}^\omega U_{y^{n-2}x}^\omega + \dots - (-1)^n (F_u {}^\omega U + \dots) x^\omega = 0.$$

Schließlich betrachten wir sämtliche Gleichungen, die wir aus (17) dadurch erhalten, daß wir für v der Reihe nach die Größen $u, u_x, \dots, u_{y^{n-2}}$ einsetzen und die Gleichungen (14) berücksichtigen, nämlich

$$(23) \quad \begin{aligned} {}^\omega U^\alpha &= {}^\omega U_x x^\alpha + {}^\omega U_y y^\alpha \\ \dots & \dots \dots \dots \dots \dots \dots \dots \\ {}^\omega U_{y^{n-2}}^\alpha &= {}^\omega U_{y^{n-2}x} x^\alpha + {}^\omega U_{y^{n-1}} y^\alpha. \end{aligned}$$

Die Gleichungen (22) und (23) sind in der Tat lineare homogene Differentialgleichungen erster Ordnung für die Größen ${}^\omega U, \dots, {}^\omega U_{y^{n-1}}$, und zwar genau soviele Gleichungen wie Unbekannte. Es sind *gewöhnliche* Differentialgleichungen, da sie nur die Ableitungen nach α enthalten. Die Determi-

nante dieses Gleichungssystems, d. h. die Determinante der Koeffizienten der *Ableitungen* der gesuchten Funktionen, ist offenbar das Produkt der Determinante der Gleichungen (22) und des Gleichungssystems (23). Die Determinante des letzteren Systems hat den Wert 1, und die Determinante der Gleichungen (22) ist gerade die Determinante (15).

Wir haben somit ein lineares homogenes gewöhnliches Differentialgleichungssystem mit nicht verschwindender Determinante für unsere Funktionen ${}^{\omega}U, \dots, {}^{\omega}U_{y^{n-1}}$ aufgestellt. Ihre Anfangswerte sind aber null. (Denn auf der Anfangslinie sind die Streifenrelationen nach der Richtung der Anfangslinie vgl. (§ 2) und nach der von ihr verschiedenen (vgl. § 2) α -Richtung erfüllt, also auch in der ω -Richtung.) Jedes stetige Lösungssystem solcher Differentialgleichungen mit verschwindenden Anfangswerten verschwindet aber überall. Damit sind sämtliche Streifenrelationen als erfüllt erkannt, und die Größen $u_x, \dots, u_{y^n} = [n]$ sind in der Tat die entsprechenden Ableitungen der Funktion $u(x, y)$.

Im ersten Teil dieser Arbeit haben wir somit für eine Lösung unserer Differentialgleichung $F = 0$ ein System von charakteristischen Differentialgleichungen erster Ordnung abgeleitet, denen sie genügen muß. Aus diesen charakteristischen Differentialgleichungen konnten wir ein System mit nicht verschwindender Determinante auswählen, dessen Lösung bei denselben Anfangsbedingungen zu einer Lösung der Gleichung $F = 0$ führt. Daß dieses System der ausgewählten Differentialgleichungen *wirklich eine Lösung besitzt, und nur eine einzige* in einem gewissen Bereiche, wird sich aus dem allgemeinen Existenzsatze des zweiten Teiles ergeben.

Aus ihm folgt, daß es ein sich an die Anfangskurve anschließendes Gebiet der (x, y) -Ebene gibt, innerhalb dessen die Differentialgleichung $F = 0$ hyperbolisch bleibt und wo eine und nur eine mit stetigen Ableitungen bis zur $n + 2$ -ten Ordnung versehene Lösung unseres Anfangswertproblems existiert; ferner, daß der Wert in einem Punkte P nur von demjenigen Teil der Anfangskurve abhängt, der von den beiden äußersten Charakteristiken durch P aus ihr ausgeschnitten wird.

Zum Schluß stellen wir die *Voraussetzungen* zusammen, die wir machen, ohne zu behaupten, daß wir nicht mit etwas schwächeren Bedingungen hätten auskommen können. Der für die Anfangswerte als hyperbolisch vorausgesetzte Ausdruck F soll nach seinen Argumenten $x, y, u, \dots, [n]$ viermal beschränkt differenzierbar sein. Die Anfangswerte $x, y, u, \dots, [n]$ sollen als Funktionen eines Parameters auf der nirgends charakteristisch gerichteten Anfangskurve dreimal beschränkt differenzierbar sein und dort der Gleichung $F = 0$ und den Streifenrelationen längs der Kurve genügen.

2. Teil.

Bevor wir unseren allgemeinen Existenzsatz formulieren, behandeln wir einen typischen *Spezialfall*. Wir betrachten ein homogenes lineares Differentialgleichungssystem erster Ordnung für drei Funktionen u, v, w , zweier Variablen α und ω einer rechtwinkligen α, ω -Ebene:

$$(24) \quad \begin{aligned} a_{11} u^\alpha + a_{12} v^\alpha + a_{13} w^\alpha &= 0 \\ a_{21} u^\beta + a_{22} v^\beta + a_{23} w^\beta &= 0 \\ a_{31} u^\omega + a_{32} v^\omega + a_{33} w^\omega &= 0. \end{aligned}$$

Die Größen a_{ik} sollen dreimal beschränkt differenzierbare Funktionen der gesuchten Funktionen u, v, w sein. $u^\alpha, v^\alpha, w^\alpha$ und $u^\omega, v^\omega, w^\omega$ bedeuten die Ableitungen von u, v, w nach α und ω , während $u^\beta, v^\beta, w^\beta$ die linearen Kombinationen

$$\begin{aligned} u^\beta &= \sigma u^\alpha + \tau u^\omega \\ v^\beta &= \sigma v^\alpha + \tau v^\omega \\ w^\beta &= \sigma w^\alpha + \tau w^\omega \end{aligned}$$

sind, wo σ und τ dreimal beschränkt differenzierbare Funktionen von u, v und w sind.

Das Anfangswertproblem, das wir lösen wollen, besteht darin, daß man auf einem Stück der Geraden $\alpha + \omega = 0$ die Werte von u, v, w dreimal beschränkt differenzierbar vorgibt, so daß dort die Determinante

$$|a_{il}| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

nicht verschwindet. Die Größen $u^\beta, v^\beta, w^\beta$ können wir als die Ableitungen von u, v, w nach einer dritten Richtung in der (α, ω) -Ebene auffassen. Wir stellen die Forderung, daß für Punkte der Anfangsgeraden $\alpha + \omega = 0$ diese dritte Richtung zwischen der α - und der ω -Richtung liegt, d. h. daß die α - und ω -Richtung äußerste Richtungen sind (vgl. § 3 S. 206). Diese Forderung ist damit gleichbedeutend, daß auf der Anfangsgeraden die Größen σ und τ positiv sind.

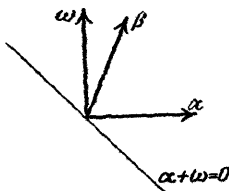


Fig. 1.

Zur Lösung des Anfangswertproblems^{9a)} überdecken wir die (α, ω) -Ebene mit einem quadratischen Gitter, das aus den Geraden $\alpha = nh$,

^{9a)} Während der Korrektur machte uns Herr Courant darauf aufmerksam, daß dieses Problem auch durch sukzessive Approximationen gelöst werden kann, indem man analog wie im folgenden vorgeht.

$\omega = mk$ besteht. $m, n = \dots - 1, 0, +1, \dots$. Es sei ABB_1 ein rechtwinkliges Dreieck aus unserem Gitter (vgl. Fig. 3), dessen Seite BA in der positiven α -Richtung zeigt, B_1A in der positiven ω -Richtung, und wo die Länge dieser Seiten die Maschenweite h ist. Dann ersetzen wir unser Differentialgleichungssystem durch das folgende Differenzgleichungssystem:

$$\begin{aligned}
 & a_{11B}(u_A - u_B) + a_{12B}(v_A - v_B) + a_{13B}(w_A - w_B) = 0 \\
 & a_{21B}[(\sigma_B + \tau_B)u_A - \sigma_B u_B - \tau_B u_{B_1}] \\
 (25) \quad & + a_{22B}[(\sigma_B + \tau_B)v_A - \sigma_B v_B - \tau_B v_{B_1}] \\
 & + a_{23B}[(\sigma_B + \tau_B)w_A - \sigma_B w_B - \tau_B w_{B_1}] = 0 \\
 & a_{31B}(u_A - u_{B_1}) + a_{32B}(v_A - v_{B_1}) + a_{33B}(w_A - w_{B_1}) = 0.
 \end{aligned}$$

Kennt man die Werte von u, v, w in den Punkten B und B_1 und ist die Determinante

$$(26) \quad |a_{ik}|_B \neq 0, \quad \text{ferner} \quad \sigma_B \geq 0, \quad \tau_B \geq 0, \quad \sigma_B + \tau_B \neq 0,$$

so können wir aus unseren Gleichungen (25) die Werte von u_A, v_A, w_A berechnen. Geht man also von den bekannten Werten auf einem Stück PQ der Anfangslinie aus, so kann man infolgedessen die Lösung der Differenzgleichung in allen Gitterpunkten desjenigen Dreiecks PQS berechnen, das man erhält, wenn man von den Punkten P und Q die Parallelen zur positiven α - bzw. ω -Richtung zieht, vorausgesetzt, daß überall in PQS unsere obigen Bedingungen (26) für irgendeinen Gitterpunkt B aus PQS erfüllt bleiben.

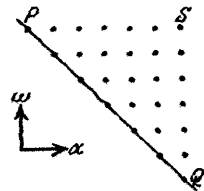


Fig. 2.

Wären die Koeffizienten a_{ik} , σ und τ konstant, so könnte man zeigen, daß im ganzen Dreieck PQS die Linearkombinationen

$$\begin{aligned}
 & a_{11}u + a_{12}v + a_{13}w, & B \xrightarrow{\alpha} A \\
 & a_{21}u + a_{22}v + a_{23}w, \\
 & a_{31}u + a_{32}v + a_{33}w
 \end{aligned}$$

unterhalb derjenigen Schranke bleiben, der sie für ihre Anfangswerte unterworfen sind. Die erste Linearkombination ist im Punkte A eines Dreiecks ABB_1 ebenso groß wie im Punkte B , und die dritte ist im Punkte A ebenso groß wie im Punkte B_1 , während die zweite Linearkombination



Fig. 3.

$$\begin{aligned}
 & a_{21}u_A + a_{22}v_A + a_{23}w_A \\
 & = \frac{\sigma}{\sigma + \tau} (a_{21}u_B + a_{22}v_B + a_{23}w_B) + \frac{\tau}{\sigma + \tau} (a_{21}u_{B_1} + a_{22}v_{B_1} + a_{23}w_{B_1})
 \end{aligned}$$

zwischen ihren Werten in den Punkten B und B_1 liegt, da σ und τ nicht negativ sind¹⁰⁾. Da die Determinante der drei Linearformen nicht verschwindet, ergibt sich, daß die drei Funktionen u, v, w selber unterhalb einer Schranke bleiben, die nur von ihren Werten auf der Anfangslinie abhängt und nicht von der Größe der zugrunde gelegten Mascheinteilung. Durch einen entsprechenden Schluß findet man, daß auch die Werte der Differenzenquotienten von u, v, w im ganzen Dreieck PQS beschränkt bleiben, wenn sie am Anfang beschränkt sind.

Im Falle nichtkonstanter Koeffizienten beginnen wir gleich damit, die Beschränktheit der ersten Differenzenquotienten aus ihrer Beschränktheit am Anfang abzuleiten.

Auf dem betrachteten Teil der Anfangsgeraden erfüllen die Anfangswerte von u, v, w ein Kurvenstück im dreidimensionalen (u, v, w) -Raum. Wir denken uns dieses Kurvenstück innerhalb eines solchen abgeschlossenen Gebietes G im (u, v, w) -Raum gelegen, daß die Größen a_{ik}, σ, τ mit ihren Ableitungen bis zur dritten Ordnung absolut für jeden Punkt unseres Gebietes beschränkt sind, daß die Determinante $|a_{ik}|$ in G absolut oberhalb einer Schranke $\delta > 0$ bleibt und daß in G $\sigma \geq 0, \tau \geq 0$ und $\sigma + \tau \neq 0$ ist.

Wir wollen jetzt zeigen, daß, solange unsere Lösungen u, v, w innerhalb des Gebietes G liegen, ihre ersten bis dritten Differenzenquotienten unabhängig von der Intervalleinteilung beschränkt bleiben, wenn sie es auf der Anfangsgeraden sind.

Vorerst bemerken wir folgendes: Setzen wir voraus, daß die Werte von u, v, w in den Punkten B, B_1, \dots innerhalb von G liegen. Ist dann $N > 0$ eine Schranke für den Betrag der Differenzenquotienten in der zur Anfangsgeraden parallelen Reihe der Punkte B, B_1, \dots , so liegen die ersten Differenzenquotienten zwischen Punkten der Reihe A, A_1, \dots und B, B_1, \dots , z. B. $\frac{u_A - u_B}{h}, \frac{u_A - u_{B_1}}{h}$ usw., unterhalb der

Schranke pN , wo p nur vom Gebiete G abhängt. Wir können nämlich die Gleichungen (25) beispielsweise so umschreiben, daß sie die Differenzen $u_A - u_B, v_A - v_B, w_A - w_B$ durch $u_B - u_{B_1}, v_B - v_{B_1}, w_B - w_{B_1}$ ausdrücken. Liegen ferner die „Linearkombinationen“

¹⁰⁾ Ein solcher Schluß ließe sich nicht unmittelbar auf die Funktionen selbst anwenden, da z. B. der Wert von u im Punkte A sich außer durch den Wert von u in B und B_1 noch durch erste Differenzen von v und w zwischen den Punkten B und B_1 ausdrückt.

$$\begin{aligned} a_{11B} \left(\frac{u_A - u_{A_1}}{h} \right) + a_{12B} \left(\frac{v_A - v_{A_1}}{h} \right) + a_{13B} \left(\frac{w_A - w_{A_1}}{h} \right), \\ a_{21B} \left(\frac{u_A - u_{A_1}}{h} \right) + a_{22B} \left(\frac{v_A - v_{A_1}}{h} \right) + a_{23B} \left(\frac{w_A - w_{A_1}}{h} \right), \\ a_{31B} \left(\frac{u_A - u_{A_1}}{h} \right) + a_{32B} \left(\frac{v_A - v_{A_1}}{h} \right) + a_{33B} \left(\frac{w_A - w_{A_1}}{h} \right) \end{aligned}$$

absolut unterhalb einer Schranke M , so liegen die Differenzenquotienten

$$\frac{u_A - u_{A_1}}{h}, \quad \frac{v_A - v_{A_1}}{h}, \quad \frac{w_A - w_{A_1}}{h}$$

absolut unterhalb der Schranke qM , wo q nur vom Gebiet G abhängt. Die Beschränktheit *dieser* Linearkombinationen können wir aber sukzessive nachweisen.

Wir können nämlich aus den Differenzgleichungen (25) Gleichungen herstellen, in denen unsere obigen Linearkombinationen für den Punkt A mit den entsprechenden Linearkombinationen für die Punkte B und B_1 verknüpft werden. Allerdings werden dabei noch Zusatzglieder auftreten, die jedoch von geringerer Größenordnung sind.

Zu dem Zweck ziehen wir von der Differenzgleichung (25)₂ im Punkte A

$$\begin{aligned} \sigma_B \left\{ a_{21B} \left(\frac{u_A - u_B}{h} \right) + a_{22B} \left(\frac{v_A - v_B}{h} \right) + a_{23B} \left(\frac{w_A - w_B}{h} \right) \right\} \\ + \tau_B \left\{ a_{21B} \left(\frac{u_A - u_{B_1}}{h} \right) + a_{22B} \left(\frac{v_A - v_{B_1}}{h} \right) + a_{23B} \left(\frac{w_A - w_{B_1}}{h} \right) \right\} = 0 \end{aligned}$$

dieselbe Gleichung für den Punkt A_1 ab:

$$\begin{aligned} \sigma_{B_1} \left\{ a_{21B_1} \left(\frac{u_{A_1} - u_{B_1}}{h} \right) + a_{22B_1} \left(\frac{v_{A_1} - v_{B_1}}{h} \right) + a_{23B_1} \left(\frac{w_{A_1} - w_{B_1}}{h} \right) \right\} \\ + \tau_{B_1} \left\{ a_{21B_1} \left(\frac{u_{A_1} - u_{B_2}}{h} \right) + a_{22B_1} \left(\frac{v_{A_1} - v_{B_2}}{h} \right) + a_{23B_1} \left(\frac{w_{A_1} - w_{B_2}}{h} \right) \right\} = 0. \end{aligned}$$

Die Differenz formen wir nun folgendermaßen um, wobei wir nur die Glieder hinschreiben, in denen u vorkommt:

$$\begin{aligned} (27a) \quad \sigma_B a_{21B} \frac{1}{h} (u_A - u_B - u_{A_1} + u_{B_1}) + (\sigma_B a_{21B} - \sigma_{B_1} a_{21B_1}) \left(\frac{u_{A_1} - u_{B_1}}{h} \right) + \dots = 0, \\ + \tau_B a_{21B} \frac{1}{h} (u_A - u_{B_1} - u_{A_1} + u_{B_2}) + (\tau_B a_{21B} - \tau_{B_1} a_{21B_1}) \left(\frac{u_{A_1} - u_{B_2}}{h} \right) \end{aligned}$$

oder schließlich (vgl. Fig. 4)

$$\begin{aligned} (27b) \quad (\sigma_B + \tau_B) a_{21B} \left(\frac{u_A - u_{A_1}}{h} \right) - \sigma_B a_{21C} \left(\frac{u_B - u_{B_1}}{h} \right) - \tau_B a_{21C_1} \left(\frac{u_{B_1} - u_{B_2}}{h} \right) \\ - h \sigma_B \left(\frac{a_{21B} - a_{21C}}{h} \right) \left(\frac{u_B - u_{B_1}}{h} \right) - h \tau_B \left(\frac{a_{21B} - a_{21C_1}}{h} \right) \left(\frac{u_{B_1} - u_{B_2}}{h} \right) + \dots = 0. \\ + h \left(\frac{\sigma_B a_{21B} - \sigma_{B_1} a_{21B_1}}{h} \right) \left(\frac{u_{A_1} - u_{B_1}}{h} \right) + h \left(\frac{\tau_B a_{21B} - \tau_{B_1} a_{21B_1}}{h} \right) \left(\frac{u_{A_1} - u_{B_2}}{h} \right) \end{aligned}$$

Analoge Umformungen erhalten wir für die Glieder mit v und w und ferner aus den Gleichungen $(25)_1$ und $(25)_3$.

Haben wir bis zur Parallelreihe der B zur Anfangsgeraden nachgewiesen, daß die Lösungen u, v, w innerhalb des Gebietes G liegen, und wissen wir ferner, daß die Linearkombinationen

$$(28) \quad \begin{aligned} a_{11}c \left(\frac{u_B - u_{B_1}}{h} \right) + \dots, & \quad a_{21}c_1 \left(\frac{u_{B_1} - u_{B_2}}{h} \right) + \dots, \\ a_{21}c \left(\frac{u_B - u_{B_1}}{h} \right) + \dots, & \quad a_{31}c_1 \left(\frac{u_{B_1} - u_{B_2}}{h} \right) + \dots \end{aligned}$$

unterhalb der Schranke M liegen, so entnehmen wir aus den drei Gleichungen (27b), daß die Linearkombinationen

$$\begin{aligned} a_{11}B \left(\frac{u_A - u_{A_1}}{h} \right) + \dots, \\ a_{21}B \left(\frac{u_A - u_{A_1}}{h} \right) + \dots, \\ a_{31}B \left(\frac{u_A - u_{A_1}}{h} \right) + \dots \end{aligned}$$

unterhalb der Schranke

$$M + hcM^2$$

liegen, wo $c > 0$ nur vom Gebiet G abhängt. Wir brauchen nämlich in (27b) die Differenzenquotienten der Koeffizienten nur nach dem Mittelwertsatz durch die Differenzenquotienten der Funktionen u, v, w abzuschätzen, unsere Vorbemerkungen (S. 216) zu beachten und entsprechend wie oben (S. 216) die Positivität von σ und τ zu benutzen.

Da diese drei Linearkombinationen auf der ersten Parallelreihe beschränkt bleiben, weil die ersten Differenzenquotienten der Anfangswerte beschränkt sind, so folgt, daß diese Kombinationen innerhalb desjenigen sich an die Gerade $\alpha + \omega = 0$ anschließenden Gebietes der (α, ω) -Ebene beschränkt bleiben, in dem die Lösungen u, v, w innerhalb von G liegen¹¹⁾. Damit ergibt sich nach den Vorbemerkungen von S. 216, daß die ersten

¹¹⁾ Bezeichnen wir nämlich die Schranken für unsere Kombinationen auf der ersten, zweiten, dritten, ..., n -ten Parallelreihe mit $M_1, M_2, M_3, \dots, M_n$, so finden wir für die Größen M_k die Differenzengleichung

$$\frac{M_{k+1} - M_k}{h} = cM_k^2,$$

deren Lösung für jede Maschenweite majorisiert wird durch die Lösung der Differentialgleichung

$$\frac{dM}{dt} = cM^2.$$

Differenzenquotienten selber in diesem Gebiete unabhängig von der Maschenweite beschränkt bleiben.

So ist also für das *Wachstum* der Funktionen u, v, w auf den Linien $\alpha = \text{konst.}$ und $\omega = \text{konst.}$ eine von der Intervalleinteilung unabhängige Schranke gegeben, woraus folgt, daß es ein an die Anfangsgerade anschließendes Gebiet B der (α, ω) -Ebene gibt, innerhalb dessen die Werte von u, v, w im Gebiet G bleiben, wie klein auch h gewählt ist.

Wir haben die Beschränktheit der ersten Differenzenquotienten im Gebiete B nachgewiesen, ohne die vorausgesetzte Beschränktheit des zweiten und dritten Differenzenquotienten am Anfang zu benutzen. Durch einen ähnlichen Schluß, indem man z. B. auf die Gleichung (27b) dasselbe Verfahren anwendet, durch das (27b) aus (25)₂ entstanden ist, kann man die Beschränktheit aller zweiten und aller dritten Differenzenquotienten im Gebiet B nachweisen, indem man diese Voraussetzungen mitbenutzt.

Aus der Beschränktheit der dritten Differenzenquotienten folgt, daß die zweiten Differenzenquotienten hinsichtlich h im ganzen Gebiet B „gleichartig stetig“ sind. Es existiert also nach dem bekannten Häufungssatz für Funktionenfolgen eine Teilfolge aus der Folge der Funktionen u, v, w bei immer kleiner werdender Maschenweite, so daß die zweiten Differenzenquotienten gleichmäßig gegen stetige Grenzfunktionen konvergieren. Die Summen, durch die sich die Funktionen u, v, w und ihre ersten Differenzenquotienten aus den zweiten Differenzenquotienten darstellen, konvergieren dabei gleichmäßig gegen Integrale, so daß wir wissen, daß die stetigen Grenzfunktionen der zweiten Differenzenquotienten die entsprechenden zweiten Ableitungen der Grenzfunktionen von u, v, w sind. Diese Grenzfunktionen genügen infolgedessen unseren Differentialgleichungen (24).

Da jede Lösung unserer Differenzgleichungen (25) in einem Punkte S nur von demjenigen Teil PQ der Anfangslinie, vgl. Fig. 2, S. 215, abhängt, der sich ergibt, wenn wir vom Punkte S aus die Geraden $\alpha = \text{konst.}$ und $\omega = \text{konst.}$ ziehen, bis sie die Anfangsgerade $\alpha + \omega = 0$ schneiden, so folgt, daß auch unsere Lösungen der Differentialgleichungen im Punkte S nur von demselben Stück der Anfangsgerade abhängen.

Daß unsere Lösungen aber die einzigen mit stetigen zweiten Ableitungen versehenen sind, erkennen wir folgendermaßen: Wir differenzieren unsere Differentialgleichungen (24) nach der zur Anfangsgeraden parallelen Richtung und ersetzen die Differentialquotienten durch Differenzenquotienten. Der Fehler ε_h , den wir dabei begehen, geht gleichmäßig im ganzen Gebiet B mit h gegen null. Wir subtrahieren die so entstehenden Differenzgleichungen für die Lösungen u, v, w unserer Differentialgleichungen von den entsprechenden Differenzgleichungen (27a) für die Lösungen $\bar{u}, \bar{v}, \bar{w}$ unserer Differenzgleichungen (25). Die so entstehenden Gleichungen formen

wir genau so um, wie wir (27a) in (27b) verwandelten. Wir erhalten dann eine Gleichung von ähnlichem Typus in $u - \bar{u}$, $v - \bar{v}$, $w - \bar{w}$ wie (27b) in u, v, w . Wir finden aus ihr: Ist M eine Schranke für die Linearkombinationen (28), genommen allerdings für die Differenzenquotienten der Funktionen $u - \bar{u}$, $v - \bar{v}$, $w - \bar{w}$ auf der n -ten Reihe, so liegen die Linearkombinationen auf der $n+1$ -ten Reihe unterhalb der Schranke $M + p h M + q h^2 + r h \varepsilon_n$, wobei die Konstanten p, q, r nur vom Gebiet G und den Schranken für die ersten und zweiten Differenzenquotienten der Funktionen u, v, w und $\bar{u}, \bar{v}, \bar{w}$ abhängen. Da nun unsere Linearkombinationen offenbar am Anfang verschwinden, so sind sie im ganzen Gebiet höchstens von der Größenordnung $h + \varepsilon_n$.

Daraus läßt sich folgern, daß die Folge der Lösungen $\bar{u}, \bar{v}, \bar{w}$ unserer Differenzgleichungen gegen jede mit stetigen zweiten Ableitungen versehene Lösungen u, v, w unserer Differentialgleichungen konvergiert; es kann also nur ein solches Lösungssystem bei vorgegebenen Anfangswerten geben.

Die Übertragung unseres Existenz- und Eindeutigkeitsbeweises auf den Fall von mehr als drei Funktionen ist unmittelbar klar. Wir können also folgenden allgemeinen Satz aussprechen:

Es seien auf einem Stück der Anfangsgeraden $\alpha + \omega = 0$ die Werte von n Funktionen u_1, \dots, u_n der beiden Veränderlichen α und ω dreimal beschränkt differenzierbar vorgegeben. Für diese Funktionen seien die folgenden n Differentialgleichungen gestellt:

$$\sum_{i=1}^n a_{1i} u_i^\alpha = 0,$$

$$\sum_{i=1}^n a_{2i} u_i^\beta = 0,$$

.....

$$\sum_{i=1}^n a_{ni} u_i^\omega = 0,$$

wo u_i^α und u_i^ω die Ableitungen von u_i nach α und ω bedeuten, wo

$$\begin{aligned} u_i^\beta &= \sigma_2 u_i^\alpha + \tau_2 u_i^\omega, & \left(\begin{array}{l} \sigma_2 \geq 0, \quad \tau_2 \geq 0, \quad \sigma_2 + \tau_2 \neq 0 \\ \sigma_3 \geq 0, \quad \tau_3 \geq 0, \quad \sigma_3 + \tau_3 \neq 0 \end{array} \right) \\ u_i^\gamma &= \sigma_3 u_i^\alpha + \tau_3 u_i^\omega, \\ & \dots \dots \dots \end{aligned}$$

ist, und wo σ_2, τ_2, \dots und die Koeffizienten a_{ik} dreimal beschränkt differenzierbare Funktionen von u_1, \dots, u_n sind; und zwar innerhalb eines gewissen Bereiches G , der die Anfangswerte im Inneren enthält. Ferner sei die Determinante $|a_{ik}|$ für die Anfangswerte von 0 verschieden.

Dann gibt es ein und nur ein zweimal stetig differenzierbares Lösungssystem u_1, \dots, u_n . Der Wert dieser Lösungen in einem Punkte S hängt nur von demjenigen Teil der Anfangswerte ab, der von den beiden Geraden $\alpha = \text{konst.}$ und $\omega = \text{konst.}$ durch diesen Punkt S aus der Anfangsgeraden ausgeschnitten wird. D. h. ändern wir die Vorgaben außerhalb dieses Stückes der Anfangsgeraden, so ändert sich der Wert im Punkte S und im ganzen Dreieck SPQ nicht mehr (vgl. Fig. 2, S. 215). Damit ist der im 1. Teil angekündigte Existenz- und Eindeutigkeitsbeweis nachgeholt.

(Eingegangen am 25. 2. 1927.)

Die Integralkurven einer gewöhnlichen Differentialgleichung erster Ordnung in der Umgebung rationaler Unbestimmtheitsstellen.

Von

Max Frommer in Göppingen.

Inhalt.

	Seite
Einleitung: Aufgabe und Methode	222
§ 1. Ausgezeichnete Richtungen; Einteilung	226
§ 2. Sondertypus	228
§ 3. Der definite Typus	234
§ 4. Die Behandlung des nicht-definiten Typus	236
§ 5. Die Umgebung einer einfachen, regulären aus- gezeichneten Richtung	238
§ 6. Die Umgebung einer mehrfachen, regulären aus- gezeichneten Richtung	241
§ 7. Die Umgebung einer singulären, ausgezeichneten Richtung	242
§ 8. Das Entscheidungsproblem	243
§ 9. Quantitative Auswertung der Methode	254
§ 10. Beispiele	265

Einleitung: Aufgabe und Methode.

Briot und Bouquet und vor allen Dingen Poincaré¹⁾ haben die Differentialgleichung

$$(1) \quad y' = \frac{ax + by + \varphi(x, y)}{cx + dy + \psi(x, y)}; \quad \begin{vmatrix} a & b \\ c & d \end{vmatrix} \neq 0; \quad \varphi(0, 0) = \psi(0, 0) = 0$$

in der Nähe des Urprungs untersucht. Dabei wurde vorausgesetzt, daß φ und ψ in Potenzreihen entwickelbar sind, deren niedrigste Glieder

¹⁾ Vgl. die in der Enzyklopädie II, A, 4a zitierten Arbeiten.

mindestens von zweiter Ordnung sein sollen. Für die Lösungen erhält man konvergente Reihen. Insbesondere hat Poincaré nachgewiesen, daß die Lösungen der Differentialgleichung

$$y' = \frac{\lambda y + \varphi(x, y)}{x + \psi(x, y)} \quad (\lambda \text{ nicht ganzzahlig und positiv})$$

Reihen sind, die nach Potenzen von x und x^{λ} fortschreiten. Aus diesen Entwicklungen wurde nachgewiesen, daß in der Nähe des Ursprungs die Gestalt der reellen Integralkurven einer solchen Differentialgleichung mit reellen Koeffizienten ähnlich ist der Gestalt der Integralkurven von

$$y' = \frac{ax + by}{cx + dy}$$

(mit Ausnahme der Fälle, die durch Koordinatentransformation auf den Fall $b = c = 0$, $a = -d$ zurückgeführt werden können).

Bendixson²⁾ bestimmte die Gestalt der Integralkurven von Differentialgleichungen von der Form:

$$(1a) \quad x^m y' = ax + by + \varphi(x, y)$$

und gab eine Methode an, die es ermöglicht, auch diese Integralkurven durch Reihen darzustellen. Er zeigte außerdem, daß man die Gleichung

$$(2) \quad y' = \frac{P(x, y)}{Q(x, y)}; \quad P(0, 0) = Q(0, 0) = 0$$

auf solche Differentialgleichungen zurückführen kann, wenn man voraussetzt, daß P und Q in konvergente Potenzreihen entwickelbar sind.

Mit Hilfe dieser Methode können demnach die Integralkurven der Differentialgleichung (2) in der Nähe des Ursprungs sowohl qualitativ als auch quantitativ bestimmt werden dadurch, daß sie auf Gleichungen von der Form (1) und (1a) zurückgeführt wird. Da aber in (1) und (1a) $\varphi(x, y)$ und $\psi(x, y)$ als Potenzreihen vorausgesetzt sind, können nur die sogenannten Briot-Bouquetschen Transformationen $x = (a + b\eta)\xi$ und $y = (c + d\eta)\xi$ zur Reduktion verwendet werden. Dadurch wird diese Reduktion etwas unübersichtlich, zumal eine geometrische Veranschaulichung der vorgenommenen Transformationen fehlt und deshalb der Zusammenhang der Differentialgleichung (2) mit den daraus abgeleiteten Gleichungen von der Form (1) und (1a) kaum mehr erkennbar ist.

In der vorliegenden Arbeit war nun das Bestreben, zur Bestimmung der Gestalt der Integralkurven von (2) ein direktes geometrisches Verfahren anzugeben. Zum Nachweis, daß dieses Verfahren immer zum Ziele führt, braucht man aber allgemeinere Transformationen von der Form $y = u(x)x^v$ und $y = x^{v(x)}$. Es ist deshalb nötig, in den Differential-

²⁾ Stockholmer Öfversigt 1898.

gleichungen (1) und (1a) die Voraussetzung, daß $\varphi(x, y)$ und $\psi(x, y)$ Potenzreihen sein sollen, fallen zu lassen.

Von dieser Voraussetzung machte sich auch Perron³⁾ in einer Arbeit frei, in der er nachwies, daß die Gestalt der Integralkurven von (1) durch die linearen Glieder bestimmt ist, wenn nur

$$\lim_{\substack{x \rightarrow 0 \\ y \rightarrow 0}} \frac{\varphi(x, y)}{|x| + |y|} = \lim_{\substack{x \rightarrow 0 \\ y \rightarrow 0}} \frac{\psi(x, y)}{|x| + |y|} = 0$$

ist; in Spezialfällen aber nur dann, wenn ein $r_0 > 1$ existiert, so daß

$$\lim_{\substack{x \rightarrow 0 \\ y \rightarrow 0}} \frac{\varphi(x, y)}{|x|^{r_0} + |y|^{r_0}} = \lim_{\substack{x \rightarrow 0 \\ y \rightarrow 0}} \frac{\psi(x, y)}{|x|^{r_0} + |y|^{r_0}} = 0$$

ist. Die dabei verwandte Methode ist in der Hauptsache funktionentheoretischer Art.

Außerdem behandelte Kuhn⁴⁾ die Differentialgleichung (1) unter der Voraussetzung, daß die beiden Grenzwerte $\lim_{\substack{x \rightarrow 0 \\ y \rightarrow 0}} \frac{\varphi(x, y)}{x^2 + y^2}$ und $\lim_{\substack{x \rightarrow 0 \\ y \rightarrow 0}} \frac{\psi(x, y)}{x^2 + y^2}$

endlich sind. In dieser Arbeit wird ein direktes geometrisches Verfahren zur Gewinnung der Gestalt der Integralkurven angewandt. Und zwar erhält man die allgemeinen Integralkurven durch Abschätzung (majorisierende und minorisierende Differentialgleichungen), die ausgezeichneten Integralkurven aber durch das Cauchysche Polygonverfahren.

In der vorliegenden Arbeit soll nun ein direktes Verfahren für die

Differentialgleichung (2) ausgebaut werden. Es kommt aber nicht die von Kuhn gebrauchte Abschätzung zur Anwendung, sondern eine von Dehn vorgeschlagene „Methode der Randsingularitäten“. Dabei betrachtet man im Richtungsfeld der Differentialgleichung eine aus differenzierbaren Kurvenstücken zusammengesetzte, geschlossene Kurve, die einen Bereich \mathfrak{B} einschließt (Fig. 1).

Liegt die Integralkurve durch einen

Punkt Q des Randes in der Nähe dieses Punktes ganz außerhalb des Be-

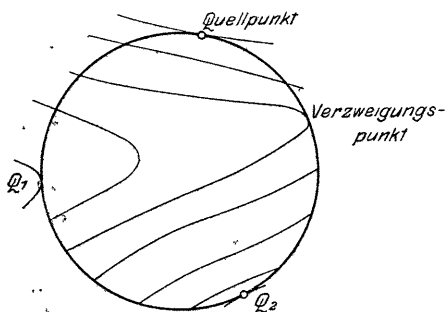


Fig. 1.

³⁾ O. Perron, Über die Gestalt der Integralkurven ... Math. Zeitschrift 15 (1922), 16 (1923).

⁴⁾ P. Kuhn, Über die Gestalt der Integralkurven einer gewöhnlichen Differentialgleichung erster Ordnung in der Umgebung gewisser singulärer Punkte. Inauguraldissertation, Frankfurt a. M. 1923.

reiches \mathfrak{B} , so nennen wir Q einen Quellpunkt für diesen Bereich. Liegt die Integralkurve durch einen Punkt V in der Nähe von V ganz innerhalb des Bereiches \mathfrak{B} , so nennen wir diesen Punkt einen Verzweigungspunkt. Quellpunkte und Verzweigungspunkte nennen wir die Randsingularitäten für den Bereich. Ist die Anzahl dieser Randsingularitäten endlich, so kann man aus der gegenseitigen Anordnung der Quell- und Verzweigungspunkte Schlüsse auf den Verlauf der Integralkurven im Innern des Bereiches ziehen. Mit wenigen Ausnahmen (vgl. § 2) werden daher alle Integralkurven aus den Eigenschaften des Feldes der Differentialgleichung und den allgemeinen Eigenschaften der Integralkurven, mit anderen Worten also mit Hilfe der Polygonmethode erhalten. Deshalb ist es nicht nötig, $P(x, y)$ und $Q(x, y)$ als Potenzreihen vorauszusetzen. Es genügt zu wissen, daß es um den Ursprung herum einen Bereich \mathfrak{B} gibt, in welchem P und Q „algebroid“ sind. Darunter soll verstanden werden, daß in \mathfrak{B} P und Q eindeutig und stetig und vom Ursprung aus durch Polynome approximierbar sind, und daß außerdem jede Gerade, die nicht durch den Ursprung geht, in \mathfrak{B} mit $P=0$ bzw. $Q=0$ nur eine endliche Anzahl von Schnittpunkten hat. Die Approximation durch die Polynome muß so sein, daß die Zusatzfunktionen gegenüber den Gliedern höchster Ordnung der Polynome noch klein sind; d. h. wenn

$$P(x, y) = \sum_{i+k \leq m} a_{ik} x^i y^k + \varphi(x, y),$$

$$Q(x, y) = \sum_{i+k \leq n} b_{ik} x^i y^k + \psi(x, y)$$

ist, so muß sein:

$$\lim_{\substack{x \rightarrow 0 \\ y \rightarrow 0}} \frac{\varphi(x, y)}{|x|^m + |y|^m} = \lim_{\substack{x \rightarrow 0 \\ y \rightarrow 0}} \frac{\psi(x, y)}{|x|^n + |y|^n} = 0.$$

Im allgemeinen genügt es, P und Q durch homogene Polynome P_m und Q_n zu approximieren; es wird sich zeigen, daß die Gestalt der Integralkurven durch die homogenen Polynome bestimmt ist, sobald nur φ und ψ den obigen Bedingungen genügen. In Spezialfällen muß zum Teil über die Zusatzfunktionen φ und ψ , zum Teil über die Approximationspolynome mehr vorausgesetzt werden, um auch hier Bestimmtes über die Gestalt der Integralkurven aussagen zu können.

Außerdem wird vorausgesetzt, daß im Bereich \mathfrak{B} der Ursprung der einzige Schnittpunkt der Kurven $P(x, y)=0$ und $Q(x, y)=0$ ist. Schließt man demnach den Ursprung durch einen beliebig kleinen Kreis K aus \mathfrak{B} aus, so entsteht ein Bereich \mathfrak{B}' mit nur regulären Punkten der Differentialgleichung. Von einem Punkt von \mathfrak{B}' aus kann also eine Integralkurve mit Hilfe des Cauchyschen Polygonverfahrens so lange fortgesetzt werden, als sie im Bereich \mathfrak{B}' bleibt. Ein innerer Punkt von \mathfrak{B}'

kann also nicht letzter Punkt einer Integralkurve sein. Die Integralkurve mündet dann in die singuläre Stelle ein, wenn mit Hilfe des Polygonverfahrens festgestellt werden kann, daß sie mit jedem noch so kleinen Kreis um den Ursprung Punkte gemeinsam hat.

Während nun Bendixson die Differentialgleichung (2) auf „Normaldifferentialgleichungen“ zurückführt, wird hier die Umgebung der singulären Stelle vom Ursprung aus in „Normalbereiche“, deren Überdeckung mit Integralkurven bekannt ist, eingeteilt. Natürlich werden sich in den komplizierten Fällen die Grenzen dieser Bereiche den Integralkurven anschmiegen, so daß diese Untersuchungen gleichzeitig das Verhalten der höheren Differentialquotienten der Integralkurven in der Nähe der singulären Stelle klarlegen. Deshalb ist auch der letzte Abschnitt dieser Arbeit der quantitativen Bestimmung der Integralkurven gewidmet. Es wird darin gezeigt, wie die Ergebnisse der Untersuchungen von Poincaré und Bendixson ganz natürlich aus dieser qualitativen Methode hervorgehen. Dabei tritt auch die Verwandtschaft dieser geometrischen Methode mit der analytischen Methode von Bendixson klar zutage.

Das Hauptziel der vorliegenden Untersuchungen ist aber nicht die quantitative Bestimmung der Integralkurven, die schon in ganz einfachen Beispielen auf komplizierte und unübersichtliche Rechnungen führen kann, sondern die Bestimmung der gestaltlichen Verhältnisse in der Umgebung der singulären Stelle. Nun sind die, unter noch allgemeineren Voraussetzungen topologisch möglichen Überdeckungen bzw. ihre topologisch invarianten Bestandteile von Brouwer⁵⁾ bestimmt worden. Es handelt sich also hier darum, festzustellen, ob Brouwers „erster“ oder „zweiter Hauptfall“ vorliegt und welches im ersten Fall die Anordnung der „elliptischen“, „parabolischen“ und „hyperbolischen“ Sektoren ist. Diese Sektoren sind nicht identisch mit den oben erwähnten „Normalbereichen“, lassen sich aber aus denselben zusammensetzen.

(Über die früheren Arbeiten berichtet Painlevé in der Enzyklopädie II, A, 4a, und auch Bieberbach in der „Theorie der Differentialgleichungen“.)

§ 1.

Ausgezeichnete Richtungen; Einteilung.

Es wird sich im folgenden ergeben, daß die Integralkurven, die in die singuläre Stelle einmünden, entweder Spiralen sind oder im Ursprung eine bestimmte Tangente besitzen. Wir fragen also: Gibt es stetig diffe-

⁵⁾ L. E. J. Brouwer, „On continuous vector distributions on surfaces“. Proceedings of the Koninklijke Akademie van Wetenschappen te Amsterdam, 1909/10.

rentierbare Funktionen $x = x(t)$, $y = y(t)$, die in der Umgebung des Ursprungs Lösungen von (2) sind und die außerdem die Eigenschaft haben, daß für einen Parameterwert z. B. $t = 0$, $x(0) = 0$ und $y(0) = 0$ ist. Bezeichnet man

$$x'(0) = \left. \frac{dx}{dt} \right|_{t=0} \text{ mit } p, \quad y'(0) = \left. \frac{dy}{dt} \right|_{t=0} \text{ mit } q,$$

so sollen p und q nicht gleichzeitig Null sein. Da die Funktionen der Differentialgleichung

$$\frac{dy}{dx} = \frac{P(x, y)}{Q(x, y)} = \frac{P_m(x, y) + \varphi(x, y)}{Q_n(x, y) + \psi(x, y)};$$

$$P_m = \sum_{i+k=m} a_{ik} x^i y^k, \quad Q_n = \sum_{i+k=n} b_{ik} x^i y^k$$

genügen, ist

$$\frac{y'(t)}{x'(t)} = \frac{P(x(t), y(t))}{Q(x(t), y(t))} = \frac{P_m(x(t), y(t)) + \varphi(x(t), y(t))}{Q_n(x(t), y(t)) + \psi(x(t), y(t))}$$

oder

$$y'(t)(Q_n + \psi) = x'(t)(P_m + \varphi),$$

$$y' \left(\frac{1}{t^n} Q_n + \frac{|x|^n + |y|^n}{t^n} \cdot \frac{\psi}{|x|^n + |y|^n} \right) = x' \left(\frac{1}{t^m} P_m + \frac{|x|^m + |y|^m}{t^m} \cdot \frac{\varphi}{|x|^m + |y|^m} \right) \cdot t^{m-n}.$$

Für $t \rightarrow 0$ ist

$$\lim \frac{1}{t^n} Q_n = \lim \sum_{i+k=n} b_{ik} \frac{x^i y^k}{t^i t^k} = \sum_{i+k=n} b_{ik} p^i q^k = Q_n(p, q),$$

$$\lim \frac{1}{t^m} P_m = \lim \sum_{i+k=m} a_{ik} \frac{x^i y^k}{t^i t^k} = \sum_{i+k=m} a_{ik} p^i q^k = P_m(p, q);$$

ferner nach Voraussetzung:

$$\lim \frac{\psi(x, y)}{|x|^n + |y|^n} = \lim \frac{\varphi(x, y)}{|x|^m + |y|^m} = 0.$$

Läßt man also in der obigen Gleichung t gegen Null gehen, so erhält man eine Gleichung zur Bestimmung von $\frac{q}{p}$, d. h. derjenigen Richtungen, längs derer Integralkurven in die singuläre Stelle einmünden können. Diese Richtungen werden ausgezeichnete Richtungen genannt; ihre Bestimmungsgleichung, $G(p, q) = 0$, charakteristische Gleichung.

Wir haben drei Fälle zu unterscheiden:

$$(3) \quad m > n, \quad G(p, q) \equiv q Q_n(p, q) = 0,$$

$$(4) \quad m < n, \quad G(p, q) \equiv p P_m(p, q) = 0,$$

$$(5) \quad m = n, \quad G(p, q) \equiv q Q_n(p, q) - p P_n(p, q) = 0.$$

Setzt man $\frac{y}{x} = u$, und bildet die Funktion

$$\psi(u, x) = \frac{P(x, ux)}{Q(x, ux)} - u \quad (\text{Feldrichtung} - \text{Richtung der Ursprungsgeraden}),$$

so sieht man ohne weiteres, daß der Zähler von $\psi(u, 0)$ identisch ist mit der Funktion $G(1, u)$. Die ausgezeichneten Richtungen sind also diejenigen Richtungen, auf denen in der Nähe des Ursprungs die Feldrichtung mit der Ursprungsgeraden zusammenfällt.

Ist $m \neq n$, dann ist $G(p, q)$ ein homogenes Polynom $(n+1)$ -ten bzw. $(m+1)$ -ten Grades. Es gibt also im allgemeinen nur endlich viele ausgezeichnete Richtungen, und zwar höchstens $(n+1)$ bzw. $(m+1)$. Nur im Falle $m = n$ kann $G(p, q)$ identisch verschwinden, wenn nämlich

$$q Q_n(p, q) \equiv p P_n(p, q)$$

ist. Dies ist der Fall, wenn

$$a_{n0} = b_{0n} = 0; \quad a_{n-1,1} = b_{n0}, \dots, a_{n-i,i} = b_{n-i+1,i-1}$$

ist.

Je nach der Beschaffenheit der charakteristischen Gleichung werden im folgenden drei Typen unterschieden:

1. Sondertypus: $G(p, q) \equiv 0$.

2. Definiter Typus: $G(p, q)$ definit; keine (reellen) ausgezeichneten Richtungen.

3. Nichtdefiniter Typus: $G(p, q)$ indefinit; es gibt eine bestimmte, endliche Anzahl von (reellen) ausgezeichneten Richtungen.

Wie man aus der Form von $G(p, q)$ leicht erkennt, kann auch der zweite Typus nur im Falle $m = n$ auftreten.

§ 2.

Sondertypus.

Zur Behandlung des Sondertypus wird zunächst keine geometrische Methode angewandt; in diesem Fall gehe man in die Differentialgleichung (2) ein mit der Briot-Bouquetschen Transformation $y = ux$. Dadurch erhält man:

$$u'x + u = \frac{P(x, ux)}{Q(x, ux)} = \frac{P_n(x, ux) + \varphi(x, ux)}{Q_n(x, ux) + \psi(x, ux)},$$

$$u' = \frac{P_n - uQ_n + \varphi - u\psi}{x(Q_n + \psi)}.$$

Nun ist in unserm Fall:

$$P_n - uQ_n \equiv 0,$$

also

$$u' = \frac{\varphi - u\psi}{x(Q_n + \psi)}.$$

Ferner sind nach Voraussetzung die Funktionen:

$$\Phi(x, u) = \frac{\varphi(x, ux)}{x^n} \quad \text{und} \quad \Psi(x, u) = \frac{\varphi(x, ux)}{x^n}$$

für $x = 0$ stetig. Also ist:

$$(6) \quad u' = \frac{\Phi(x, u) - u \Psi(x, u)}{x \left(\frac{Q_n(x, ux)}{x^n} + \Psi(x, u) \right)} = \frac{\Phi(x, u) - u \Psi(x, u)}{x(\Pi(u) + \Psi(x, u))}.$$

Da $b_{0n} = 0$ ist, ist

$$\frac{Q_n(x, ux)}{x^n} = b_{n0} + b_{n-1,1} u + \dots + b_{1, n-1} u^{n-1} = \Pi(u)$$

ein Polynom von höchstens $(n-1)$ -tem Grad in u .

Setzt man über Φ und Ψ nicht mehr voraus, als daß sie für $x = 0$ verschwinden, so kann man keine allgemeinen Schlüsse über die Gestalt der Integralkurven ziehen. Ich will zunächst voraussetzen, daß Φ und Ψ den Faktor x enthalten, Es ist also:

$$\Phi(x, u) = x F(x, u); \quad \Psi(x, u) = x G(x, u),$$

wobei $F(0, u)$ und $G(0, u)$ stetige und beschränkte Funktionen von u seien. Die Differentialgleichung (6) nimmt dann die Form:

$$(7) \quad u' = \frac{F(x, u) - u G(x, u)}{\Pi(u) + x G(x, u)}$$

an. Für $x = 0$ verschwindet der Nenner dieses Bruches nur an den Stellen, an denen $\Pi(u) = 0$ ist, also an höchstens $(n-1)$ Punkten. Jeder andere Punkt der u -Achse ist ein regulärer Punkt der Differentialgleichung, bestimmt also eindeutig eine Integralkurve. In der (xy) -Ebene mündet also längs jeder regulären Richtung eine und nur eine Integralkurve in die singuläre Stelle ein. Dabei werden als singuläre Richtungen diejenigen bezeichnet, deren Richtungskoeffizienten u_i die Gleichung $\Pi(u) = 0$ befriedigen.

Um die Verhältnisse in der Nähe der y -Achse zu untersuchen, wende man auf die Differentialgleichung die Transformation $x = vy$ an und verfähre entsprechend.

Um zu sehen, wie sich die Integralkurven in der Nähe der singulären Richtungen verhalten, betrachte man die entsprechenden Singularitäten in der (x, u) -Ebene. An einer solchen Stelle verschwindet der Nenner von (7). Verschwindet der Zähler nicht zugleich, so geht durch diesen Punkt eine Integralkurve, die dort die u -Achse zur Tangente hat. Durchsetzt diese Integralkurve die u -Achse gleichzeitig, so mündet in der (x, y) -Ebene längs dieser Richtung ebenfalls nur eine Integralkurve ein. Bleibt aber diese Integralkurve zunächst in einer Halbebene, z. B. $x \geq 0$, so ist

in der (x, y) -Ebene diese singuläre Richtung für $x \geq 0$ Rückkehrtangente einer Integalkurve, während für $x \leq 0$ keine Integalkurve längs dieser Richtung einmündet.

Wenn aber in einem solchen Punkt Zähler und Nenner gleichzeitig verschwinden, so haben wir hier eine neue Unbestimmtheitsstelle. Diese gehört entweder zu den später zu behandelnden Typen oder wiederum zum Sondertypus. In diesem Fall kommen aber, da

$$\Pi(u) = (u - u_i)^k \Pi_1(u); \quad \Pi_1(u_i) \neq 0$$

ist, im Nenner Glieder von der Ordnung $k \leq n - 1$ vor. Wenn man demnach durch die Briot-Bouquetschen Transformationen immer wieder auf den Sondertypus geführt wird, so kommt man nach endlich vielen Schritten zu einer Differentialgleichung, die auf der Ordinatenachse nur reguläre Punkte enthält. Damit ist der Sondertypus erledigt, bzw. auf später zu behandelnde Probleme zurückgeführt.

Es wurde allerdings bis jetzt vorausgesetzt, daß die Funktionen Φ und Ψ den Faktor x enthalten. Ich werde im folgenden nun zeigen, daß man topologisch dasselbe Kurvenbild erhält, wenn man nur voraussetzt, daß eine Zahl $r > 0$ existiert, so daß

$$\Phi(x, u) = |x|^r F(x, u), \quad (6)$$

$$\Psi(x, u) = x^r G(x, u)$$

ist und $F(0, u)$ bzw. $G(0, u)$ stetige und beschränkte Funktionen von u sind. Die Differentialgleichung (6) erhält dann die Form:

$$(8) \quad u' = \frac{F(x, u) - uG(x, u)}{x^{1-r}(\Pi(u) + x^r G(x, u))} = \frac{Z(x, u)}{x^{1-r} N(x, u)}.$$

Der Fall $r \geq 1$ ist erledigt; ist $r < 1$, so ist die u -Achse Integalkurve, aber in ihrer Umgebung ist die Lipschitzsche Bedingung in bezug auf keine Variable befriedigt. In Anlehnung an das Vorhergehende werden die Punkte $(0, u_i)$, für die $\Pi(u_i) = 0$ ist, als singuläre Punkte der u -Achse, alle andern als reguläre Punkte bezeichnet. Ich werde nun, und zwar mit Hilfe der Methode der Randsingularitäten zeigen, daß durch jeden regulären Punkt $(0, u_0)$ außer dem singulären Integral noch ein reguläres Integral hindurchgeht, in der (x, y) -Ebene also längs jeder regulären Richtung eine und nur eine Integalkurve in die singuläre Stelle einmündet.

Da $\Pi(u_0) \neq 0$ ist, gibt es um den Punkt $(0, u_0)$ einen konvexen Bereich \mathfrak{B} , in dem

$$|u'| < \frac{M}{x^{1-r}}$$

. 6) Im folgenden soll unter x^r immer $|x|^r$ verstanden werden.

ist. Die beiden Parabeln (siehe Fig. 2):

$$u_1(x) = u_0 + \frac{M}{r} x^r \quad \text{und} \quad u_2(x) = u_0 - \frac{M}{r} x^r$$

begrenzen also mit einer passenden Geraden $x = \delta$ einen Bereich S :

$$u_1 \geq u \geq u_2, \quad 0 \leq x \leq \delta,$$

der ganz in \mathfrak{B} liegt und auf seinem Rande nur in den Punkten $A(\delta, u_1(\delta))$ und $B(\delta, u_2(\delta))$ Randsingularitäten und zwar Quellpunkte besitzt. Betrachtet man nun die Integralkurve durch einen Punkt $X_1(x_1, u_1(x_1))$ bzw. $X_2(x_2, u_2(x_2))$, so wissen wir, da diese Integralkurve in S eine eindeutige Funktion von x ist, daß sie im Bereich S nur zu solchen Punkten gelangen kann, deren Abszisse größer ist als die Abszisse x_1 bzw. x_2 des Eintrittspunktes. Es ist also nicht möglich, daß die Integralkurve durch X_1 den Bereich S in X_2 verläßt, weil sonst sowohl $x_1 > x_2$, als auch $x_2 < x_1$ sein müßte. Also verlassen die Integralkurven durch X_1 und X_2 den Bereich S in Punkten Y_1 und Y_2 von AB . Bewegt sich X_1 auf u_1 von A gegen P , so bewegt sich Y_1 auf AB von A gegen B , ohne den Punkt B zu erreichen. Also ist ein Punkt Y bestimmt durch:

$$Y = \lim_{X_1 \rightarrow P} Y_1.$$

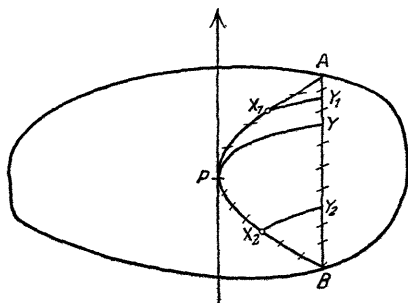


Fig. 2.

Die Integralkurve durch Y geht durch den Punkt P und berührt dort die u -Achse. Es ist zu zeigen, daß dies die einzige Integralkurve durch P ist. Aus der obigen Abschätzung von $|u'|$ geht hervor, daß jedes reguläre Integral durch P im Bereich S verlaufen muß. Nun bilde man durch die Transformation $u - u_0 = v x^r$ den Bereich S auf ein Rechteck der (x, v) -Ebene ab. Durch diese Transformation geht die Differentialgleichung (8) über in:

$$v' x^r + v r x^{r-1} = \frac{Z(x, u)}{N(x, u)} x^{r-1}$$

$$v' x = \frac{Z(x, u)}{N(x, u)} - r v = f(x, v).$$

Die Integralkurve PY habe die Gleichung $v = v(x)$, eine benachbarte die Gleichung $v = \bar{v}(x)$. Dann ist:

$$v'x = f(x, v)$$

$$\bar{v}'x = f(x, \bar{v}).$$

Also

$$(v - \bar{v})'x = f(x, v) - f(x, \bar{v}) = (v - \bar{v}) \frac{\partial f}{\partial v}(x, \bar{v})$$

\bar{v} zwischen v und \bar{v} .

Nun ist:

$$\frac{\partial f}{\partial v} = \frac{\partial}{\partial u} \left(\frac{Z(x, u)}{N(x, u)} \right) x^r - r,$$

also:

$$\frac{\partial f}{\partial v}(0, v) = -r.$$

Demnach gibt es einen Bereich $0 \leq x \leq \delta_1$, in welchem

$$\frac{\partial f}{\partial v}(x, v) < -\epsilon; \quad \epsilon > 0$$

ist. Also ist in diesem Bereich:

$$\frac{w'}{w} \equiv \frac{(v - \bar{v})'}{v - \bar{v}} < -\frac{\epsilon}{x}$$

$$w(x) = v(x) - \bar{v}(x).$$

Nun sind von der Kurvenschar

$$\frac{w'}{w} = -\frac{\epsilon}{x}$$

die Geraden $w = 0$ und $x = 0$ die einzigen Integralkurven durch den Punkt $(0, 0)$, während alle anderen asymptotisch an diese Geraden herangehen. Deshalb ist auch

$$\left| \lim_{x \rightarrow 0} w(x) \right| = \infty, \quad \text{wenn } w(\delta_1) \neq 0 \text{ ist.}$$

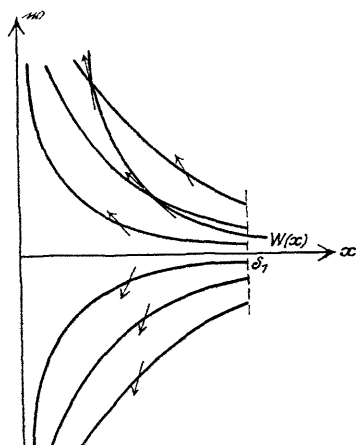


Fig. 3.

Das heißt aber für die (x, u) -Ebene, daß alle Kurven durch die Nachbarpunkte von Y , die nicht auf PY liegen, den Bereich S verlassen, also nicht durch P gehen. PY ist also das einzige reguläre Integral durch P . Also mündet in der (x, y) -Ebene längs jeder regulären Richtung durch den Ursprung eine und nur eine Integralkurve in die singuläre Stelle ein. Singuläre Richtungen kommen nur in endlicher Anzahl vor. Die Integralkurven in ihrer Umgebung erhält man durch das Studium der betreffenden Singularitäten der (u, x) -Ebene. Da $r < 1$ ist, kann in diesem Fall der Sondertypus nicht auftreten, weil die charakteristische Gleichung nicht identisch verschwindet. Längs dieser Richtungen münden entweder keine, endlich viele oder unendlich viele Integralkurven in die singuläre Stelle ein. Um dieses Kurvenbild zu erhalten, haben wir als hinreichende Bedingungen erkannt:

1. Die charakteristische Gleichung muß identisch befriedigt sein.

2. Die Zusatzfunktionen müssen klein werden wie eine Potenz von x ; deren Exponent höher ist als der Grad des homogenen Approximationspolynoms.

Perron²⁾ hat an einem Beispiel, das wir sogleich angeben werden, gezeigt, daß die erste Bedingung ohne die zweite nicht ausreichend ist, dieses Kurvenbild zu erzwingen.

Beispiele.

1. Beispiel ohne eine singuläre Richtung:

$$y' = \frac{y + \varphi(x, y)}{x + \psi(x, y)}.$$

Längs jeder Ursprungsgeraden mündet eine und nur eine Integralkurve in die singuläre Stelle ein, wenn

$$\lim_{\substack{x \rightarrow 0 \\ y \rightarrow 0}} \frac{\varphi(x, y)}{x^r + y^r} = \lim_{\substack{x \rightarrow 0 \\ y \rightarrow 0}} \frac{\psi(x, y)}{x^r + y^r} = 0$$

und $r > 1$ ist. Bei dem Beispiel von O. Perron

$$y' = \frac{y + \frac{2x}{\lg(x^2 + y^2)}}{x - \frac{2y}{\lg(x^2 + y^2)}}$$

ist $r = 1$; die Integralkurven sind Spiralen, die sich asymptotisch dem Ursprung nähern.

2. Beispiel mit einer singulären Richtung, längs der keine Integralkurve in den Ursprung einmündet:

$$y' = \frac{y^2 - x^4}{xy}.$$

Durch die Briot-Bouquetsche Transformation erhält man:

$$u' = \frac{-x}{u}$$

$$x^2 + u^2 = c^2:$$

sich auf einen Punkt zusammenziehende, lemniskatenartige Kurven 4. Grades, deren Achse die x -Achse ist. In der Richtung der Achse mündet keine Kurve in den Ursprung ein.

3. Beispiel mit einer singulären Richtung, längs der zwei Integralkurven in den Ursprung einmünden:

$$y' = \frac{y^2 + x^4}{xy}$$

$$u' = \frac{x}{u}$$

$$x^2 - u^2 = c:$$

längs der x -Achse münden zwei Parabeln in die singuläre Stelle ein.

4. Beispiel mit einer singulären Richtung, längs der unendlich viele Integralkurven in den Ursprung einmünden (Fig. 4):

$$y' = \frac{y^2 - 6x^2y + x^4}{xy - 3x^2}$$

$$u' = \frac{3u - x}{-u + 3x}.$$

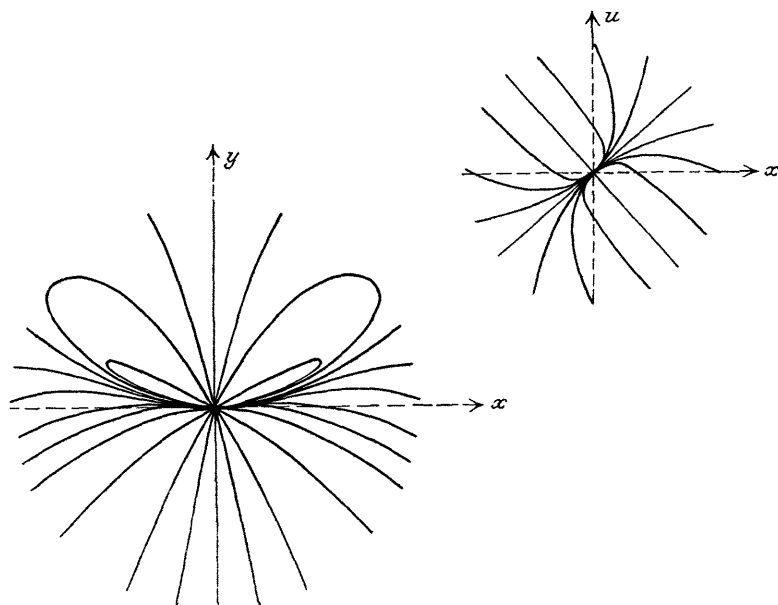


Fig. 4.

§ 3.

Der definite Typus.

Beim definiten Typus gibt es um den Ursprung einen Kreis, in dem die Feldrichtung nie mit der Ursprungsgeraden durch diesen Punkt zusammenfällt; im Innern dieses Kreises sind also die Integralkurven eindeutige Funktionen des Winkels φ der Ursprungsgeraden gegen die x -Achse⁷⁾. Führt man Polarkoordinaten ein, so geht der Kreis der (x, y) -Ebene über in ein Rechteck $ABCD$ der (r, φ) -Ebene. Diesen Bereich teile man durch die Geraden $r = \frac{1}{n}$ ($n = 1, 2, 3, \dots$) in verschiedene Bereiche. Trennt man die Strecke AB ($r = 0$) durch irgendeine dieser Geraden vom Rechteck $ABCD$ ab, so entsteht ein Bereich

⁷⁾ Der analytische Beweis dieses Satzes bietet keine Schwierigkeiten und ist deshalb hier weggelassen.

$$\frac{1}{n} \leq r \leq [r_0]; \quad 0 \leq \varphi \leq 2\pi,$$

in welchem nach dem Polygonverfahren die Integralkurve vollständig konstruiert werden kann.

Ich werde nun zeigen, daß die Integralkurve durch einen Punkt $X_0(r_0, 0)$ das Rechteck $ABCD$ in einem Punkt Y_0 von BC oder CD verlassen muß. Trennt man nämlich durch die Gerade $r = \frac{1}{n_0} AB$ vom Rechteck ab, so kann die Integralkurve durch X_0 verfolgt werden bis zu ihrem Austritt aus dem neuen Bereich. Geschieht dies in einem Punkt Y von BC oder CD , so ist der Beweis erbracht. Andernfalls geschieht es in einem Punkt $P_0\left(\frac{1}{n_0}, \varphi_0\right)$. Dann trenne man AB durch die nächste Gerade $r = \frac{1}{n_0+1}$ ab und mache dieselbe Fallunterscheidung. Entweder ist nun der ausgesprochene Satz richtig, oder man erhält auf jeder Geraden $\frac{1}{n_0+v}$ einen

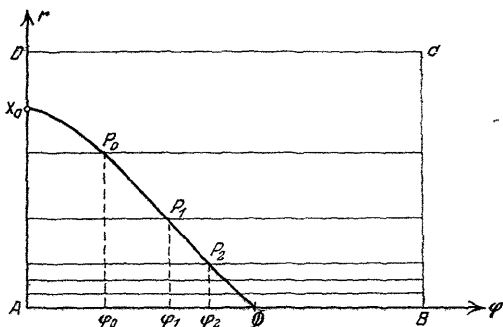


Fig. 5.

Punkt $P_v\left(\frac{1}{n_0+v}, \varphi_v\right)$. Die Argumente φ_v bilden dann eine monoton wachsende Folge, die nach oben beschränkt ist, also einen bestimmten Grenzwert Φ besitzt. Diese Integralkurve würde also durch den Punkt $(0, \Phi)$ stetig ergänzt; in der (x, y) -Ebene hätte also diese Kurve im Ursprung eine bestimmte Tangentenrichtung $\text{tg } \Phi$, was wider die Voraussetzung ist.

Die Integralkurve durch den Punkt $X_0(r_0, 0)$ verläßt also den Bereich $ABCD$ in einem Punkt Y_1 von BC oder CD . Liegt Y_1 auf BC , so ist er identisch mit einem Punkt X_1 auf AD . Die Integralkurve durch $X_1(r_1, 0)$ bestimmt vielleicht auf BC einen neuen Punkt

$$Y_2(r_2, 2\pi) \equiv X_2(r_2, 0)$$

usw. Ebenso ist

$$X_0(r_0, 0) \equiv Y_0(r_0, 2\pi).$$

Die Integralkurve durch Y_0 bestimmt eventuell auf AD den Punkt

$$X_{-1}(r_{-1}, 0) \equiv Y_{-1}(r_{-1}, 2\pi)$$

usw. Zu jedem Punkt $X_0(r_0, 0)$ gehört also eine unendliche Folge von Punkten $X_v(r_v, 0)$, wobei die r_v eine monotone Folge bilden. Sind zwei r_v einer Folge einander gleich, so sind notwendigerweise alle gleich; sind zwei voneinander verschieden, so sind notwendigerweise auch alle vonein-

ander verschieden. Es kann demnach in der (x, y) -Ebene geschlossene Integralkurven geben, die den Ursprung einschließen; alle andern winden sich unendlich oft um die singuläre Stelle herum. Es können also folgende Fälle eintreten:

1. Alle Integralkurven sind geschlossen (Wirbel).

Beispiel: $y' = -\frac{x}{y}$ (konzentrische Kreise).

2. In hinreichender Nähe des Ursprungs gibt es keine geschlossenen Integralkurven; diese sind Spiralen und nähern sich asymptotisch dem Ursprung (Strudel).

Beispiel: $y' = \frac{x+y}{x-y}$ (logarithmische Spiralen).

3. In beliebiger Nähe des Ursprungs gibt es noch geschlossene Kurven, denen sich andere asymptotisch nähern. (Strudel mit Grenzzyklen.)

Beispiel⁸⁾: $y' = \frac{x+y(x^2+y^2) \sin \frac{1}{x^2+y^2}}{-y+x(x^2+y^2) \sin \frac{1}{x^2+y^2}}$ oder in Polarkoordinaten:

$$\frac{dr}{d\varphi} = r^3 \sin \frac{1}{r^2}.$$

Als Grenzzyklen treten auf die Kreise $r = \sqrt{\frac{1}{k\pi}}$.

Dabei nähern sich die Spiralen von innen und von außen den Kreisen mit ungeradem k im Sinne des Uhrzeigers, den Kreisen mit geradem k im entgegengesetzten.

Diese drei Fälle sind die einzigen topologischen Möglichkeiten, die beim definiten Typus und, wie sich später ergeben wird, in jedem Fall, bei dem keine Integralkurve mit bestimmter Tangentenrichtung in die singuläre Stelle einmündet (Brouwers „zweiter Hauptfall“), vorkommen können. Es gibt jedoch keine allgemeine Methode, die es ermöglicht, bei einer vorgegebenen Differentialgleichung zu entscheiden, welcher der drei Fälle vorliegt. Im folgenden wird gezeigt werden, daß die Entscheidung zwischen den topologischen Möglichkeiten immer getroffen werden kann, sobald es Integralkurven gibt, die mit bestimmter Tangentenrichtung in die singuläre Stelle einmünden (Brouwers „erster Hauptfall“).

§ 4.

Die Behandlung des nicht-definiten Typus.

Beim nicht-definiten Typus kommt eine endliche Anzahl von ausgezeichneten Richtungen vor, längs welchen ihrer Bestimmung nach Integralkurven in die singuläre Stelle einmünden können. Anstatt nun die

⁸⁾ Siehe Bieberbach, Differentialgleichungen.

ganze Umgebung der singulären Stelle auf einmal zu betrachten, teile ich diesen Bereich durch passende Ursprungsgeraden in verschiedene sektorförmige Teilbereiche ein, die entweder eine oder keine ausgezeichnete Richtung enthalten.

Ist A_1OA ein Bereich ohne eine ausgezeichnete Richtung, so geht die Integralkurve durch einen Punkt P auf OA in genügender Nähe von O durch einen Punkt Q von OA_1 und umgekehrt. Der Beweis ist dem im § 3 vollständig analog (dem Sektor A_1OA entspricht dort die ganze Umgebung der singulären Stelle). Diese Bereiche sind demnach regulär mit Integralkurven überdeckt. Die Gestalt der Integralkurven in der Nähe der singulären Stelle ist also im wesentlichen bestimmt durch ihr Verhalten in den, die ausgezeichneten Richtungen umgebenden Bereichen.

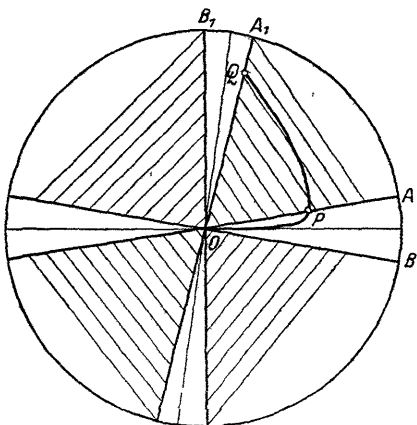


Fig. 6.

Weiterhin ist klar, daß eine Integralkurve durch den Punkt P , von der ich nachweisen kann, daß sie in dem, eine ausgezeichnete Richtung einschließenden Sektor AOB in die singuläre Stelle einmündet, dort diese ausgezeichnete Richtung berühren muß. Denn wählt man anstatt des Sektors AOB einen anderen Sektor $A'OB'$ mit kleinerem Öffnungswinkel, der die ausgezeichnete Richtung ebenfalls einschließt, so wissen wir, daß in hinreichender Nähe des Ursprungs diese Integralkurve in diesem Sektor verlaufen muß, denn nur in diesem Sektor kann eine Integralkurve, die im Sektor AOB bleibt, in den Ursprung einmünden. Die Integralkurve bleibt also in hinreichender Nähe des Ursprungs in einem beliebig kleinen, die ausgezeichnete Richtung umgebenden Sektor, d. h. sie berührt im Ursprung die ausgezeichnete Richtung.

Um die Gestalt der Integralkurven in den, die ausgezeichneten Richtungen umgebenden Bereichen zu finden, mache ich die zu untersuchende ausgezeichnete Richtung zur x -Achse. Ich werde dann zwei Arten von ausgezeichneten Richtungen unterscheiden:

1. Die reguläre ausgezeichnete Richtung (allgemeiner Fall).
2. Die singuläre ausgezeichnete Richtung. Die x -Achse wird dann singuläre ausgezeichnete Richtung genannt, wenn $y = 0$ zugleich die charakteristische Gleichung $G(x, y) = 0$ und $Q_m(x, y) = 0$ befriedigt, wenn also $Q(x, y) = 0$ im Ursprung die x -Achse berührt. Dementsprechend kommen

dann in beliebiger Nähe der ausgezeichneten Richtung und gleichzeitig in beliebiger Nähe des Ursprungs zur ausgezeichneten Richtung senkrechte Feldrichtungen vor. Ist die letzte Bedingung nicht befriedigt, so wird die x -Achse reguläre ausgezeichnete Richtung genannt. Es kann leicht nachgewiesen werden, daß durch Einführung einer neuen y -Achse es nicht möglich ist, eine singuläre Richtung in eine reguläre überzuführen und umgekehrt. Es handelt sich hier also um Merkmale der ausgezeichneten Richtungen, die gegen Koordinatentransformationen invariant sind.

Bei den regulären ausgezeichneten Richtungen will ich wiederum zwei Arten unterscheiden:

a) Die einfache, reguläre ausgezeichnete Richtung, die man als einfache Wurzel der charakteristischen Gleichung erhält.

b) Die mehrfache, reguläre ausgezeichnete Richtung, die man als mehrfache Wurzel der charakteristischen Gleichung erhält.

Diese einzelnen Arten werden im folgenden der Reihe nach behandelt werden.

§ 5.

Die Umgebung einer einfachen, regulären ausgezeichneten Richtung.

Es sei $y = 0$ eine reguläre ausgezeichnete Richtung. Da in diesem Fall $Q(x, y) = 0$ im Ursprung die x -Achse nicht berührt, gibt es einen Bereich OAB

$$-\varepsilon \leq \frac{y}{x} = u \leq \varepsilon; \quad 0 \leq x \leq \delta,$$

in welchem überall mit Ausnahme des Punktes $(0, 0)$ $Q(x, y) \neq 0$ ist. Bildet man außerdem die im § 1 eingeführte Funktion $\psi(u, x)$ ⁹⁾, die die Differenz von Feldrichtung und Richtung der Ursprungsgeraden darstellt, so ist $\psi(0, 0) = 0$, wenn die x -Achse ausgezeichnete Richtung ist. Ist sie aber eine einfache, reguläre ausgezeichnete Richtung, so ist außerdem $\frac{\partial \psi}{\partial u}(0, 0) \neq 0$. Die obigen Größen ε und δ können nun so klein gewählt werden, daß auch im ganzen Bereich OAB $\frac{\partial \psi}{\partial u}(u, x) \neq 0$ ist. Dann ist $\psi(\pm \varepsilon, 0) \neq 0$; also kann δ noch so bestimmt werden, daß $\psi(\pm \varepsilon, x) \neq 0$ ist für $0 \leq x \leq \delta$. Auf den Schenkeln OA und OB stimmt also die Feldrichtung nie mit der Geradenrichtung überein, ebensowenig auf der Strecke AB . Der Bereich OAB kann also höchstens in den Punkten A und B Randsingularitäten enthalten. Es sind nun zwei Fälle zu unterscheiden:

⁹⁾ In der (x, u) -Ebene lautet die Differentialgleichung: $u'x + u = y'(x, u)$

$$u' = \frac{y'(x, u) - u}{x} = \frac{\psi(x, u)}{x}.$$

$$1. \frac{\partial \psi}{\partial u}(0, 0) > 0.$$

Dann ist:

$$\psi(\varepsilon, x) > 0,$$

$$\psi(-\varepsilon, x) < 0.$$

Der Bereich OAB (Fig. 7) ist frei von Randsingularitäten. Trennt man nun die singuläre Stelle O durch eine Gerade A_1B_1 ($x = \delta_1$) vom Bereich OAB ab, so entsteht ein Bereich A_1B_1BA , der nur reguläre Punkte der Differentialgleichung enthält, und in dem die Integralkurven eindeutige Funktionen von x sind. Nach bekannter Schlußweise kann man also die Integralkurven von einem Punkt des Randes dieses Bereiches durch den Bereich hindurch konstruieren bis zu einem andern Randpunkt. Die Integralkurve durch den Punkt X_1 (mit der Abszisse x_1) des nicht-geschlossenen Streckenzugs A_1AB_1 kann im betrachteten Bereich nur zu solchen Punkten gelangen, deren Abszissen kleiner sind als x_1 ; daraus folgt, daß sie den Bereich nur in einem Punkt von A_1B_1 verlassen kann. Denn verließ sie den Bereich in einem Punkt X_2 mit der Abszisse x_2 , der auf dem obigen Streckenzug läge, so wäre sowohl $x_1 > x_2$ als auch $x_2 > x_1$. Also hat die Integralkurve durch X_1 mit der Strecke A_1B_1 noch einen Punkt gemeinsam, so klein auch δ_1 sein mag. Sie mündet also in den Ursprung ein und berührt dort die x -Achse (§ 4). Ist also $\frac{\partial \psi}{\partial u}(0, 0) > 0$, so mündet jede Integralkurve durch einen Punkt des Randes von OAB mit wagrechter Tangente in die singuläre Stelle ein.

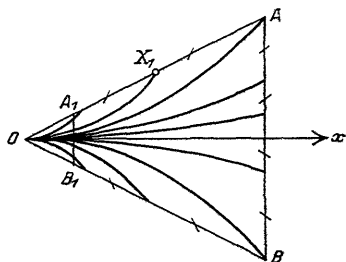


Fig. 7.

$$2. \frac{\partial \psi}{\partial u}(0, 0) < 0.$$

Dann ist:

$$\psi(\varepsilon, x) < 0,$$

$$\psi(-\varepsilon, x) > 0.$$

Der Bereich OAB besitzt in den Punkten A und B Quellpunkte. In diesem Fall gibt es eine und nur eine Integralkurve, die in diesem Bereich in die singuläre Stelle einmündet (Beweis siehe § 2).

Ist $m > n$ und ist die x -Achse reguläre ausgezeichnete Richtung, so ist:

$$\psi(u, 0) = -u.$$

Also:

$$\frac{\partial \psi}{\partial u}(0, 0) = -1.$$

In diesem Falle mündet also längs der x -Achse immer nur eine Integralkurve in die singuläre Stelle ein.

Betrachtet man nun den Winkelraum zwischen zwei aufeinanderfolgenden, einfachen, regulären ausgezeichneten Richtungen, so können je nach der Beschaffenheit der diese Richtungen umgebenden Bereiche drei Fälle eintreten (vgl. § 4):

1. Fall. Längs jeder ausgezeichneten Richtung münden unendlich viele Integralkurven in die singuläre Stelle ein.

Die Integralkurven sind in der Nähe des Ursprungs blattförmig geschlossene Kurven (Fig. 8).

2. Fall. Längs einer ausgezeichneten Richtung münden unendlich viele Integralkurven, längs der anderen nur eine in die singuläre Stelle ein.

Die Integralkurven besitzen eine „Grenzkurve“ (Fig. 9).

3. Fall. Längs jeder ausgezeichneten Richtung mündet nur eine Integralkurve in die singuläre Stelle ein.

Die Integralkurven besitzen zwei Grenzkurven (Fig. 10).

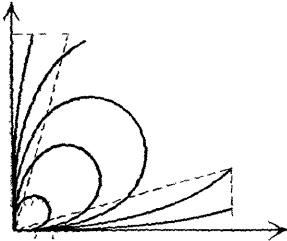


Fig. 8.

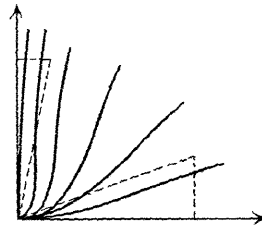


Fig. 9.

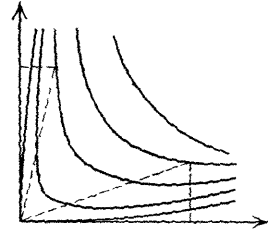


Fig. 10.

Diese drei Fälle bilden die für alle Überdeckungen charakteristischen Elemente.

Beispiele.

$$1. \quad y' = \frac{ay + \Theta(x, y)}{x + H(x, y)}; \quad a \neq 0 \text{ und } \neq 1.$$

$$\lim_{\substack{x \rightarrow 0 \\ y \rightarrow 0}} \frac{\Theta(x, y)}{|x| + |y|} = \lim_{\substack{x \rightarrow 0 \\ y \rightarrow 0}} \frac{H(x, y)}{|x| + |y|} = 0.$$

$$\psi(u, 0) = (a - 1)u.$$

$$\frac{\partial \psi}{\partial u}(0, 0) = a - 1.$$

Für die Umgebung der y -Achse tritt an Stelle von a sein reziproker Wert $\frac{1}{a}$.

Ist $a > 0$, so ist entweder $a - 1 > 0$ und $\frac{1}{a} - 1 < 0$, oder umgekehrt (Knoten, 2. Fall).

Ist $a < 0$, so ist sowohl $(a - 1) < 0$ als auch $\left(\frac{1}{a} - 1\right) < 0$ (Sattel, 3. Fall).

$$2. \quad y' = \frac{y}{x(x+y)(x-y)(x-2y)}.$$

$$\psi(u, x) = \frac{u}{x^3(1+u)(1-u)(1-2u)} - u.$$

$$\frac{\partial \psi}{\partial u}(0, 0) = +\infty \text{ für } x > 0, \\ -\infty \text{ für } x < 0.$$

Längs der y -Achse mündet nach S. 239 nur eine Integralkurve in den Ursprung ein, links von der y -Achse ist die Überdeckung sattelförmig, rechts knotenförmig (2. und 3. Fall).

$$3. \quad y' = \frac{y(2x-y)(2x+y)}{x(x-2y)(x+2y)};$$

Charakteristische Gleichung: $3xy(x^2 + y^2) = 0$.

$$\psi(u, x) = \frac{u(2-u)(2+u)}{(1-2u)(1+2u)} - u.$$

$$\frac{\partial \psi}{\partial u}(0, 0) = 3 > 0.$$

Vertauscht man x mit y und umgekehrt, so geht die Gleichung in sich selbst über. Also haben die Integralkurven in der Umgebung der y -Achse dieselbe Gestalt wie in der Umgebung der x -Achse.

Die Überdeckung zeigt in jedem Quadranten blattförmig geschlossene Kurven (1. Fall).

§ 6.

Die Umgebung einer mehrfachen, regulären ausgezeichneten Richtung.

Auch in diesem Falle kann ein Bereich OAB konstruiert werden, der höchstens in den Punkten A und B Randsingularitäten besitzt. Man braucht nur ε und δ so klein zu wählen, daß $\psi(u, 0)$ im Intervall $|u| \leq \varepsilon$ nur für $u = 0$ verschwindet, und daß $\psi(\pm \varepsilon, x) \neq 0$ ist für $0 \leq x \leq \delta$. Da aber $\frac{\partial \psi}{\partial u}(0, 0) = 0$ ist, ist es nicht nötig, daß, wenn A_1 bzw. B_1 Quellpunkt ist, auch B_1 bzw. A_1 dieselbe Eigenschaft hat; wir haben also drei Fälle zu unterscheiden:

1. Der Bereich OAB ist frei von Randsingularitäten. Nach S. 239 münden die Integralkurven durch sämtliche Punkte des Randes mit wag-rechter Tangente in den Ursprung ein.

2. Der Bereich OAB hat in A und B je einen Quellpunkt (Fig. 11). Nach S. 231 gibt es mindestens eine Integralkurve OY , die im Bereich OAB in die singuläre Stelle einmündet. Gibt es eine davon verschiedene Integralkurve $O\bar{Y}$, die in den Ursprung einmündet, so münden sämtliche Integralkurven durch die Punkte der Strecke $Y\bar{Y}$ mit wagrechter Tangente in den Ursprung ein. Entweder gibt es also nur eine Integralkurve OY , die in die singuläre Stelle einmündet, oder unendlich viele, die von zwei Grenzkurven OY und $O\bar{Y}$ eingeschlossen werden; eine andere Möglichkeit ist nicht vorhanden. Das Problem der Entscheidung wird später behandelt.

3. Der Bereich OAB hat nur in einem Eckpunkt, z. B. dem Punkt A , einen Quellpunkt (Fig. 12). Die Integralkurve durch einen Punkt $X_1(x_1, \varepsilon x_1)$ verläßt den Bereich in einem Punkt Y_1 , dessen Abszisse größer ist als x_1 ,

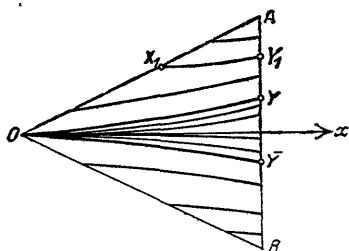


Fig. 11.

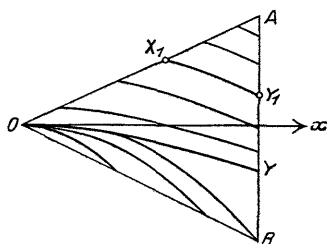


Fig. 12.

also notwendigerweise in einem Punkt von AB oder OB . Bewegt sich nun X_1 auf AO von A gegen O , so bewegt sich Y_1 auf ABO von A gegen O . Es können dann zwei Fälle eintreten:

a) $\lim_{x_1 \rightarrow 0} Y_1 = 0$; es mündet keine Integralkurve aus OAB in die singuläre Stelle ein.

b) $\lim_{x_1 \rightarrow 0} Y_1 = Y \neq 0$; sämtliche Integralkurven durch die Punkte des Randstückes OY , das den Punkt A nicht enthält, münden mit wagrechter Tangente in die singuläre Stelle ein.

Eine andere Möglichkeit gibt es nicht. Die Entscheidung wird in § 8 geliefert.

§ 7.

Die Umgebung einer singulären ausgezeichneten Richtung.

Im vorliegenden Fall berührt $Q(x, y) = 0$ die x -Achse; es ist also $Q_x(x, 0) \equiv 0$. Außerdem ist $G(x, 0) \equiv 0$, also auch $\psi(0, 0) = 0$. Man kann aber eine Zahl ε so bestimmen, daß für $-\varepsilon \leq u = \frac{y}{x} \leq \varepsilon$ die Ge-

rade $y = 0$ die einzige Nullstelle der Funktionen $G(x, y)$ und $Q_n(x, y)$ ist. Dann ist auch

$$\psi(\varepsilon, 0) \neq 0 \quad \text{und} \quad \psi(-\varepsilon, 0) \neq 0.$$

Also kann δ so gewählt werden, daß für $0 < x \leq \delta$ die Funktionen $Q(x, \varepsilon x)$ und $Q(x, -\varepsilon x)$, ebenso $\psi(\varepsilon, x)$ und $\psi(-\varepsilon, x)$ nicht verschwinden. Der so definierte Bereich OAB (Fig 13) besitzt demnach auf seinen Schenkeln OA und OB keine Randsingularitäten, und außerdem kommt auf ihnen keine zur y -Achse parallele Feldrichtung vor. Randsingularitäten entstehen höchstens in den Punkten A und B und in den Punkten Q_1, Q_2, \dots, Q_n der Strecke AB , in denen $Q(x, y)$ verschwindet. Man kann nun annehmen, daß diese Punkte Q_i auf solchen Zweigen von $Q(x, y) = 0$ liegen, die im Ursprung die x -Achse berühren, denn alle anderen Punkte Q_i können durch Verkleinerung von δ zum Wegfall gebracht werden. Nun schneide man den Bereich OAB längs dieser durch die Punkte Q_i gehenden Kurvenzweige von $Q(x, y) = 0$ auf. Es entsteht dadurch eine endliche Anzahl von Teilbereichen, die nur in den Eckpunkten auf AB Randsingularitäten (und zwar Quellpunkte) besitzen können. Es können demnach folgende drei Fälle eintreten:

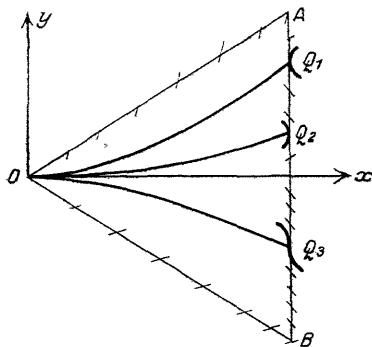


Fig. 13.

1. Ein Teilbereich hat keine Randsingularitäten.
2. Ein Teilbereich hat einen Quellpunkt.
3. Ein Teilbereich hat zwei Quellpunkte.

Da im Innern dieser Teilbereiche $Q(x, y)$ nie verschwindet, sind darin die Integralkurven eindeutige Funktionen von x . Die Überlegungen des vorigen Paragraphen lassen sich also vollständig übertragen. Demnach ist das Problem, die Gestalt der Integralkurven im Bereich OAB zu finden, übergeführt in das Problem, zu entscheiden, ob ein Bereich mit einem oder mit zwei Quellpunkten auf die eine oder die andere, in § 6 dargestellte Art mit Integralkurven überdeckt ist.

§ 8.

Das Entscheidungsproblem.

Um die Gestalt der Integralkurven in einem, die x -Achse als ausgezeichnete Richtung umgebenden Sektor weiter zu untersuchen, werde ich im folgenden diesen Sektor durch krummlinige Grenzen weiter einengen,

resp. zerlegen. Die Grenzen dieser Gebiete müssen sich den Integralkurven näher anschmiegen als die Tangente. So werden wir zunächst auf das Problem geführt, die Krümmungsverhältnisse der Integralkurven in einem Sektor zu untersuchen. Liegt eine Integralkurve für kleine x unterhalb jeder Parabel $y = x^\nu$, $\nu < \nu_1$ und oberhalb jeder Parabel $y = x^\nu$, $\nu > \nu_1$, so nennen wir ν_1 die „Krümmungsordnung“ dieser Integralkurve. Liegt eine Integralkurve für kleine x unterhalb *jeder* Parabel $y = x^\nu$, so nennen wir ihre Krümmungsordnung Unendlich. Liegt eine Integralkurve von der Krümmungsordnung ν_1 für kleine x unterhalb jeder Parabel $y = ux^{\nu_1}$, $u > u_1$ und oberhalb jeder Parabel $y = ux^{\nu_1}$, $u < u_1$, so nennen wir u_1 ihr „Krümmungsmaß“. Liegt sie unterhalb *jeder* Parabel $y = ux^{\nu_1}$, so nennen wir ihr Krümmungsmaß Null; wir nennen es Unendlich, wenn sie oberhalb aller dieser Parabeln liegt. Haben Integralkurven die Krümmungsordnung 1, und berühren sie die x -Achse, so ist ihr Krümmungsmaß notwendigerweise gleich Null (vgl. § 10, 1. Beispiel).

Um nun festzustellen, welche Krümmungsordnungen überhaupt vorkommen können, gehe ich in die Differentialgleichung (2) ein mit dem Ansatz $y = x^{\nu(x)}$. Dadurch erhält man:

$$(9) \quad \frac{d\nu}{dx} = \nu' = \frac{P(x, x^\nu) - \nu x^{\nu-1} Q(x, x^\nu)}{x^\nu Q(x, x^\nu) \lg x}.$$

Verzichtet man zunächst einmal auf diejenigen Integralkurven der (x, y) -Ebene, deren Krümmungsordnung sehr groß, also größer als eine feste Zahl N ist, so hat man die Differentialgleichung (9) zu untersuchen in einem Bereich

$$1 \leq \nu \leq N; \quad 0 \leq x \leq \delta.$$

Insbesondere soll festgestellt werden, ob es auf der ν -Achse, die sich als Integralkurve ergeben wird, solche Punkte $(0, \nu_i)$ gibt, durch die noch andere Integralkurven hindurchgehen können. Diese Werte ν_i sind dann die möglichen Krümmungsordnungen.

Zu diesen Untersuchungen genügen die ursprünglichen Voraussetzungen über $P(x, y)$ und $Q(x, y)$ nicht. Um nicht zu wenig vorauszusetzen, nehme ich zunächst an, daß P und Q in der Umgebung des Ursprungs in konvergente Potenzreihen entwickelbar sind. Es ist also:

$$P(x, y) = \sum a_{ik} x^i y^k; \quad P(x, x^\nu) = \sum a_{ik} x^{i+k\nu}, \\ Q(x, y) = \sum b_{ik} x^i y^k; \quad Q(x, x^\nu) = \sum b_{ik} x^{i+k\nu}.$$

Unter diesen Voraussetzungen kann man den Quotienten der Differentialgleichung (9) mit einer bestimmten Potenz von x kürzen. Zu diesem Zweck muß festgestellt werden, welcher Exponent der Summen $P(x, x^\nu)$ und $Q(x, x^\nu)$ jeweils der kleinste ist. Diese Exponenten sind von der

Form $i + k\nu$, haben also im Intervall $1 \leq \nu \leq S$ ihren kleinsten Wert bei $\nu = 1$, ihren größten bei $\nu = N$. Nun können für ein bestimmtes ν dieses Intervalls nur diejenigen Exponenten den kleinsten Wert haben, deren Minimum kleiner ist als das Maximum M irgend eines Exponenten; es kommen also nur die in Betracht, für die

$$i + k < M,$$

also nur eine endliche Anzahl (Fig. 14). Das ganze Intervall kann also in endlich viele Teilintervalle geteilt werden, die so beschaffen sind, daß man im Innern eines solchen Intervalls aus der Summe $a_{ik} x^{i+k\nu}$ eine bestimmte Potenz $x^{J_1+K_1\nu}$ herausziehen kann. Die Werte ν_i , die diese Einteilung bewirken, werden aus linearen Gleichungen erhalten; diese ν_i sind also rational. Ebenso erhält man für $\sum b_{ik} x^{i+k\nu}$ eine solche Einteilung. Überlagert man diese beiden Einteilungen, so erhält man eine endliche Anzahl rationaler ν_i , die das Intervall $1 \leq \nu \leq N$ in solche Teilintervalle teilen, in denen man aus $\sum a_{ik} x^{i+k\nu}$ eine Potenz $x^{e_1} \equiv x^{J_1+K_1\nu}$, aus $\sum b_{ik} x^{i+k\nu}$ eine Potenz $x^{e_2} \equiv x^{J_2+K_2\nu}$ herausziehen kann. Ich werde nun zeigen, daß es im Innern eines solchen Teilintervalls höchstens einen Wert ν geben kann, der Krümmungsordnung von Integralkurven sein könnte. Als überhaupt mögliche Krümmungsordnungen haben wir dann die Menge dieser rationalen Einteilungspunkte und die ebenfalls endliche Menge dieser noch zu bestimmenden inneren Punkte.

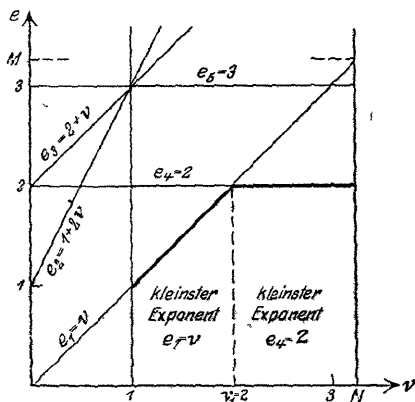


Fig. 14.

Um den eben ausgesprochenen Satz zu beweisen, betrachte ich die Differentialgleichung (9) in einem solchen Teilintervall. Darin ist:

$$P(x, x^\nu) = x^{e_1} \sum a_{ik} x^{i+k\nu-e_1}; \quad Q(x, x^\nu) = x^{e_2} \sum b_{ik} x^{i+k\nu-e_2}.$$

Also:

$$(10) \quad \nu' = \frac{x^{e_1} \sum a_{ik} x^{i+k\nu-e_1} - \nu x^{e_2+\nu-1} \sum b_{ik} x^{i+k\nu-e_2}}{x^{e_2+\nu} \lg x \sum b_{ik} x^{i+k\nu-e_2}}.$$

Ist $e_1 \leq e_2 + \nu - 1$, so kann dieser Quotient mit x^{e_1} gekürzt werden, andernfalls mit $x^{e_2+\nu-1}$. Dadurch erhält man:

$$\nu' = \frac{\mathfrak{Z}(x, \nu)}{x^\nu \mathfrak{X}(x, \nu) \lg x}; \quad \nu \geq 1; \quad \mathfrak{X}(0, \nu) = b_{J_2, K_2} \neq 0.$$

Ich werde nun zeigen, daß es nur an den Punkten $(0, \nu)$, an denen $\mathfrak{Z}(0, \nu) = 0$ ist, vorkommen kann, daß eine von der ν -Achse verschiedene

Integralkurve Punkte mit derselben gemeinsam hat. Dies ist evident, wenn $r > 1$ ist; denn in diesem Fall ist in allen Punkten, in denen $\mathfrak{Z}(0, \nu) \neq 0$ ist, die Lipschitzsche Bedingung in bezug auf ν als unabhängig Variable befriedigt. Ist dagegen $r = 1$, so strebt die dabei auftretende partielle Ableitung nach x gegen Unendlich wie $\lg x$. Trotzdem geht durch diesen

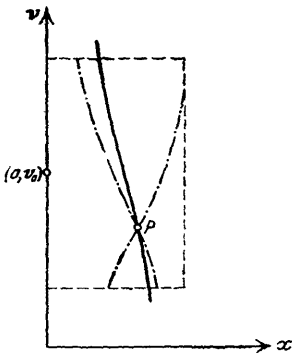


Fig. 15.

Punkt keine von $x = 0$ verschiedene Integralkurve. Denn um diesen Punkt gibt es einen Bereich (Fig. 15) $|\nu - \nu_0| \leq \varepsilon, 0 \leq x \leq \delta < 1$, in dem

$$|\nu'| > \frac{c}{x |\lg x|} = c \frac{1}{|\lg x|}.$$

Die Integralkurve durch einen Punkt $P(x, \nu)$ dieses Bereiches bleibt also zwischen den Kurven der beiden Scharen

$$\nu = C + c \lg |\lg x| \quad \text{und} \quad \nu = C - c \lg |\lg x|,$$

die durch den Punkt P gehen, geht also nicht durch den Punkt $(0, \nu_0)$.

Es können also nur in solche Punkte $(0, \nu)$ Integralkurven einmünden, für die $\mathfrak{Z}(0, \nu) = 0$ ist. Nun verschwindet $\mathfrak{Z}(0, \nu)$ nicht identisch (die ν -Achse ist also Integralkurve der Differentialgleichung (10)), denn es ist

für $e_1 < e_2 + \nu - 1$	$\mathfrak{Z}(0, \nu) \equiv a_{J_1, K_1} \neq 0,$
für $e_1 > e_2 + \nu - 1$	$\mathfrak{Z}(0, \nu) \equiv -\nu b_{J_2, K_2} \neq 0$ für $1 \leq \nu \leq A,$
für $e_1 = e_2 + \nu - 1$	$\mathfrak{Z}(0, \nu) \equiv a_{J_1, K_1} - \nu b_{J_2, K_2},$
	$\mathfrak{Z}(0, \nu) = 0$ für $\nu = \frac{a_{J_1, K_1}}{b_{J_2, K_2}}.$

Demnach ist es nur im letzten Falle, wenn also

$$e_1 \equiv J_1 + K_1 \nu = J_2 - 1 + \nu(K_2 + 1) \equiv e_2 + \nu - 1,$$

also $J_1 = J_2 - 1$ und $K_1 = K_2 + 1$ ist, möglich, daß $\mathfrak{Z}(0, \nu)$ eine Nullstelle im Innern des betrachteten Intervalls besitzt. Die so erhaltenen Werte von ν sind rational im Bereich der Koeffizienten von P und Q , während die anderen rational im Bereich der ganzen Zahlen sind. Jedenfalls gibt es auf der ν -Achse nur eine endliche Anzahl von Punkten, in die Integralkurven einmünden können. Dies bedeutet für die (x, y) -Ebene, daß jede Integralkurve, die oberhalb der Parabel $y = x^N$ bzw. unterhalb der Parabel $y = -x^N$ längs der x -Achse in die singuläre Stelle einmündet, eine eindeutig bestimmte Krümmungsordnung hat. Die Anzahl der dabei möglichen Krümmungsordnungen ist endlich

Bis jetzt wurden nur diejenigen Integralkurven berücksichtigt, deren Krümmungsordnung unterhalb einer festen Schranke N bleibt. Es fragt

sich nun, wie man eine solche Schranke N bestimmen kann. Da P und Q als teilerfremd vorausgesetzt wurden, enthalten nicht beide zugleich den Faktor y . Zwei Fälle sind zu unterscheiden:

1. $P(x, y) = \sum a_{i,k} x^i y^k$ enthalte den Faktor y nicht; dann gibt es in dieser Summe Glieder von der Form $a_{r,0} x^r$, und unter diesen ein Glied $a_{\rho,0} x^\rho$, dessen Exponent ρ von allen diesen Exponenten r der kleinste ist. Geht man in (2) ein mit der Transformation $y = u(x) x^{\rho+1}$, so erhält man, wenn man noch mit x^ρ kürzt:

$$u' = \frac{a_{\rho,0} + x R(x, u)}{x Q(x, u x^{\rho+1})}.$$

Diese Differentialgleichung ist aber in bezug auf u als unabhängige Variable regulär für $x = 0$. Also gibt es in einem Bereich $|u| < A$ keine Integralkurven, die mit der u -Achse Punkte gemeinsam haben und nicht mit ihr zusammenfallen. Für die (x, y) -Ebene bedeutet dies, daß in einem Bereich $-\Delta x^{\rho+1} \leq y \leq \Delta x^{\rho+1}$ keine Integralkurven in die singuläre Stelle einmünden. Die obere Schranke N kann also gleich $\rho + 1$ gesetzt werden.

2. $P(x, y)$ enthalte den Faktor y ; dann enthält ihn $Q(x, y) = \sum b_{i,k} x^i y^k$ nicht. Also gibt es in dieser Summe ein Glied $b_{\rho,0} x^\rho$, wo ρ wieder die Minimaleigenschaft hat. Geht man nun mit derselben Transformation $y = u x^{\rho+1}$ in (2) ein, so erhält man:

$$u' = \frac{P(x, u x^{\rho+1}) - (\rho + 1) u x^\rho Q(x, u x^{\rho+1})}{x^{\rho+1} Q(x, u x^{\rho+1})}.$$

Nun enthält $Q(x, u x^{\rho+1})$ den Faktor x^ρ ; also der zweite Bestandteil des Zählers den Faktor $x^{2\rho}$. Der erste enthalte den Faktor x^l . Es sind nun drei Fälle zu unterscheiden:

a) $l > 2\rho$; der Bruch kann mit $x^{2\rho}$ gekürzt werden. Dadurch erhält man:

$$u' = \frac{-(\rho + 1) u (b_{\rho,0} + x R_1(x, u))}{x (b_{\rho,0} + x R_2(x, u))}.$$

Nach § 5 ist in der (x, u) -Ebene die x -Achse die einzige Integralkurve dieser Differentialgleichung, die in einem Bereich $|u| \leq A$ Punkte mit der u -Achse gemein hat, ohne mit ihr zusammenzufallen. In der (x, y) -Ebene ist also die x -Achse die einzige Integralkurve, die aus dem Bereich $-\Delta x^{\rho+1} \leq y \leq \Delta x^{\rho+1}$ in die singuläre Stelle einmündet. Auch hier kann $N = \rho + 1$ gesetzt werden.

b) $l < 2\rho$; da $l = J + K(\rho + 1) < 2\rho$ ist, ist notwendigerweise $K = 1$ und $J + 1 < \rho$. Der Bruch kann mit x^l gekürzt werden; dadurch erhält man:

$$u' = \frac{a_{J,1} u + u x R_1(x, u)}{x^{2\rho-l+1} (b_{\rho,0} + x R_2(x, u))}.$$

Diese Differentialgleichung kann nach § 5 in einem Bereich $|u| \leq A$, $0 \leq x \leq \delta$ untersucht werden. Ist $\frac{a_{J1}}{b_{e0}} < 0$, so erhält man topologisch dasselbe Ergebnis wie im Fall a). Ist dieses Verhältnis aber positiv, so münden alle Integralkurven durch den Rand des Bereiches $-Ax^{e+1} \leq y \leq Ax^{e+1}$, $0 < x \leq \delta$ in diesem Bereich in die singuläre Stelle ein. Wendet man anstatt der Transformation $y = ux^{e+1}$ irgendeine andere $y = ux^{e+i}$ ($i > 1$) an, so findet man, daß diese Integralkurven die Krümmungsordnung Unendlich besitzen, sich also der x -Achse näher anschmiegen als irgendeine Parabel. (Beispiel: $y' = \frac{2y}{x^3}$. Lösungen: $y = ce^{-\frac{1}{x^2}}$.) Setzen wir auch hier $N = e + 1$, so schließen wir nur Integralkurven unendlich hoher Krümmungsordnung aus.

c) $l = 2e$; also $J = e - 1$; $K = 1$. In diesem Fall wende man die Transformation $y = ux^{e+i}$ ($i > 1$) an. Aus dem zweiten Teil des Zählers kann dann der Faktor x^{2e+i-1} herausgezogen werden, aus dem ersten der Faktor $x^{J+e+i} = x^{2e+i-1}$. Also kann mit diesem Faktor gekürzt werden. Dadurch erhält man:

$$u' = \frac{u(a_{J1} - (e+i)b_{e0}) + uxR_1(x, u)}{x(b_{e0} + xR_2(x, u))}.$$

Nun wähle man i so groß, daß der Quotient $\frac{a_{J1} - (e+i)b_{e0}}{b_{e0}}$ negativ wird; dann treten dem Fall a) analoge Verhältnisse ein. Die Zahl $e + i$ wähle man dann als obere Schranke N .

Jetzt ist diese Schranke N in allen Fällen so bestimmt, daß nur die Integralkurven mit unendlich hoher Krümmungsordnung ausgeschlossen werden. Aus diesen Untersuchungen sehen wir auch zugleich, daß die Anzahl der überhaupt möglichen Krümmungsordnungen endlich ist; denn unterhalb N sind es nur endlich viele und oberhalb N höchstens noch die Ordnung ∞ . Wir können nach dem Vorhergehenden auch immer feststellen, ob solche Integralkurven von der Krümmungsordnung Unendlich vorkommen; es handelt sich also im weiteren nur noch um die Integralkurven mit endlicher Krümmungsordnung ν .

Es sei nun ν_1 der Wert einer solchen möglichen Krümmungsordnung. Um das Krümmungsmaß der Integralkurven von der Ordnung ν_1 festzustellen, gehe man in die Differentialgleichung (2) ein mit dem Ansatz: $y = u(x)x^{\nu_1}$. Dadurch erhält man:

$$(11) \quad u' = \frac{P(x, ux^{\nu_1})}{x^{\nu_1} Q(x, ux^{\nu_1})} - \nu_1 \frac{u}{x} = \frac{x^{\lambda_1} (P_1(u) + x^{\lambda_1} R_1(x, u))}{x^{\lambda_2} (P_2(u) + x^{\lambda_2} R_2(x, u))},$$

wobei P_1 und P_2 Polynome in u sind. Es sind nun zwei Fälle zu unterscheiden:

1. $l_1 > l_2 - 1$; die Bedingungen des § 2 (Sondertypus, vgl. Gl. (8)) sind befriedigt. Zu jedem Wert u , für den $P_2(u) \neq 0$ ist, gehört eine und nur eine Integralkurve. Da in den in § 6 und § 7 konstruierten Bereichen keine zur y -Achse parallele Feldrichtung vorkommt, ist in den jetzt zu untersuchenden Fällen $P_2(u)$ immer von Null verschieden.

Dieser Fall tritt insbesondere immer dann ein, wenn die Zahl ν_1 nicht aus den Exponenten, sondern aus den Koeffizienten der Differentialgleichung berechnet wurde. Denn nach Seite 245/246 ist dann:

$$P(x, u x^{\nu_1}) = x^{e_1} (a_{J, K_1} u^{K_1} + x^{\lambda_1} R_1(x, u)),$$

$$Q(x, u x^{\nu_1}) = x^{e_2} (b_{J_2, K_2} u^{K_2} + x^{\lambda_2} R_2(x, u)).$$

Setzt man diese Werte in (11) ein und beachtet dabei, daß $e_1 = e_2 + \nu_1 - 1$,

$\nu_1 = \frac{a_{J_1, K_1}}{b_{J_2, K_2}}$ und $K_1 = K_2 + 1$ ist, so erhält man:

$$\begin{aligned} u' &= \frac{x^{e_1} (a_{J_1, K_1} u^{K_1} - \nu_1 b_{J_2, K_2} u^{K_2+1} + x^{\lambda_1} R_1 - x^{\lambda_2} \nu_1 u x^{\nu_1-1} R_2)}{x^{e_1+1} (b_{J_2, K_2} u^{K_2} + x^{\lambda_2} R_2)} \\ &= \frac{x^{\lambda_1} R_1 - x^{\lambda_2} \nu_1 u x^{\nu_1-1} R_2}{x (b_{J_2, K_2} u^{K_2} + x^{\lambda_2} R_2)}. \end{aligned}$$

Es ist also: $l_2 - 1 = 0$, $l_1 = \lambda_1$ oder $l_1 = \lambda_2 + \nu_1 - 1$. Folglich ist in diesem Falle stets: $l_1 > 0$, und damit $l_1 > l_2 - 1$. Ferner ist $P_2(u) \equiv b_{J_2, K_2} u^{K_2}$, verschwindet also nur für $u = 0$. Wir haben demnach das Ergebnis, daß jedesmal, wenn eine Parabel mit irrationaler Krümmungsordnung ν_1 Schmiegungsparabel einer Integralkurve ist, längs jeder Parabel der Schar $y = p x^{\nu_1}$ ($p \neq 0$) eine und nur eine Integralkurve in die singuläre Stelle einmündet. (Gleichzeitig sieht man, was für später wichtig ist, daß für $\varepsilon \leq |u| \leq N$, $0 \leq x \leq \delta$ die partielle Ableitung $\frac{\partial u'}{\partial u}$ einer Ungleichung: $\left| \frac{\partial u'}{\partial u} \right| \leq \frac{M}{x}$ genügt.) (Man vergleiche auch § 10, 8. Beispiel.)

2. $l_1 \leq l_2 - 1$; nur längs der Parabeln $y = u_i x^{\nu_i}$, für deren Koeffizienten u_i das Polynom $P_1(u)$ verschwindet, können Integralkurven in die singuläre Stelle einmünden. Es gibt also nur eine endliche Anzahl solcher Parabeln, die Schmiegungsparabeln von Integralkurven sein können.

Ebenso wie vorhin diejenigen Integralkurven speziell behandelt wurden, deren Krümmungsordnung Unendlich ist, werden nun auch diejenigen speziell untersucht, deren Krümmungsordnung ν_1 zwar endlich, deren Krümmungsmaß u aber Null oder Unendlich ist. Durch die Transformation $y = u x^{\nu_1}$ wird eine Kurve $y = c x^{\nu_2}$ abgebildet in eine Kurve $u = c x^{\nu_2 - \nu_1}$ und umgekehrt. Eine Kurve vom Krümmungsmaß Null und der Krümmungsordnung ν_1 wird demnach abgebildet in eine Kurve, die

durch den Punkt $(0, 0)$ geht, sich aber dort näher an die u -Achse anschmiegt als irgendeine Parabel $u = cx^v$. Mit Integralkurven dieser Art haben wir uns aber vorhin beschäftigt und haben gefunden, daß es, so oft solche auftreten, einen Bereich $0 \leq x \leq Au^{e+1}$, $0 \leq u \leq \delta$ gibt, der frei von Randsingularitäten ist. In der (x, y) -Ebene begrenzen also in diesem

Fall die Parabeln $y = \delta x^v$ und $y = \left(\frac{1}{A}\right)^{\frac{1}{e+1}} x^{v+\frac{1}{e+1}} = cx^{v+\varepsilon}$ einen Bereich ohne Randsingularitäten. Um also zu untersuchen, ob Integralkurven vom Krümmungsmaß Null vorkommen, genügt es, festzustellen, ob die Parabel $y = \delta x^v$ mit einer Parabel $y = cx^{v+\varepsilon}$ einen Bereich ohne Randsingularitäten bildet (δ und ε hinreichend klein). (Diese Feststellung wird nach unseren Voraussetzungen durch algebraische Methoden erledigt.) (Vgl. § 10, 1. Beispiel.)

Um auf Integralkurven vom Krümmungsmaß Unendlich zu untersuchen, wende man die Transformation $y = \frac{1}{u} x^v$ an. Ohne weiteres erhält man als entsprechendes Ergebnis, daß solche Integralkurven nur dann vorkommen, wenn die Parabeln $y = Nx^v$ und $y = cx^{v-\varepsilon}$ (N hinreichend groß und ε hinreichend klein) einen Bereich ohne Randsingularitäten bilden.

Zusammenfassend haben wir das Ergebnis, daß auch das Krümmungsmaß der Integralkurven eindeutig bestimmt, und zwar null, endlich oder unendlich ist.

Mit der Bestimmung der Schmiegungsparabel kennt man die Gestalt der Integralkurven in den erledigten Sonderfällen (Sondertypus, Krümmungsordnung Unendlich, Krümmungsmaß Null und Unendlich). Es bleibt eine endliche Anzahl von Parabeln $y = u_1 x^v$ übrig, die Schmiegungsparabeln von Integralkurven sein können und in deren Umgebung wir die Gestalt der Integralkurven noch nicht kennen. Zur weiteren Untersuchung gehe man nun in die ursprüngliche Differentialgleichung ein mit dem Ansatz $y = u_1 x^v + x^{\nu(x)}$, bestimme wiederum die singulären Punkte $(0, \nu_2)$ der ν -Achse und wende dann die Transformation $y = u_1 x^v + u(x)x^{\nu_2}$ an, um die singulären Punkte $(0, u_2)$ der u -Achse zu bestimmen. Dabei können alle, aber auch nur die Fälle wieder eintreten, die bei den einfacheren Transformationen unterschieden wurden. Dieses Verfahren kann fortgesetzt werden, wenn man nicht auf einen der obigen Sonderfälle geführt wird. In diesem Fall kennen wir aber die Gestalt der Integralkurven in der Umgebung der Schmiegungsparabel. Die andern Schmiegungsparabeln können wir auf beliebig viele Glieder genau bestimmen.

Nach diesen Vorbereitungen können wir nun an das eigentliche Entscheidungsproblem herangehen. Es handelt sich dabei darum, die Gestalt

der Integralkurven in einem, die x -Achse als ausgezeichnete Richtung umgebenden Sektor, oder einem Teilbereich davon (siehe § 7), zu untersuchen. Zu diesem Zweck stelle man zunächst fest, ob es im Innern dieses Bereiches Parabeln gibt, an die sich Integralkurven anschmiegen können. Außerdem untersuche man, ob es Integralkurven von der Krümmungsordnung Unendlich oder auch vom Krümmungsmaß Null oder Unendlich gibt. Gibt es im Bereich eine unendliche Anzahl solcher Schmiegeparabeln, so haben wir den Fall des Sondertypus; es münden also in diesem Bereich unendlich viele Integralkurven in die singuläre Stelle ein. Andernfalls

haben wir nur eine endliche Anzahl solcher Parabeln. Es sei $\eta = \sum_{i=1}^n u_i x^i$

eine davon. Um die Gestalt der Integralkurven in ihrer Umgebung zu finden, wendet man auf die Differentialgleichung (2) die Transformation

$y = \sum_{i=1}^{n-1} u_i x^i + u(x) x^n$ an. Die Differentialgleichung der neuen (x, u) -

Ebene hat dann im Punkt $(0, u_n)$ eine singuläre Stelle. Da aber die Parabel η im Innern eines Bereiches liegt, in welchem keine zur y -Achse parallele Feldrichtung vorkommt, gibt es einen Bereich $0 < x \leq \delta; |u - u_n| \leq \varepsilon$, in welchem der Nenner der neuen Differentialgleichung nicht verschwindet. Diese Parabel η ist also das Analogon einer regulären, ausgezeichneten Richtung (vgl. § 5). Als neue Differentialgleichung erhält man:

$$u' x^n = y' - \sum_{i=1}^{n-1} u_i v_i x^{i-1} - v_n u x^{n-1}.$$

Die rechte Seite dieser Differentialgleichung ist aber die Funktion Feldrichtung minus Parabelrichtung, die in Anlehnung an das Vorhergehende mit $\Psi_n(u, x)$ bezeichnet werden soll. Es ist also:

$$u' x^n = \Psi_n(u, x); \quad u' x = \frac{\Psi_n(u, x)}{x^{n-1}} = \psi_n(u, x).$$

Nun sind drei Fälle zu unterscheiden:

1. $\frac{\partial \psi_n}{\partial u}(u_n, 0) > 0$; längs dieser Parabel münden unendlich viele Integralkurven in die singuläre Stelle ein (vgl. § 5).

2. $\frac{\partial \psi_n}{\partial u}(u_n, 0) < 0$; nach § 5 gibt es nur eine Integralkurve, die längs dieser Parabel einmündet.

3. $\frac{\partial \psi_n}{\partial u}(u_n, 0) = 0$; diese Parabel ist das Analogon einer mehrfachen, regulären ausgezeichneten Richtung. Deshalb versuche man, die Schmiegeparabel näher zu bestimmen. Ist dies nicht möglich, weil es kein weiteres v gibt oder das Krümmungsmaß Null oder Unendlich ist, so kann man nach dem Vorhergehenden die Gestalt der Integralkurven in der Umgebung

dieser Parabel ermitteln. Andernfalls sei $\eta_1 = \eta + u_{n+1} x^{r_{n+1}}$ eine Näherungsparabel höherer Ordnung. Für die Umgebung dieser Parabel haben wir dieselben Fallunterscheidungen zu machen. Es ist nun zu zeigen, daß dieses Verfahren zum Ziele führt, daß man also nach einer endlichen Anzahl von Schritten zu einer Schmiegungsparabel kommen muß, die einer einfachen, regulären ausgezeichneten Richtung entspricht.

Setzt man $y' = f(x, y)$, so ist $\frac{\partial \psi_n}{\partial u}(u, x) = \frac{\partial f}{\partial y} x - v_n$. Ist nun $\frac{\partial \psi_n}{\partial u}(u, 0)$ in der Umgebung von $u = u_n$ eine stetige Funktion von u , so ist

$$\lim_{\substack{y=\eta_1 \\ x \rightarrow 0}} \frac{\partial f}{\partial y} x = \lim_{\substack{y=\eta \\ x \rightarrow 0}} \frac{\partial f}{\partial y} x = v_n,$$

denn die Kurven η_1 und η gehören zu demselben Wert u_n . Durch die Transformation $y = \eta + u x^{r_{n+1}}$ erhält man also eine Funktion $\psi_{n+1}(u, x)$, deren Ableitung

$$\frac{\partial \psi_{n+1}}{\partial u}(u_{n+1}, 0) = v_n - v_{n+1} < 0$$

ist. Längs dieser Parabel mündet also nur eine Integralkurve in den Ursprung ein (§ 5). In diesem Falle führt also schon der zweite Schritt zum Ziel.

Im andern Falle ist $\frac{\partial \psi_n}{\partial u}(u, 0)$ in der Umgebung von $u = u_n$ unstetig. Nun ist

$$\frac{\partial f}{\partial y} x = \frac{(Q P_y - P Q_y) x}{Q^2} = \frac{Z(x, y)}{N(x, y)}.$$

Setzt man $y = \sum_{i=1}^{n-1} u_i x^{r_i} + u(x) x^{r_n}$, so geht $Z(x, y)$ über in $x^r \mathfrak{Z}(x, u)$, $N(x, y)$ in $x^s \mathfrak{N}(x, u)$. Dabei sind $\mathfrak{Z}(x, u)$ und $\mathfrak{N}(x, u)$ stetige Funktionen der Variablen, die für $x = 0$ nicht identisch verschwinden. Insbesondere ist $\mathfrak{N}(0, u_n) \neq 0$, und s unabhängig von n . Wenn also $\frac{\partial \psi_n}{\partial u}(u, 0)$ bei $u = u_n$ unstetig ist, so muß notwendigerweise $s > r$ und $\mathfrak{Z}(0, u_n) = 0$ sein. Nun ist $\lim_{\substack{y=\eta \\ x \rightarrow 0}} \frac{\partial f}{\partial y} x = v_n$. Da aber $N(x, u)$ wie x^s verschwindet, muß $Z(x, \eta)$

ebenfalls wie x^s verschwinden. Es ist also $|Z(x, \eta)| < C x^s$. Man bestimme nun in der (x, y) -Ebene diejenigen Gebiete im betrachteten Bereich, für die $|Z(x, y)| < C x^s$ ist. In einem dieser Gebiete muß meine Näherungsparabel η liegen, vorausgesetzt, daß ich in hinreichender Nähe des Ursprungs bleibe. Die Kurven, die dieses Gebiet begrenzen, seien y_1 und y_2 . Dann ist $|Z(x, y_1)| = |Z(x, y_2)| = C x^s$. Auf Grund dieser Gleichungen können y_1 und y_2 in allgemeine Potenzreihen entwickelt werden. Dadurch erhält man $y_1 = \sum a_i x^{e_i}$ und $y_2 = \sum b_i x^{e_i}$. Da y_1 und y_2 nicht

identisch sein können, können diese Entwicklungen auch nur in endlich vielen Gliedern miteinander übereinstimmen. Man kann also n so groß wählen, daß für $i < n$ $a_i = b_i$ und $\varrho_i = \sigma_i$ ist, während entweder ϱ_n und σ_n oder a_n und b_n voneinander verschieden sind. Da nun η zwischen y_1 und y_2 liegt, ist notwendigerweise $u_i = a_i = b_i$; $v_i = \varrho_i = \sigma_i$ für $i < n$. Es sei nun $\varrho_n = \sigma_n$. (Ist $\varrho_n > \sigma_n$ resp. $\sigma_n > \varrho_n$, so setze man $a_n = 0$ resp. $b_n = 0$.) Dann ist $a_n \neq b_n$. Ist nun $v_n = \varrho_n$, so liegt u_n zwischen a_n und b_n . Alle Kurven $y = \sum_{i=1}^{n-1} u_i x^{v_i} + u x^{v_n}$, deren Parameterwert u zwischen a_n und b_n liegt, verlaufen also zwischen y_1 und y_2 ; für sie ist also $Z(x, y) = x^s \bar{Z}(x, u)$.

Ist aber $v_n > \varrho_n$, so liegen die Nachbarkurven von $\eta = \sum_{i=1}^n u_i x^{v_i}$ ebenfalls zwischen y_1 und y_2 , weil die Differenz $|\eta - y_1|$ oder $|\eta - y_2|$ klein wird wie x^{v_n} , während die Differenz zweier Nachbarkurven klein wird wie x^{v_n} .

Setzt man also $y = \sum_{i=1}^{n-1} u_i x^{v_i} + u x^{v_n}$, wo n genügend groß ist, so hat $Z(x, y)$ den Faktor x^s . Also wird $\frac{\partial y_n}{\partial u}(u, 0)$ in der Umgebung von $u = u_n$ eine stetige Funktion von u . Nach dem Vorhergehenden entspricht dann die Parabel $\eta = \sum_{i=1}^{n+1} u_i x^{v_i}$ notwendigerweise einer einfachen, regulären ausgezeichneten Richtung, längs der eine und nur eine Integralkurve in die singuläre Stelle einmündet. Damit ist nachgewiesen, daß man in allen Fällen nach einer endlichen Anzahl von Schritten entweder zum Sonder-typus oder zu solchen Schmiegungsparabeln kommt, die einfachen, regulären, ausgezeichneten Richtungen entsprechen.

Damit ist das Entscheidungsproblem gelöst. Die Lösung erfolgt durch die Bestimmung der „Normalbereiche“. Darunter verstehen wir Bereiche, die durch Parabeln von der Form $y = \sum u_i x^{v_i}$ begrenzt sind, und in denen

- a) alle Integralkurven,
- b) eine Integralkurve,
- c) keine Integralkurven

in die singuläre Stelle einmünden. Durch paarweises Zusammensetzen dieser Normalbereiche erhält man die Brouwerschen Sektoren⁵⁾ als topologisch invariante Grundelemente für das Integralkurvenfeld. Der Bereich c) hat dabei die Eigenschaft, daß er, mit einem andern Bereich zusammengesetzt, dessen Charakter nicht ändert. Besteht die Umgebung der singulären Stelle nur aus Bereichen c), so liegt der „zweite Hauptfall“ vor. Wird der Bereich a) mit einem gleichartigen zusammengenommen, so erhält man den „elliptischen Sektor“ (Fig. 8); wird er mit b) zusammengenommen, so erhält man den „parabolischen Sektor“ (Fig. 9); wird b) mit einem gleich-

artigen zusammengenommen, so erhält man den „hyperbolischen Sektor“ (Fig. 10). Es ist also möglich, die Reihenfolge und Anordnung der topologischen Grundelemente zu bestimmen.

Da es sich bei der praktischen Durchführung nur darum handelt, ob aus einem Bereich endlich oder unendlich viele Integralkurven in die singuläre Stelle einmünden, genügt es also, nach Bereichen ohne Randsingularitäten zu suchen; denn nur in diesem Fall kann es nach obigem vorkommen, daß aus einem Bereich unendlich viele Integralkurven in den Ursprung einmünden. Die Begrenzungen sind Parabeln von der Form $y = \sum u_i x^{v_i}$.

Es wurde zunächst vorausgesetzt, daß P und Q Potenzreihen sind, doch brauchte man für das Entscheidungsproblem nur ein endliches Polynom zu berücksichtigen. Die Überlegungen gelten daher für alle Funktionen P und Q , die in der Nähe des Ursprungs durch ein Polynom genügend hohen Grades approximierbar sind.

§ 9.

Quantitative Auswertung der Methode.

Im vorhergehenden Abschnitt wurde gezeigt, daß man mit Hilfe der Methode der Randsingularitäten und der allgemeinen Transformationen jeweils die Gestalt der Integralkurven bestimmen kann. Ich werde nun noch zeigen, daß diese Methoden auch eine analytische Darstellung derselben ermöglichen, wenn man voraussetzt, daß $P(x, y)$ und $Q(x, y)$ konvergente Potenzreihen sind.

Es wurde schon darauf hingewiesen, daß man die Näherungsparabeln auf beliebig viele Glieder genau bestimmen kann, wenn nicht einer der „Sonderfälle“ eintritt. Als Sonderfälle wurden bezeichnet:

1. Der Fall des Sondertypus, bei dem, mit Ausnahme einer endlichen Anzahl, jede Parabel der Schar $\eta(x, p) = \sum_{i=1}^{n-1} u_i x^{v_i} + p x^n$ Schmiegungsparabel einer und nur einer Integralkurve ist. Es ist ohne weiteres klar, daß man auch hier von jeder Parabel das nächste Glied mit Hilfe der Transformationen

$$y = \sum_{i=1}^{n-1} u_i x^{v_i} + c x^{v_n} + x^{v(x)} \quad \text{und} \quad y = \sum_{i=1}^{n-1} u_i x^{v_i} + c x^{v_n} + u(x) x^{v_{n+1}}$$

bestimmen kann. Dabei ist zu beachten, daß sich die „singulären“ Parabeln eventuell auf mehrere Arten fortsetzen lassen. Diese Schmiegungsparabeln lassen sich demnach auch beliebig genau bestimmen.

2. Der Fall der unendlich hohen Krümmungsordnung.

3. Der Fall des Krümmungsmaßes Null oder Unendlich, der aber auf den Fall unendlich hoher Krümmungsordnung zurückgeführt wurde.

Wir haben demnach zwei Typen von Integralkurven zu unterscheiden:

a) solche, deren Schmiegungsparabel $\eta = \sum_{i=1}^n u_i x^{\nu_i}$ auf beliebig viele Glieder bestimmt werden kann;

b) solche, die sich an eine Schmiegungsparabel $y = \sum_{i=1}^n u_i x^{\nu_i}$ näher anschmiegen als irgendeine Parabel höherer Ordnung.

Betrachtet man nun Integralkurven vom ersten Typus, so wissen wir, daß man nach einer endlichen Anzahl von Schritten zu einer Parabel

$\eta_n = \sum_{i=1}^n u_i x^{\nu_i}$ kommt, die einer einfachen, regulären, ausgezeichneten Richtung entspricht. Geht man also in die Differentialgleichung $y' = \frac{P(x, y)}{Q(x, y)}$

ein mit der Transformation $y = \sum_{i=1}^{n-1} u_i x^{\nu_i} + (u(x) - u_n) x^{\nu_n}$, so gehen P

und Q über in Reihen von der Form $\sum f_i(u) x^{\alpha_i}$, wobei die f_i Polynome in u sind und die α_i sich ganzzahlig linear aus den Exponenten $\nu_1, \nu_2, \dots, \nu_n$ zusammensetzen. Da aber die Parabel η_n einer regulären ausgezeichneten Richtung entspricht, wissen wir außerdem, daß in der Reihe $Q(x, y) = x^l \{a_0 + x^{\alpha_1} f_1(u) + \dots\}$ das erste Glied a_0 eine von Null verschiedene Konstante ist. Es ist demnach möglich, den Ausdruck

$\frac{1}{a_0 + x^{\alpha_1} f_1(u) + \dots}$ wiederum in eine Reihe zu entwickeln, in der die Ex-

ponenten von x ebenfalls ganzzahlig linear aus $\nu_1, \nu_2, \dots, \nu_n$ zusammengesetzt und deren Koeffizienten $f(u)$ wiederum Polynome in u sind. Bildet man nun die Funktion $\Psi_n(u, x)$, die die Differenz Feldrichtung

minus Richtung der Parabel $y = \sum_{i=1}^{n-1} u_i x^{\nu_i} + (u(x) - u_n) x^{\nu_n}$ darstellt, so erhält man:

$$\Psi_n(u, x) = x^{\alpha_n} f_n(u) + x^{\beta_n} g_n(u) + x^{\gamma_n} h_n(u) + \dots$$

Die Differentialgleichung in der (u, x) -Ebene lautet dann

$$u' x^{\nu_n} = \Psi_n(u, x) = x^{\alpha_n} f_n(u) + x^{\beta_n} g_n(u) + x^{\gamma_n} h_n(u) + \dots$$

oder

$$(12) \quad u' = x^{\alpha'_n} f'_n(u) + x^{\beta'_n} g'_n(u) + x^{\gamma'_n} h'_n(u) + \dots,$$

wobei auch die Exponenten $\alpha'_n, \beta'_n \dots$ ganzzahlig linear aus $\nu_1, \nu_2, \dots, \nu_n$ zusammengesetzt sind. Um ν_{n+1} zu erhalten, geht man in die Differentialgleichung (12) ein mit dem Ansatz $u = v x^d$. Dadurch erhält man:

$$v' x^d + d v x^{d-1} = x^{\alpha'_n} f'_n(v x^d) + x^{\beta'_n} g'_n(v x^d) + x^{\gamma'_n} h'_n(v x^d) + \dots$$

Da die Parabel η_n einer *einfachen* regulären ausgezeichneten Richtung entspricht, beginnt $f_n(u)$ mit einem linearen Glied cu . Es ist also:

$$(13) \quad v'x^d = -dvx^{d-1} + cvx^{a_n+d} + x^{a_n}f'_n(vx^d) + x^{\beta_n}g_n(vx^d) + \dots$$

Nach § 8 wird nun die Zahl $d = \nu_{n+1} - \nu_n$ entweder aus den Koeffizienten von (13) oder aus den Exponenten bestimmt. Erhält man sie aber aus den Koeffizienten, so wissen wir nach Seite 249, daß jede Parabel $y = \eta_n + cx^{\nu_{n+1}}$ ($c \neq 0$) Schmiegungsparabel einer und nur einer Integralkurve ist.

Nehmen wir nun an, die Parabel η_n sei schon so beschaffen, daß sie die Schmiegungsparabel *nur einer* Integralkurve ist, so wird offenbar die Differenz d aus den in der obigen Differentialgleichung vorkommenden Exponenten berechnet und die Parabel η_{n+1} entspricht wiederum einer einfachen regulären ausgezeichneten Richtung, längs der nur eine Integralkurve in die singuläre Stelle einmündet. Ist dabei η_n nicht selber Lösung der Differentialgleichung, so gibt es in (13) Glieder von der Form $a_i x^{i\lambda'_n}$, und unter diesen Gliedern ein Glied $a x^{\lambda'_n}$, dessen Exponent λ'_n von allen Exponenten λ'_{i_n} der kleinste ist. Dann ist nach § 8

$$d = \lambda'_n - \alpha'_n, \quad \text{wenn } \alpha'_n < -1,$$

$$\text{oder} \quad d = \lambda'_n + 1, \quad \text{„ } \alpha'_n \geq -1 \text{ ist.}$$

Da sich nun λ'_n und α'_n ganzzahlig linear aus $\nu_1, \nu_2, \dots, \nu_n$ zusammensetzen, können wir dasselbe über d und hiermit auch über ν_{n+1} aussagen. Weil aber die Parabel η_{n+1} dieselben Eigenschaften hat, wie η_n , so folgt hieraus, daß ν_{n+2} ganzzahlig linear aus $\nu_1, \nu_2, \dots, \nu_{n+1}$ und damit aus $\nu_1, \nu_2, \dots, \nu_n$ zusammengesetzt ist. Also sind alle Exponenten ν_i der Reihe $\sum_{i=1}^{\infty} u_i x^{\nu_i}$ ganzzahlig linear aus $\nu_1, \nu_2, \dots, \nu_n$ zusammensetzbar.

Sind nun $\nu_1, \nu_2, \dots, \nu_n$ rational, so haben die Zahlen $(1, \nu_1 \dots \nu_n)$ ein größtes gemeinschaftliches Maß ϱ . Alle Exponenten der Reihe $\sum_{i=1}^{\infty} u_i x^{\nu_i}$, wo die ν_i monoton wachsen, sind also ganzzahlige Vielfache von ϱ ; mithin ist $\lim_{i \rightarrow \infty} \nu_i = \infty$.

Sind aber die Exponenten $\nu_1, \nu_2, \dots, \nu_n$ nicht alle rational, so kann man annehmen, ν_n sei die erste und damit die einzige irrationale Zahl; denn nach Seite 249 entspricht jede solche Parabel $y = \eta_{n-1} + cx^{\nu_n}$ ($c \neq 0$) einer einfachen regulären ausgezeichneten Richtung, längs der nur eine Integralkurve in die singuläre Stelle einmündet. Also sind nach dem Vorhergehenden alle Exponenten der Reihe $\sum_{i=1}^{\infty} u_i x^{\nu_i}$ ganzzahlig linear durch $\nu_1, \nu_2, \dots, \nu_n$ darstellbar. Ist also ϱ das größte gemeinschaftliche Maß

der Zahlen $(1, \nu_1 \dots \nu_{n-1})$, so ist $\nu_m = i\rho + k\nu_n$, wobei i und k ganze Zahlen unabhängig von ν_n sind. Setzt man für ν_n einen rationalen Näherungswert $\bar{\nu}_n$ ein, so erhält man für ν_m den rationalen Näherungswert $\bar{\nu}_m = i\rho + k\bar{\nu}_n$. Da nun $\lim_{m \rightarrow \infty} \bar{\nu}_m = \infty$ ist, ist auch $\lim_{m \rightarrow \infty} \nu_m = \infty$.

Bis jetzt wurde vorausgesetzt, daß die Parabel η_n einer einfachen, regulären ausgezeichneten Richtung entspricht, längs der nur eine Integralkurve in die singuläre Stelle einmündet. Schmiegen sich nun an η_n unendlich viele Integralkurven an, so kann η_n Integralkurve sein. In der Gleichung (13) gibt es dann kein Glied von der Form αx^{ν_n} ; es ist also nicht möglich, die Zahl d aus den Exponenten zu berechnen. Wird sie aber aus den Koeffizienten berechnet, so schmiegt sich an jede Parabel $y = \eta_n + cx^{\nu_{n+1}}$ ($c \neq 0$) eine und nur eine Integralkurve an; dieser Fall ist also auf den vorhergehenden zurückgeführt. Läßt sich aber d auch nicht aus den Koeffizienten berechnen, so schmiegen sich nach § 8 die Integralkurven an die Parabel η_n näher an, als an jede andere Parabel.

Ist aber η_n nicht Integralkurve, so wird ν_{n+1} entweder aus den Exponenten oder aus den Koeffizienten berechnet. Werden alle ν_i aus den Exponenten berechnet, so ist nach der vorhergehenden Überlegung ν_i ein ganzzahliges Vielfaches vom größten gemeinschaftlichen Maß ρ der Zahlen $(1, \nu_1, \nu_2, \dots, \nu_n)$. Ist aber ν_m ein Exponent, der aus den Koeffizienten berechnet wird, so ist

$$\text{für } i < m \quad \nu_i = k\rho \quad (k \text{ positive ganze Zahl}),$$

$$\text{für } i > m \quad \nu_i = k\rho + l\nu_m \quad (k \text{ und } l \text{ ganze Zahlen}).$$

Also ist auch hier $\lim_{i \rightarrow \infty} \nu_i = \infty$.

Die bisherigen Ergebnisse können wir folgendermaßen zusammenfassen: Wird die Schmiegungsparabel einer Integralkurve hinreichend weit bestimmt, so erhält man eine Reihe $\sum u_i x^{\nu_i}$. Mit zunehmendem i wachsen dabei die Exponenten ν_i über jede Schranke und alle Exponenten ν_i sind entweder ganzzahlige Vielfache einer rationalen Zahl ρ oder sind linear darstellbar durch eine Irrationalität.

Nun werde ich zeigen, daß, wenn die Reihe $\sum_{i=1}^{\infty} u_i x^{\nu_i}$ konvergiert, die Kurve $\varphi(x) = \sum_{i=1}^{\infty} u_i x^{\nu_i}$ Lösung der Differentialgleichung ist. Aus Gleichung (13) erhält man durch Multiplikation mit x^{ν_n} die Funktion Feldrichtung minus Parabelrichtung:

$$\begin{aligned} \Psi_{n+1}(v, x) &= v' x^{\nu_{n+1}} = -d v x^{\nu_{n+1}-1} + c v x^{\alpha_n+d} + \dots \\ &= x^{\alpha_{n+1}} f_{n+1}(v) + x^{\beta_{n+1}} g_{n+1}(v) + \dots \end{aligned}$$

Daraus ersieht man, daß

$$\alpha_{n+1} = \nu_{n+1} - 1 \quad \text{ist, wenn} \quad \nu_{n+1} - 1 < \alpha_n + d$$

oder

$$\alpha_{n+1} = \alpha_n + d, \quad \text{wenn} \quad \nu_{n+1} - 1 \geq \alpha_n + d \quad \text{ist.}$$

Hieraus folgt nun, daß der niedrigste Exponent α_m ($m > n$) in $\Psi_m(u, x)$ entweder $\nu_m - 1$ oder $\alpha_n + \nu_m - \nu_n$ ist. Also ist $\lim_{m \rightarrow \infty} \alpha_m = \infty$.

Für die Feldrichtung längs der Parabel η_m erhält man aus der ursprünglichen Differentialgleichung $y' = f(x, y)$ eine Reihe, die nach den Voraussetzungen innerhalb der Konvergenzkreise von P und Q konvergiert, bis η_m die Kurve $Q(x, y) = 0$ schneidet. Diese Reihe stimmt aber nach dem Vorhergehenden mit η'_m in allen Gliedern überein, deren Exponenten kleiner als α_m sind. Da nun $\lim_{m \rightarrow \infty} \alpha_m = \infty$ ist, ist auch $\lim_{m \rightarrow \infty} \Psi_m(\eta_m, x) = 0$, folglich:

$$\begin{aligned} \varphi'(x) &= \lim_{m \rightarrow \infty} \eta'_m = \lim_{m \rightarrow \infty} [f(x, \eta_m) - \Psi_m(\eta_m, x)] \\ &= f(x, \lim_{m \rightarrow \infty} \eta_m) - \lim_{m \rightarrow \infty} \Psi(\eta_m, x) = f(x, \lim_{m \rightarrow \infty} \eta_m) = f(x, \varphi(x)). \end{aligned}$$

Die Integralkurve $\varphi(x)$ wird also in diesem Fall durch eine Reihe $\sum u_i x^{v_i}$ dargestellt, die so lange konvergiert, bis $\varphi(x)$ den Konvergenzbereich von P und Q verläßt oder die Kurve $Q(x, y) = 0$ schneidet.

Da wir wissen, daß in allen Fällen $\lim_{i \rightarrow \infty} v_i = \infty$ ist, so folgt, daß immer, wenn eine Integralkurve $\varphi(x)$ einer analytischen Differentialgleichung $y' = f(x, y)$ von einer Unbestimmtheitsstelle aus in eine Potenzreihe entwickelbar ist, $\varphi(x)$ die Grenzlage der Näherungsparabeln η_m ist und daß die dadurch erhaltene Reihe für $\varphi(x)$ innerhalb des Konvergenzbereiches von P und Q so lange konvergiert, als $\varphi'(x)$ endlich ist.

Wie nun aus einfachen Beispielen hervorgeht, lassen sich aber nicht alle Integralkurven von der singulären Stelle aus in Potenzreihen entwickeln. Ich werde nun zeigen, daß diese Entwicklung nach steigenden Potenzen nur ein Spezialfall einer analytischen Darstellung ist, mit deren Hilfe man alle in die singuläre Stelle einmündenden Integralkurven einer analytischen Differentialgleichung erfassen kann. Es wird sich ergeben, daß man für die Lösungen der Differentialgleichung (12) Reihenentwicklungen angeben kann. Damit ist eine Möglichkeit gegeben, alle Integralkurven darzustellen, mit Ausnahme derjenigen, bei denen das Krümmungsmaß Null oder Unendlich auftritt. Wendet man in diesem Fall die Transformation $y = \sum_{i=1}^{n-1} u_i x^{v_i} + u(x) x^{v_n}$ bzw. $y = \sum_{i=1}^{n-1} u_i x^{v_i} + \frac{1}{u(x)} x^{v_n}$ an (vgl. S. 249/250), so erhält man für $\frac{dx}{du}$ eine Differentialgleichung von der Form (12).

In diesem Fall erhält man die Integralkurve der (x, y) -Ebene in der Parameterform $x = R(u)$, $y = \sum_{i=1}^{n-1} u_i x^{r_i} + u x^{r_n}$ bzw. $\sum_{i=1}^{n-1} u_i x^{r_i} + \frac{1}{u} x^{r_n}$.

Zu diesem Zweck nehme ich zunächst an, daß in der Transformationsgleichung $y = u_1 x^{r_1} + u_2 x^{r_2} + \dots + (u(x) - u_n) x^{r_n}$ alle Exponenten rational seien. Ist dann ρ das größte gemeinschaftliche Maß der Zahlen $(1, r_1, \dots, r_n)$, und setzt man $\bar{x} = x^\rho$, so erhält man in der Differentialgleichung (12) nur ganzzahlige Exponenten. Schreibt man der Einfachheit halber statt \bar{x} wieder x und statt u wieder y , so lautet diese Differentialgleichung:

$$y' = \frac{f_1(y)}{x^m} + \frac{f_2(x, y)}{x^{m-1}} = f(x, y).$$

Dabei ist $f_1(0) = 0$; $f_1'(0) = c \neq 0$; $m \geq 1$. Ohne Beschränkung der Allgemeinheit kann man annehmen, daß $f_1(y) \equiv cy$ ist; denn ist dies bei Zugrundelegung der Differentialgleichung (12) noch nicht der Fall, so tritt es jedenfalls bei der Differentialgleichung (13) ein. Die zu betrachtende Differentialgleichung hat demnach die Form:

$$(14) \quad y' = \frac{cy}{x^m} + \frac{f_2(x, y)}{x^{m-1}} = f(x, y).$$

Betrachtet man nun einen Bereich \mathfrak{B} : $0 \leq x \leq a$; $|y| \leq b$, in welchem $f_2(x, y)$ regulär ist, so genügt darin die partielle Ableitung $\frac{\partial f_2}{\partial y}$ einer Ungleichung: $\left| \frac{\partial f_2}{\partial y} \right| \leq M$. Also ist in \mathfrak{B} : $\frac{c}{x^m} - \frac{M}{x^{m-1}} \leq \frac{\partial f}{\partial y} \leq \frac{c}{x^m} + \frac{M}{x^{m-1}}$. Für die folgende Untersuchung müssen wir a noch durch die Ungleichung $a < \frac{|c|}{2M}$ einschränken.

Um nun eine Integralkurve durch den singulären Punkt zu erhalten, nehme man, wie bei der Methode der sukzessiven Approximation, eine Näherungskurve $\eta_1(x)$, etwa $\eta_1(x) \equiv 0$. Dann wird eine Integralkurve y dargestellt durch $y = \eta_1 + u$; also:

$$\eta_1' + u' = f(x, \eta_1 + u) = f(x, \eta_1) + u \frac{\partial f}{\partial y}(x, \eta_1) + \frac{u^2}{2} \frac{\partial^2 f}{\partial y^2}(x, \eta_1) + \dots$$

Um die zweite Näherungskurve η_2 zu erhalten, könnte man den Zusatz u_1 bestimmen aus der Differentialgleichung:

$$\eta_1' + u_1' = f(x, \eta_1) + u_1 \frac{\partial f}{\partial y}(x, \eta_1).$$

Entwickelt man aber $\frac{\partial f}{\partial y}(x, \eta_1)$ nach Potenzen von x und begnügt sich auch hier mit dem ersten Glied, so erhält man zur Bestimmung von u_1 die Differentialgleichung:

$$(15) \quad u_1' = f(x, \eta_1) - \eta_1' + u_1 \frac{c}{x^m} \cdot^{10)}$$

Da die erste Näherungskurve im Ursprung die x -Achse berührt, enthält der Ausdruck $\eta_1(x)$ sicher den Faktor x . Nach (14) ist dann für $0 \leq x \leq a$ $|f(x, \eta_1) - \eta_1'| \leq \frac{N}{x^{m-1}}$. Man kann demnach die Zahl a so klein wählen, daß sie den obigen Bedingungen genügt, und daß außerdem für $0 \leq x \leq a$ $|f(x, \eta_1) - \eta_1'| < \frac{b|c|}{2x^m}$ ist.

Betrachtet man nun das Paar von Geraden $|u_1| = \frac{b}{2}$, so sieht man aus Gleichung (15), daß auf diesen im Bereich \mathfrak{B} u_1' dasselbe Vorzeichen hat wie $u_1 c$. Nun sind zwei Fälle zu unterscheiden:

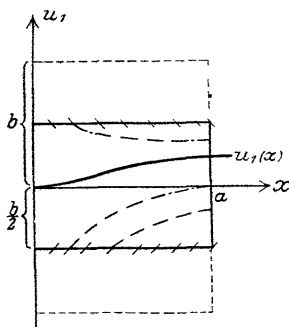


Fig. 16.

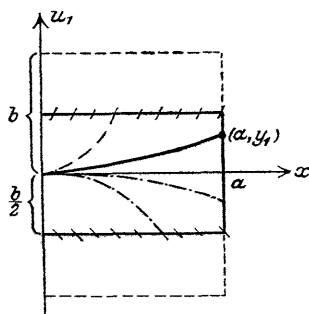


Fig. 17.

1. $c < 0$. Der Bereich $|u_1| \leq \frac{b}{2}$; $0 \leq x \leq a$ besitzt auf seinem Rand zwei Quellpunkte (Fig. 16). Es gibt nur eine Integralkurve $u_1(x)$, die in den Ursprung einmündet. Die Zusatzfunktion $u_1(x)$ ist demnach eindeutig bestimmt. Für sie gilt in \mathfrak{B} die Abschätzung:

$$|u_1(x)| \leq \frac{b}{2}.$$

2. $c > 0$. Der obige Bereich ist frei von Randsingularitäten (Fig. 17). Alle Integralkurven durch den Rand des Bereiches laufen durch den Ursprung. Es gibt demnach unendlich viele Zusatzfunktionen, die man mit η_1 zusammennehmen kann. Dies stimmt überein mit dem Ergebnis der qualitativen Untersuchung, daß unendlich viele Integralkurven in die singuläre Stelle einmünden, wenn in der Umgebung des Ursprungs $\frac{\partial f}{\partial y}$ positiv ist.

¹⁰⁾ Diese Methode hat auch Bendixson zur quantitativen Bestimmung der Integralkurven angewandt.

Um eine Integralkurve eindeutig zu bestimmen, muß eine weitere Randbedingung angegeben werden. Will man z. B. die Integralkurve durch den Punkt (a, y_1) ($|y_1| \leq \frac{b}{2}$), so bestimme man u_1 so, daß die Näherungskurve η_2 durch diesen Punkt hindurchgeht. Alle weiteren Zusatzfunktionen werden dann so bestimmt, daß sie für $x = a$ verschwinden. (Hat aber die zu bestimmende Integralkurve mit der Strecke $x = a$, $|y_1| \leq \frac{b}{2}$ keinen Punkt gemeinsam, so kann man durch Verkleinerung von a immer erreichen, daß sie diese Strecke schneidet.)

Ist nun $\eta_1 \equiv 0$, so ist im Bereich \mathfrak{B} $|\eta_2| \leq \frac{b}{2}$. Die zweite Zusatzfunktion $u_2(x)$ erhält man nun aus der Differentialgleichung

$$(16) \quad u_2' = f(x, \eta_2) - \eta_2' + u_2 \frac{c}{x^m}$$

und der Bedingung $u_2(0) = 0$, wenn $c < 0$,

bzw. $u_2(a) = 0$, wenn $c > 0$ ist.

Nun folgt aus Gleichung (15):

$$\eta_2' = \eta_1' + u_1' = f(x, \eta_1) + \frac{u_1 c}{x^m}.$$

Also:

$$\begin{aligned} f(x, \eta_2) - \eta_2' &= f(x, \eta_2) - f(x, \eta_1) - u_1 \frac{c}{x^m} \\ &= (\eta_2 - \eta_1) \frac{\partial f}{\partial y}(x, \bar{\eta}_1) - u_1 \frac{c}{x^m} = u_1 \left(\frac{\partial f}{\partial y}(x, \bar{\eta}_1) - \frac{c}{x^m} \right) \end{aligned}$$

$$|f(x, \eta_2) - \eta_2'| \leq |u_1| \frac{M}{x^{m-1}} \leq \frac{b}{2} \frac{M}{x^{m-1}} \leq \frac{b}{2} \frac{aM}{x^m} < \frac{b|c|}{4x^m}.$$

Aus Gleichung (16) folgt jetzt, daß für $0 \leq x \leq a$ auf dem Geradenpaar $|u_2| = \frac{b}{4}$ das Vorzeichen von u_2' übereinstimmt mit dem Vorzeichen von $c u_2$. Die Zusatzfunktion $u_2(x)$, die gemäß den obigen Bedingungen bestimmt wird, genügt also der Ungleichung $|u_2(x)| \leq \frac{b}{4}$. Folglich ist

$$|\eta_3(x)| = |\eta_2(x) + u_2(x)| \leq \frac{3b}{4},$$

bleibt also für $0 \leq x \leq a$ im Bereich \mathfrak{B} .

Wird das Verfahren fortgesetzt, so findet man, daß $|u_i| \leq \frac{b}{2^i}$ ist.

Denn ist dies der Fall für alle $i \leq n$, so ist $|\eta_{n+1}| \leq \frac{(2^n - 1)b}{2^n}$. Die

Kurve η_{n+1} verläuft also für $0 \leq x \leq a$ im Bereich \mathfrak{B} . Den Zusatz u_{n+1} erhält man aus der Differentialgleichung

$$u_{n+1}' = f(x, \eta_{n+1}) - \eta_{n+1}' + u_{n+1} \frac{c}{x^m}.$$

Nun ist:

$$\begin{aligned}\eta'_{n+1} &= \eta'_n + u'_n = f(x, \eta_n) + u_n \frac{c}{x^m}, \\ f(x, \eta_{n+1}) - \eta'_{n+1} &= f(x, \eta_{n+1}) - f(x, \eta_n) - u_n \frac{c}{x^m} \\ &= (\eta_{n+1} - \eta_n) \frac{\partial f}{\partial y}(x, \bar{\eta}_n) - u_n \frac{c}{x^m} \\ &= u_n \left(\frac{\partial f}{\partial y}(x, \bar{\eta}_n) - \frac{c}{x^m} \right).\end{aligned}$$

Da nun η_{n+1} und η_n im Bereich \mathfrak{B} verlaufen, verläuft darin auch $\bar{\eta}_n$; also gilt für $\frac{\partial f}{\partial y}(x, \bar{\eta}_n)$ die frühere Abschätzung. Es ist also:

$$|f(x, \eta_{n+1}) - \eta'_{n+1}| \leq |u_n| \frac{M}{x^{m-1}} \leq \frac{b}{2^n} \frac{M}{x^{m-1}} < \frac{b}{2^{n+1}} \frac{|c|}{x^m}.$$

Also stimmt für $0 \leq x \leq a$ auf dem Geradenpaar $|u_{n+1}| = \frac{b}{2^{n+1}}$ das Vorzeichen von u'_{n+1} überein mit dem Vorzeichen von $c u_{n+1}$. Die Zusatzfunktion $u_{n+1}(x)$, die gemäß den vorgeschriebenen Bedingungen bestimmt wird, genügt also der Ungleichung

$$|u_{n+1}(x)| \leq \frac{b}{2^{n+1}}.$$

Daraus folgt, daß für $0 \leq x \leq a$ die Folge $\eta_1, \eta_2, \dots, \eta_i, \dots$ absolut und gleichmäßig gegen eine Grenzkurve η konvergiert. Diese Grenzkurve ist notwendigerweise Integralkurve, denn es ist

$$\lim_{i \rightarrow \infty} u_i = 0;$$

folglich ist für $x > 0$

$$\lim_{i \rightarrow \infty} (f(x, \eta_i) - \eta'_i) = 0, \quad \lim_{i \rightarrow \infty} f(x, \eta_i) = f(x, \eta) = \eta' = \lim_{i \rightarrow \infty} \eta'_i.$$

Außerdem genügt die Funktion $\eta(x)$ nach der Art ihrer Entstehung den vorgeschriebenen Anfangsbedingungen.

Es wird also hier eine Integralkurve einer analytischen Differentialgleichung mit einer Unbestimmtheitsstelle im Ursprung nach Lösungen linearer Differentialgleichungen entwickelt, für die der Ursprung ebenfalls eine Unbestimmtheitsstelle ist, aber weiterhin im allgemeinen auch eine wesentlich singuläre Stelle für einen der Koeffizienten. Das i -te Glied genügt der Differentialgleichung

$$u'_i = f(x, \eta_i) - \eta'_i + u_i \frac{c}{x^m},$$

also ist

$$u_i = e^{\int_1^x \frac{c}{\xi^m} d\xi} \int_1^x e^{-\int_1^{\xi} \frac{c}{\xi^m} d\xi} [f(x, \eta_i) - \eta'_i] dx.$$

Dabei ist bei dem Hauptintegral die obere Grenze x ; die untere Grenze ist für alle i gleich Null, wenn $c < 0$ ist. Sie ist für alle $i > 1$ gleich a , wenn $c > 0$ ist; für $i = 1$ muß sie so gewählt werden, daß die Kurve η_2 durch den Punkt (a, y_1) geht, dessen Integralkurve man bestimmen will. Demnach ist

$$\eta = e^1 \int_{\frac{c}{\xi^m}}^x d\xi \int_0^x e^{-\int_{\frac{c}{\xi^m}}^x d\xi} \Sigma(f(x, \eta_i) - \eta'_i) dx, \quad \text{wenn } c < 0 \text{ ist,}$$

und

$$\eta = \eta_2 + e^1 \int_{\frac{c}{\xi^m}}^x d\xi \int_a^x e^{-\int_{\frac{c}{\xi^m}}^x d\xi} \Sigma(f(x, \eta_i) - \eta'_i) dx \quad \text{für } c > 0.$$

Hieraus sieht man ohne weiteres, daß im Spezialfall $m = 1$ und $c < 0$ alle Näherungskurven η_i und damit auch η durch Potenzreihen darstellbar sind.

Bis jetzt wurde vorausgesetzt, daß die Exponenten $\nu_1, \nu_2, \dots, \nu_n$ rational seien. Ich werde nun durch die Methode der sukzessiven Approximation nachweisen, daß jede Integralkurve $\varphi(x)$, die sich an eine Näherungsparabel $y = \sum_{i=1}^n u_i x^{\nu_i}$ mit irrationalem ν_n anschmiegt, als Reihe nach Potenzen von x^ρ und x^{ν_n} (ρ gemeinschaftliches Maß von $1, \nu_1, \dots, \nu_{n-1}$) darstellbar ist.

Die Existenz dieser Integralkurve ist in § 8 nachgewiesen. Wir wissen ferner, daß man von der zugehörigen Näherungsparabel beliebig viele Glieder bestimmen kann, und daß in ihrer Umgebung für $0 \leq x \leq a$ die partielle Ableitung $\frac{\partial f}{\partial y}$ der Ungleichung $\left| \frac{\partial f}{\partial y} \right| \leq \frac{M}{x}$ genügt. Man kann also eine Näherungsparabel η_0 finden, so daß für $0 \leq x \leq \delta \leq a$

$$|\varphi(x) - \eta_0| \leq x^{2M}$$

ist. Setzt man dann

$$\eta'_1 = f(x, \eta_0),$$

so ist

$$\varphi' - \eta'_1 = f(x, \varphi) - f(x, \eta_0) = (\varphi - \eta_0) \frac{\partial f}{\partial y}(x, \bar{y}_0),$$

$$|\varphi' - \eta'_1| \leq x^{2M} \frac{M}{x} = M x^{2M-1},$$

$$|\varphi - \eta_1| \leq \frac{x^{2M}}{2}.$$

Daraus folgt

$$|\varphi - \eta_i| \leq \frac{x^{2M}}{2^i}.$$

Also

$$\lim_{i \rightarrow \infty} (\varphi - \eta_i) = 0,$$

$$\varphi = \lim_{i \rightarrow \infty} \eta_i.$$

Wird also bei der Bestimmung der Näherungsparabeln ein Exponent aus den Koeffizienten der ursprünglichen Differentialgleichung berechnet, so sind die entsprechenden Integralkurven in Reihen nach Potenzen von x^ρ und x^λ (ρ rational, λ irrational) entwickelbar. Der einfachste Fall dieser Art ist die Differentialgleichung (14) für $m = 1$. Es ist also:

$$(17) \quad y' = \frac{cy}{x} + f_2(x, y).$$

Ist $c < 0$, so kann die Schmiegungsparabel der die x -Achse berührenden Integralkurve beliebig genau bestimmt werden. Die dabei auftretenden Exponenten sind ganzzahlig. Die vorhergehende Überlegung zeigt, daß im vorliegenden Fall diese Integralkurve als Potenzreihe in x dargestellt werden kann (siehe auch S. 263).

Ist $c > 0$ und nicht ganzzahlig, so können die Näherungsparabeln ebenfalls beliebig genau bestimmt werden. Dabei wird man notwendigerweise zum Sondertypus geführt, bei dem c als Exponent auftritt. In diesem Fall sind also die Integralkurven in Reihen nach Potenzen von x und x^c entwickelbar.

Ist $c > 0$ und ganzzahlig, so ist es möglich, daß die Näherungsparabeln beliebig genau bestimmt werden können. Die Integralkurven sind dann als Potenzreihen in x darstellbar ($y' = \frac{2y+x^2}{x}$; $y = cx^2 + x^3$). Bei der Bestimmung der Näherungsparabeln kann man aber auch auf den Fall unendlichen Krümmungsmaßes geführt werden. In diesem Fall sind die Integralkurven in der Parameterform darzustellen ($y' = \frac{2y+x^2}{x}$; $x = ce^{\frac{1}{u}}$, $y = \frac{1}{u}x^2$).

Aus Beispielen ist ersichtlich, daß es auch im Falle $m > 1$ vorkommen kann, daß einzelne Integralkurven durch Potenzreihen darstellbar sind; dies sind jedoch nur Spezialfälle der in diesem Abschnitt hergeleiteten Darstellungsmöglichkeit.

Daß es vorkommen kann, daß eine nach der Methode von § 8 bestimmte Potenzreihe keine Integralkurve darstellt, zeigt das Beispiel $y' = \frac{y-x^2}{x^2}$. Für alle Integralkurven mit Ausnahme der y -Achse erhält man als n -te Näherungskurve die Parabel $y = \sum_{i=2}^n (i-1)! x^i$.

Die Voraussetzung, daß $P(x, y)$ und $Q(x, y)$ Potenzreihen sind, wurde nur zur Bestimmung der Exponenten der Potenzreihen gebraucht. Zum Konvergenzbeweis für das hier angegebene Approximationsverfahren genügt es, vorauszusetzen, daß die Funktion $f_2(x, y)$ die Lipschitzsche Bedingung

$$|f_2(x, y_2) - f_2(x, y_1)| \leq (y_2 - y_1) M$$

befriedigt.

§ 10.

Beispiele ¹¹⁾.

1. Beispiel.

$$y' = \frac{y + \Theta(x, y)}{x + y + H(x, y)};$$

$$\lim_{\substack{x \rightarrow 0 \\ y \rightarrow 0}} \frac{\Theta(x, y)}{|x|^{r_0} + |y|^{r_0}} = \lim_{\substack{x \rightarrow 0 \\ y \rightarrow 0}} \frac{H(x, y)}{|x|^{r_0} + |y|^{r_0}} = 0; \quad r_0 > 1.$$

Charakteristische Gleichung $y^2 = 0$; der Bereich OAB ($|\frac{y}{x}| \leq \varepsilon, 0 \leq x \leq \delta$) hat im Punkte A einen Quellpunkt.

$$y = x^\nu(x); \quad \nu' = \frac{x^\nu(1-\nu) - \nu x^{2\nu-1} + \Theta - \nu x^{\nu-1} H}{x^\nu \lg x (x + x^\nu + H)}.$$

Mögliche Krümmungsordnung für $1 \leq \nu \leq r_0$ nur $\nu = 1$. Also

$$y = ux; \quad u' = \frac{-u^2 + \frac{\Theta}{x} - \frac{uH}{x}}{x + ux + H}.$$

Mögliches Krümmungsmaß: $u = 0$ (wie auch auf Seite 244 für diesen Fall gefolgert wurde).

Nun begrenzen die Geraden $y = -\varepsilon x$, $x = \delta$ und die Parabel $y = -x^{r_0}$ einen Bereich ohne Randsingularitäten. Also gibt es unendlich viele Integralkurven von der Krümmungsordnung 1 und dem Krümmungsmaß Null.

Um die Integralkurven zu erhalten, deren Krümmungsordnung größer als r_0 ist, wende man die Transformation $y = u(x)x^{r_0}$ an. Dadurch erhält man

$$u' = \frac{u(1-r_0) - r_0 u^2 x^{r_0-1} + \frac{\Theta}{x^{r_0}} - r_0 u \frac{H}{x}}{x + u x^{r_0} + H}.$$

Nach § 5 gibt es nur eine Integralkurve, die im Bereich $|u| \leq 1$ in die singuläre Stelle $(0, 0)$ einmündet. Also mündet in der (x, y) -Ebene im

¹¹⁾ In den folgenden Beispielen beschränken wir uns auf die qualitative Bestimmung der Integralkurven. Ihre analytische Darstellung ist meist kompliziert und unübersichtlich.

Bereich $|y| \leq x^{\tau_0}$ (Fig. 18) nur eine Integralkurve in den Ursprung ein. Wenn über Θ und H nicht mehr vorausgesetzt wird, kann man nur aussagen, daß die Krümmungsordnung dieser Integralkurve größer als τ_0 ist. Sind aber Θ und H Potenzreihen, so läßt sich dieses Integral am Ursprung in eine Potenzreihe nach x entwickeln.

2. Beispiel.

$$y' = \frac{y(y-x^2)(y+x^2)}{-x^6} = \frac{y^3 - x^4 y}{-x^6}.$$

Charakteristische Gleichung $x y^3 = 0$. Die y -Achse ist einfache, reguläre ausgezeichnete Richtung, längs der nur eine Integralcurve in den Ursprung einmündet (§ 5)¹²⁾. Die x -Achse ist dreifache, reguläre ausgezeichnete Richtung. Der Bereich OAB ($x > 0$) hat zwei Quellpunkte (Fig. 19), der Bereich $OA'B'$ ($x < 0$) ist frei von Randsingularitäten.

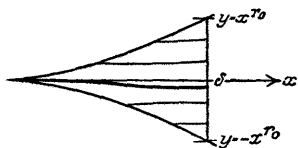


Fig. 18.

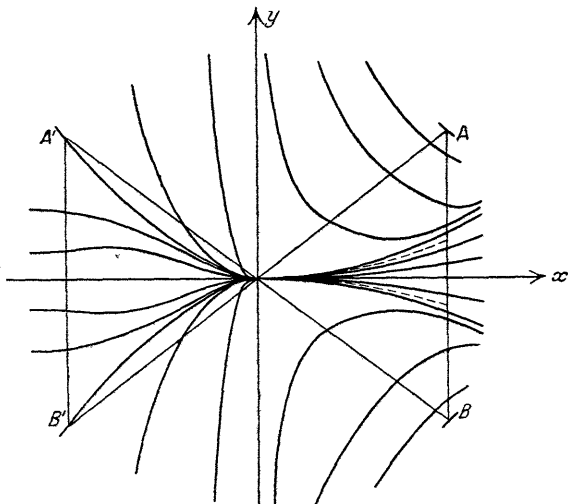


Fig. 19.

$$y = x^{\nu}; \quad \nu' = \frac{x^{2\nu} - x^4 + \nu x^5}{-x^6 \lg x}.$$

Mögliche Krümmungsordnung $\nu = 2$.

$$y = u(x)x^2; \quad u' = \frac{u(u^2 - 1) + 2ux}{-x^2}.$$

Mögliche Krümmungsmaße $u = 0, \pm 1$. Nach § 5 gehen durch den Punkt $(0, 0)$ der (x, u) -Ebene unendlich viele Integralkurven, durch die Punkte $(0, \pm 1)$ dagegen außer der u -Achse je nur eine. Also sind die Parabeln $y = \pm x^2$ Schmiegeparabel je einer Integralkurve, während sich an die x -Achse unendlich viele anschmiegen, deren Krümmungsordnung unendlich ist. Für $x < 0$ schmiegen sich dagegen je unendlich viele Integralkurven an die Parabeln an.

¹²⁾ In Fig. 19 und einigen der folgenden Figuren ist versehentlich die y -Achse nicht als Integralkurve eingezeichnet worden.

3. Beispiel.

$$y' = \frac{y^2 + ax^6}{x^4}.$$

Charakteristische Gleichung $xy^2 = 0$. Längs der y -Achse mündet nach § 5 nur eine Integralkurve in die singuläre Stelle ein. Die x -Achse ist zweifache, reguläre ausgezeichnete Richtung, die sowohl für $x > 0$ als auch $x < 0$ von einem Bereich mit einem Quellpunkt umgeben wird. Durch $\xi = -x$, $\eta = -y$ geht die Differentialgleichung in sich über. Die Integralkurven sind demnach zentrisch-symmetrisch. Es genügt also, ihre Gestalt für $x > 0$ zu untersuchen.

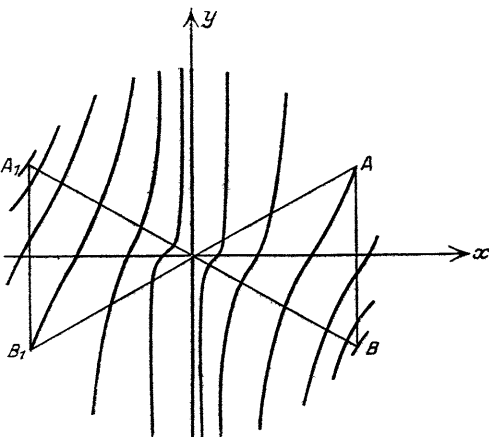


Fig. 20.

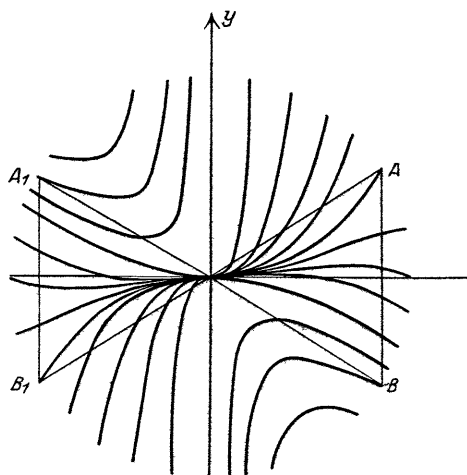


Fig. 21.

$$y = x^r; \quad r' = \frac{x^{2r} + ax^6 - rx^{3+r}}{x^{4+r} \lg x}.$$

Mögliche Krümmungsordnung $r = 3$.

$$y = ux^3; \quad u' = \frac{u^2 - 3u + a}{x}.$$

Nun sind drei Fälle zu unterscheiden:

a) $a > \frac{9}{4}$; $u^2 - 3u + a$ ist definit.

Es gibt keine Schmiegungsparabeln, folglich keine Integralkurven, die aus diesen Bereichen in den Ursprung einmünden (Fig. 20).

b) $a < \frac{9}{4}$; $\psi(u, 0) \equiv u^2 - 3u + a$ hat zwei Nullstellen u_1 und u_2 .

Es sei dabei $\frac{\partial \psi}{\partial u}(u_1, 0) > 0$; dann ist $\frac{\partial \psi}{\partial u}(u_2, 0) < 0$. Also schmiegen sich unendlich viele Integralkurven an die Parabel $y = u_1 x^3$ an und nur eine an die Parabel $y = u_2 x^3$ (Fig. 21).

$$c) \quad a = \frac{9}{4}; \quad u^2 - 3u + \frac{9}{4} = \left(u - \frac{3}{2}\right)^2.$$

Die Parabel $y = \frac{3}{2}x^3$ entspricht einer doppelten ausgezeichneten Richtung; also wende man die Transformation $y = \frac{3}{2}x^3 + x^{\nu(x)}$ an; dadurch erhält man

$$\nu' = \frac{(3 - \nu) + x^{\nu-3}}{x \lg x}.$$

Mögliche Krümmungsordnung nur $\nu = 3$. Daher

$$y = \left(\frac{3}{2} + u\right)x^3; \quad u' = \frac{u^2}{x};$$

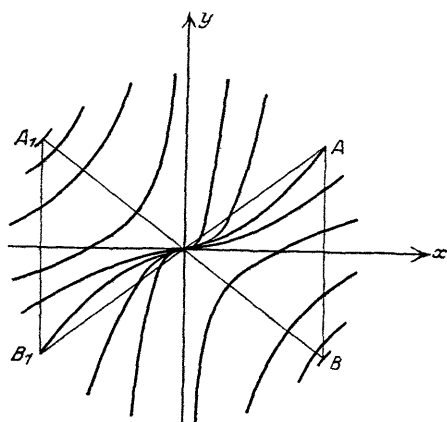


Fig. 22.

mögliches Krümmungsmaß nur $u = 0$. Nun begrenzen die Parabeln $y = \left(\frac{3}{2} + \delta\right)x^3$ und $y = \frac{3}{2}x^3 + x^4$ mit der Geraden $x = \delta_1$ einen Bereich ohne Randsingularitäten. Also gibt es unendlich viele Integralkurven, die sich an die Parabel $y = \frac{3}{2}x^3$ anschmiegen (Fig. 22). (Da diese Parabel selbst Lösung der Differentialgleichung ist, handelt es sich hier um eine integrierbare Riccatische Gleichung.)

4. Beispiel.

$$y' = \frac{x^2}{y}.$$

Charakteristische Gleichung $y^2 = 0$. Die x -Achse ist singuläre ausgezeichnete Richtung. Die sie umgebenden Bereiche zerfallen durch $y = 0$ für $x > 0$ in zwei Bereiche mit je zwei Quellpunkten, für $x < 0$ in zwei Bereiche mit je einem Quellpunkt.

$$y = x^\nu; \quad \nu' = \frac{x^2 - \nu x^{2\nu-1}}{x^{2\nu} \lg x}.$$

Mögliche Krümmungsordnung $\nu = \frac{3}{2}$.

$$y = ux^{\frac{3}{2}}; \quad u' = \frac{1 - \frac{3}{2}u^2}{ux} \quad (x > 0).$$

Mögliche Krümmungsmaße $u = \pm \sqrt{\frac{2}{3}}$.

Nun ist

$$\frac{\partial \psi}{\partial u} \left(\pm \sqrt{\frac{2}{3}}, 0 \right) = -3;$$

also mündet längs jeder dieser Parabeln eine Integralkurve in die singuläre Stelle ein.

Für $x < 0$ setze man $\xi = -x$; $\eta = y$; dann ist

$$\eta' = -\frac{\eta^2}{\xi}; \quad \eta = u \xi^{\frac{3}{2}}; \quad u' = \frac{-1 - \frac{3}{2} u^2}{u \xi}.$$

Es gibt also keine Schmiegungsparabeln für $x < 0$, also auch keine Integralkurven, die aus diesem Bereich in die singuläre Stelle einmünden.

5. Beispiel.

$$y' = -\frac{x^3}{y}.$$

Charakteristische Gleichung $y^2 = 0$. Die x -Achse ist singuläre, ausgezeichnete Richtung. Die sie umgebenden Bereiche zerfallen in je zwei Bereiche mit einem Quellpunkt.

$$y = x^{\nu}; \quad \nu' = \frac{-x^3 - \nu x^{2\nu-1}}{x^{2\nu} \lg x}.$$

Mögliche Krümmungsordnung $\nu = 2$.

$$y = u x^2; \quad u' = \frac{-1 - 2u^2}{u x}.$$

Da der Zähler definit ist, gibt es für $x > 0$ keine Schmiegungsparabeln; also gibt es auch keine Integralkurven, die für $x > 0$ in den Ursprung einmünden. Für $x < 0$ verhält es sich ebenso, da durch $\xi = -x$, $\eta = y$ die Gleichung in sich übergeht. *Das Beispiel zeigt, daß die Figur des Wirbels oder des Strudels nicht ein charakteristisches Merkmal für den definiten Typus ist, sondern daß sie auch beim nicht-definiten Typus auftreten kann.*

6. Beispiel.

$$y' = \frac{-(y-x^2)(y+x^2)(2y-x^2)(2y+x^2)}{xy} = \frac{-4y^4 + 5y^2x^4 - x^8}{xy}.$$

Charakteristische Gleichung $x y^2 = 0$. Nach § 5 mündet längs der y -Achse nur eine Integralkurve in die singuläre Stelle ein. Die x -Achse ist singuläre ausgezeichnete Richtung; sie wird umgeben von Bereichen, die durch $y = 0$ in je zwei Teilbereiche mit einem Quellpunkt zerfallen (Fig. 23).

$$y = x^{\nu}; \quad \nu' = \frac{-4x^{4\nu} + 5x^{4+2\nu} - x^8 - \nu x^{2\nu}}{x^{2\nu+1} \lg x}.$$

Mögliche Krümmungsordnung $\nu = 4$.

$$y = u x^4; \quad u' = \frac{-1 - 4u^2 + 5u^2 x^4 - 4u^4 x^8}{u x}.$$

Da $(-1 - 4u^2)$ definit ist, gibt es für $x > 0$ keine Schmiegungsparabeln. Für $x < 0$ verhält es sich ebenso. Also münden längs der x -Achse keine Integralkurven in die singuläre Stelle ein.

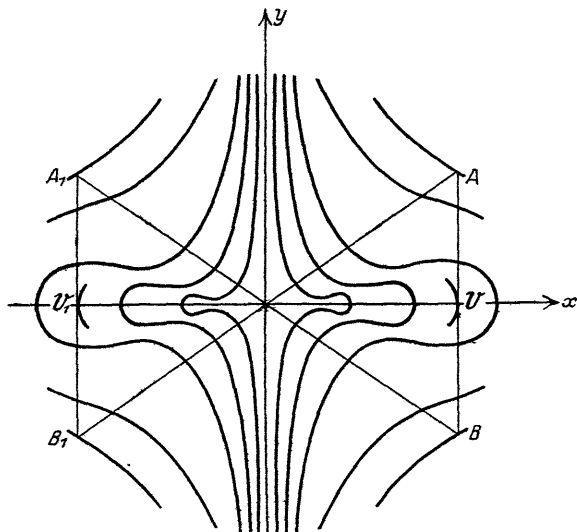


Fig. 23.

7. Beispiel.

$$y' = \frac{(y - x^2 + x^3)(y - x^2 - x^3)}{(y - x^2)(y - x^2 + 2x^3)(y - x^2 - 2x^3)} = \frac{y^2 - 2x^2y + x^4 - x^6}{y^3 - 3x^2y^2 + 3x^4y - x^6 - 4x^6y + 4x^8}.$$

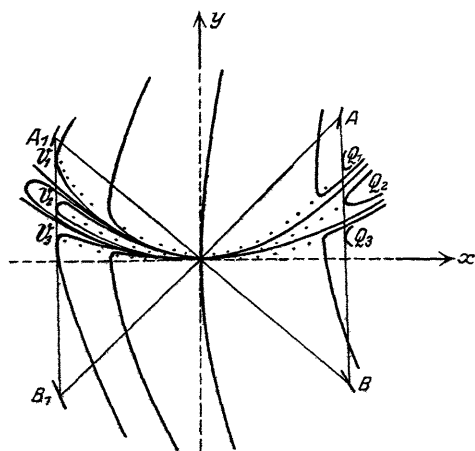


Fig. 24.

Charakteristische Gleichung $xy^2 = 0$. Längs der y -Achse mündet nach § 5 nur eine Integralkurve in die singuläre Stelle ein. Die x -Achse ist singuläre ausgezeichnete Richtung. Die sie umgebenden Bereiche zerfallen für $x > 0$ in zwei Bereiche mit je zwei Quellpunkten und zwei Bereiche mit je einem Quellpunkt; für $x < 0$ in zwei Bereiche ohne Randsingularitäten und zwei Bereiche mit je einem Quellpunkt (Fig. 24).

$$y = x^{\nu}; \quad \nu' = \frac{\sum a_{ik} x^{i+k\nu}}{\lg x \sum b_{ik} x^{i+k\nu}}.$$

Im Zähler kommen als Exponenten, die am kleinsten werden können, nur in Betracht die Exponenten 2ν , $2 + \nu$ und 4 . Sie sind einander gleich für $\nu = 2$; also ist $\nu = 2$ die einzige mögliche Krümmungsordnung.

$$y = u x^2; \quad u' = \frac{u^2 - 2u + 1 + x R_1(x, u)}{x^4 (u^3 - 3u^2 + 3u - 1) - x^6 (4u - 4)}.$$

Mögliches Krümmungsmaß also nur $u = 1$. Nun entspricht aber die Parabel $y = x^2$ noch einer singulären ausgezeichneten Richtung. Also wende man die nächst höhere Transformation an:

$$y = x^2 + x^{\nu}; \quad \nu' = \frac{(x^{\nu} + x^2)(x^{\nu} - x^2)}{x^{2\nu}(x^{\nu} + 2x^2)(x^{\nu} - 2x^2) \lg x} - \frac{2x + \nu x^{\nu-1}}{x^{\nu} \lg x}.$$

Mögliche Krümmungsordnung $\nu = 3$.

$$y = x^2 + u x^3; \quad u' = \frac{(u+1)(u-1)}{u x^6 (u+2)(u-2)} - \frac{2+3u x}{x^2}.$$

Mögliches Krümmungsmaß $u = \pm 1$. Als Näherungsparabeln erhält man also $y_1 = x^2 + x^3$ und $y_2 = x^2 - x^3$. Diese Parabeln entsprechen einfachen, regulären ausgezeichneten Richtungen. Nach § 5 mündet für $x > 0$ längs jeder Parabel eine Integralkurve in die singuläre Stelle ein, dagegen für $x < 0$ längs jeder Parabel unendlich viele.

8. Beispiel.

$$y' = \frac{\sqrt{2} y^2 + x^5}{x y}.$$

Charakteristische Gleichung $(\sqrt{2} - 1) x y^2 = 0$. Längs der y -Achse mündet nach § 5 nur eine Integralkurve in die singuläre Stelle ein. Die x -Achse ist singuläre ausgezeichnete Richtung. Die sie umgebenden Bereiche zerfallen für $x > 0$ in zwei Bereiche mit je einem Quellpunkt, für $x < 0$ in zwei Bereiche ohne Randsingularitäten (Fig. 25).

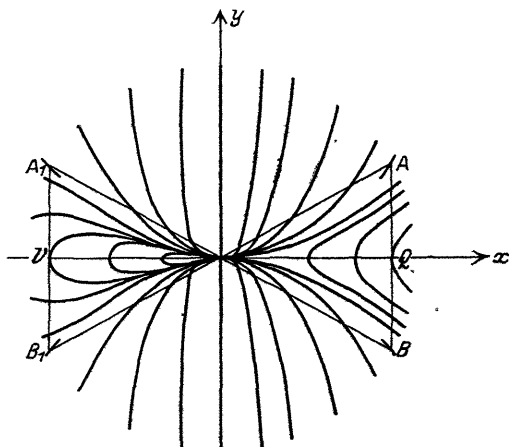


Fig. 25.

$$y = x^{\nu}; \quad \nu' = \frac{(\sqrt{2} - \nu) x^{2\nu} + x^5}{x^{2\nu+1} \lg x}.$$

Mögliche Krümmungsordnungen $\nu = \frac{5}{2}$ und $\nu = \sqrt{2}$, weil $1 < \sqrt{2} < \frac{5}{2}$ ist.

$$y = u x^{\frac{5}{2}}; \quad u' = \frac{\left(\sqrt{2} - \frac{5}{2}\right)u^2 + 1}{u x}.$$

Daher Krümmungsmaß

$$u = \pm \frac{1}{\sqrt{\frac{5}{2} - \sqrt{2}}};$$

da aber die Parabeln

$$y = \pm \frac{1}{\sqrt{\frac{5}{2} - \sqrt{2}}} x^{\frac{5}{2}}$$

selbst Integralkurven sind, sind es nach § 5 die einzigen Integralkurven von der Krümmungsordnung $\frac{5}{2}$.

$$y = u x^{\sqrt{2}}; \quad u' = \frac{x^5}{u x^{2\sqrt{2}} + 1} = \frac{x^{(4-2\sqrt{2})}}{u}.$$

Diese Differentialgleichung ist in jedem Punkt $(0, u \neq 0)$ regulär. Also gehört zu jedem u eine Integralkurve. Es gibt also Integralkurven von der Krümmungsordnung $\sqrt{2}$, und zwar so, daß zu jedem von Null verschiedenen Krümmungsmaß u eine und nur eine Kurve gehört (vgl. S. 249). In diesem Fall sind sämtliche Integrale am Nullpunkt in Potenzreihen entwickelbar.

(Eingegangen am 19. 12. 26.)

Die parabolische Kurve.

Beitrag zur Geometrie der Berührungstransformationen, der partiellen Differentialgleichungen zweiter Ordnung und der Flächenverbiegung.

Von

Stefan Cohn-Vossen in Göttingen.

Überblick.

Unterwirft man eine stetig gekrümmte Fläche (im reellen Euklidischen dreidimensionalen Raum) einer Legendre-Transformation, so ist das Bild ebenfalls stetig gekrümmt dort, wo die Gaußsche Krümmung des Originals nicht verschwindet; dagegen geben parabolische Punkte des Originals singuläre Bildpunkte. Etwas Analoges erscheint bei sämtlichen Berührungstransformationen, soweit sie nicht bloß erweiterte Punkttransformationen sind.

In vorliegender Arbeit soll im ersten Teil der geometrische Grund dieser Singularitäten und einiges über ihr Aussehen ermittelt und in Formeln gebracht werden.

Diese Formeln werden gestatten, umgekehrt gewisse Singularitäten auf Flächen, die bei der Integration partieller Differentialgleichungen zweiter Ordnung durch besondere Anfangsstreifen auftreten müssen, mittels passender Berührungstransformationen zu regularisieren. Hierdurch wird der Cauchy-Kowalewskasche Existenzsatz und der Integralbegriff überhaupt ähnlich erweitert, wie man in der Funktionentheorie durch Regularisieren der Pole vom Begriff der regulären Funktion zu dem der meromorphen fortschreitet. Diese Erörterung ist der zweite Teil der Arbeit.

Im dritten Teil wird die Regularisierungsmethode benutzt, um etwas über die singulären Ränder auszusagen, die bei der Verbiegung krummer Flächen in der Regel auftreten (man denke zum Beispiel an die Rückkehrkanten der abwickelbaren Flächen und die Ränder und Kegelspitzen der explizit bekannten Flächen konstanter Gaußscher Krümmung). Es ergeben sich sowohl über das Auftreten als auch über die Gestalt dieser Ränder einige anschauliche neue Sätze.

Um den Gedankengang nicht durch Nebenbetrachtungen aufzuhalten, sind die vorkommenden geometrischen und analytischen Gebilde immer als regulär analytisch angenommen, soweit nichts anderes bemerkt ist. In Wirklichkeit gelten die aufzustellenden Sätze auch unter schwächeren Voraussetzungen. Dies wird aber im Verlauf der Arbeit nicht mehr jedesmal erwähnt.

Die Formeln sind meist zu Systemen zusammengefaßt, die in den zwei ersten Teilen mit (A), (B), . . ., im dritten Teil mit (I), (II), . . ., bezeichnet sind; die einzelnen Gleichungen der Systeme sind 1., 2., . . . numeriert.

In der Literatur scheint die hier behandelte Frage noch nicht gestellt worden zu sein. Rückkehrkanten spezieller Flächen sind unter ganz anderm Gesichtspunkt schon von Monge¹⁾ eingehend untersucht worden.

Erster Teil.

§ 1.

Auf einer stetig gekrümmten Fläche im dreidimensionalen Euklidischen Raum heißen bekanntlich diejenigen Punkte parabolisch, in denen die Gaußsche Krümmung der Fläche verschwindet. Wenn die Fläche Stellen sowohl positiver als auch negativer Gaußscher Krümmung hat, bilden die parabolischen Punkte der Fläche gewöhnlich Kurvenzüge, die *parabolische Kurve* der Fläche.

Zum Beispiel betrachte man die ringförmige Drehfläche eines Kreises um eine in seiner Ebene liegende Gerade, die den Kreis nicht schneidet. Diese Fläche hat einen Bereich positiver und einen negativer Gaußscher Krümmung. Die Bereiche werden getrennt durch die beiden Breitenkreise, die der höchste und der tiefste Punkt bei der Drehung beschreiben (wenn die Achse als vertikal angenommen wird). Diese beiden Breitenkreise sind die parabolische Kurve der Ringfläche. Die Fläche hat also längs beiden parabolischen Kurvenzügen je eine und dieselbe Tangentialebene; oder nach Lie: Die parabolischen Streifen der Fläche sind eben.

Es soll nun gezeigt werden, daß diese Erscheinung weniger speziell ist, als es zunächst scheinen mag. Es gilt nämlich der Satz:

Ein parabolischer Streifen einer Fläche ist eben, wenn seine Kurve Enveloppe der Asymptotenlinien der Fläche ist.

Beweis. Jede Asymptotenlinie einer Fläche genügt bekanntlich der Enneperschen Gleichung

$$\frac{1}{\tau} = \sqrt{-K} \quad \left(\frac{1}{\tau} = \text{Torsion} \right),$$

¹⁾ Application de l'analyse à la géométrie. 5. Aufl. 1850 von Liouville.

und nach ihrer Definition ist ihre Schmiegeebene Tangentialebene der Fläche. Wie man sich leicht überzeugt, folgen beide Eigenschaften ausschließlich daraus, daß die zweite Fundamentalförm²⁾ der Fläche in der Asymptotenrichtung verschwindet. Daher gelten beide Eigenschaften auch für die parabolische Kurve, die als Enveloppe von Asymptotenlinien stets Asymptotenrichtung hat. Auf ihr ist aber definitionsgemäß $K = 0$; also $\frac{1}{\tau} = 0$, d. h. ihre Schmiegeebene ist fest. Diese ist aber gleichzeitig Tangentialebene der Fläche, womit alles bewiesen ist.

Wenn das Netz der Asymptotenlinien einer Fläche andererseits überhaupt eine Enveloppe hat, so ist diese eine parabolische Kurve der Fläche. Denn die Asymptotenlinien sind durch die Gleichung $Ldu^2 + 2Mdudv + Ndv^2 = 0$ definiert, die etwaige Enveloppe erfüllt also $LN - M^2 = 0$, die Gleichung der parabolischen Kurve.

Trivial ist die Umkehrung des eben Bewiesenen: Jeder ebene Streifen einer Fläche ist parabolisch, und seine Kurve ist Asymptotenlinie der Fläche.

§ 2.

Der eben bewiesene Satz ist ein Einzelfall eines allgemeineren, der zur Theorie der Flächenkomplexe und Berührungstransformationen überleitet.

Sei irgendein hinreichend stetiges System von ∞^3 Flächen, ein *Flächenkomplex* (S) gegeben, ferner irgendeine nicht zu (S) gehörige Fläche F . Dann gibt es, wie aus einer einfachen Konstantenabzählung folgt, zu jedem Punkt P auf F genau eine Komplexfläche S_0 , die F in P berührt; die leicht übersehbaren Ausnahmefälle, die durch das Verschwinden gewisser Funktionaldeterminanten bedingt werden, braucht man für das folgende nicht zu berücksichtigen.

Beim Übergang von P zu einem Nachbarpunkt P' wird die zugehörige Komplexfläche S_0 in eine Nachbarfläche S'_0 übergehen, die gewöhnlich von S_0 verschieden ist. Es kann aber sein, daß es eine Richtung d durch P gibt, so daß S_0 auch in dem zu d gehörenden Nachbarpunkt P'_d noch F berührt. In diesem Falle nennen wir im folgenden P einen *parabolischen Punkt von F bezüglich (S)*, und nennen d die *Oskulationsrichtung in P* .³⁾ Die Verbindungslinie der bezüglich (S) parabolischen Punkte von F nennen wir die *parabolische Kurve von F bezüglich (S)*.

²⁾ Bezeichnungen wie bei Blaschke: Vorlesungen über Differentialgeometrie I, Berlin 1921.

³⁾ Hierbei und im folgenden bleibt der Fall unerörtert, daß d unbestimmt ist, daß also S_0 in mehr als einer Richtung F in P oskuliert.

Die gewöhnliche parabolische Kurve ist in dieser Definition enthalten; sie bezieht sich auf den Komplex aller Ebenen. In der Tat sind parabolisch im gewöhnlichen Sinne diejenigen Punkte, deren Tangentialebene die Fläche noch in einem Nachbarpunkte berührt. d ist dann Hauptkrümmungs- und Haupttangentialrichtung; bekanntlich fallen beide Richtungen in parabolischen Punkten zusammen.

Der Satz aus § 1 legt es daher nahe, auch bei den allgemeinen parabolischen Kurven zu unterscheiden, ob sie in jedem Punkt die zugehörige Oskulationsrichtung haben oder nicht.

In der Tat: Im zweiten Fall wird die oskulierende Komplexfläche beim Fortschreiten längs der parabolischen Kurve sich ändern. Im ersten Fall aber folgt unmittelbar aus der Definition der Oskulationsrichtung, daß die Fläche F längs der ganzen parabolischen Kurve von einer und derselben Komplexfläche S_0 berührt wird. Anders ausgedrückt: der parabolische Streifen gehört ganz zu (S) .

Offenbar gibt § 1 einen Sonderfall dieses Satzes. Als neues Beispiel wählen wir für (S) den Komplex aller Kugeln vom Radius a . Die parabolischen Punkte bezüglich dieses Komplexes sind die, in denen eine der Hauptkrümmungen gleich $\frac{1}{a}$ ist. d ist die zur Krümmung $\frac{1}{a}$ gehörige Hauptkrümmungsrichtung. Es gilt also der (aus der Theorie der Kanalfächen bekannte) Satz: Ist längs einer Krümmungslinie die zugehörige Hauptkrümmung fest gleich $\frac{1}{a}$, so liegt der Krümmungsstreifen auf einer Kugel vom Radius a .

§ 3.

Irgendwelche hinreichend unabhängige dreiparametrische Darstellung des Komplexes (S) definiert — die Parameter als Punktkoordinaten eines neuen Raumes gedeutet — eine Abbildung der Komplexflächen auf Punkte. Bekanntlich wird durch jede solche Abbildung eine *Berührungstransformation* erzeugt, nach folgendem Lie'schen Verfahren: x, y, z seien kartesische Koordinaten des Ausgangsraumes, den wir im folgenden als „ x -Raum“ bezeichnen; ξ, η, ζ seien geeignete Parameter von (S) . (S) werde durch die Gleichung bestimmt:

$$f(x, y, z; \xi, \eta, \zeta) = 0.$$

Bezeichnet man dann in der üblichen Weise ein Flächenelement des x -Raumes durch $x, y, z; p, q$, und faßt man analog $\xi, \eta, \zeta; \pi, \kappa$ als Flächenelement eines „ ξ -Raumes“ auf, in dem ξ, η, ζ kartesische Punktkoordinaten sind, so transformieren sich die Elemente durch die Formeln:

$$(A) \quad \begin{array}{l} 1. f(x, y, z; \xi, \eta, \zeta) = 0, \\ 2. f_x + p f_z = 0, \\ 3. f_y + q f_z = 0, \\ 4. f_\xi + \alpha f_\zeta = 0, \\ 5. f_\eta + \kappa f_\zeta = 0. \end{array} \quad \left(f_x = \frac{\partial f(x, y, z; \xi, \eta, \zeta)}{\partial x}, \text{ usw.} \right)$$

Mittels dieser Transformation geht die Fläche F des x -Raumes in eine Fläche Φ des ξ -Raumes über. *Es zeigt sich nun, daß die bezüglich (S) parabolische Kurve l von F in eine singuläre Kurve λ von Φ übergeht; singulär insofern, als die zweiten Ableitungen von Φ bei Annäherung an λ unendlich werden, wenn die Ausgangsfläche und die Berührungstransformation selbst in der Umgebung von l sich regulär verhält; diese Annahme wird im folgenden stets gemacht.*

Wir erörtern nun — vorläufig beweislos — das Zustandekommen jener Singularität; wir betrachten dabei erst die Transformation von Streifen, dann die von Flächenstücken; diese kann man ja aus Streifen aufbauen.

Ein Punkt P im x -Raum durchlaufe irgendeinen Streifen l . Eine zum Flächenkomplex gehörige BT.⁴⁾ führt l gewöhnlich in einen Streifen λ über, und dem laufenden Punkt P entspricht ein Punkt II , der λ durchläuft. Sei etwa die Bogenlänge s von l als Parameter auf l und λ gewählt. Sei $d\sigma$ das Bogendifferential auf λ und $d\tau$ das Bogendifferential des Gaußschen sphärischen Bildes von λ . Dann ist $\frac{d\sigma}{ds}$ und $\frac{d\tau}{ds}$ beschränkt und nicht gleichzeitig $d\sigma = 0$, $d\tau = 0$. Eine Unstetigkeit der „Streifenkrümmung“ $\frac{d\tau}{d\sigma}$ auf λ findet daher genau da statt, wo $d\sigma = 0$, wo also II stillsteht, während P läuft. II bezeichnet aber definitionsgemäß die Komplexfläche, die jeweils durch das Element P hindurchgeht. Der Stillstand von II bedeutet also, daß zwei oder mehrere benachbarte Elemente von l auf derselben Komplexfläche liegen.

Das Bild eines Streifens l mittels einer zum Flächenkomplex (S) gehörigen BT. wird da singulär, wo l eine Komplexfläche oskuliert.

Bekannter als dieser Satz ist der Sonderfall, daß l im ganzen auf einer festen Komplexfläche liegt. Dann wird l auf ein Bündel von Elementen abgebildet, die alle durch einen festen Punkt gehen, den Bildpunkt von P .

Die so gefundenen Singularitäten werden, wie man nach Analogie zum Verhalten ebener Kurven erwarten darf, in zwei Klassen zerfallen, je nachdem II in dem stationären Punkt seine Wegrichtung umkehrt oder

⁴⁾ Berührungstransformation.

nicht; nur Singularitäten der ersten Art werden, grob ausgedrückt, mit bloßem Auge sichtbar und vom Charakter von Rückkehrpunkten sein, entsprechend der Neilschen Parabel $y = x^{\frac{3}{2}}$ im Nullpunkt. Die Singularitäten der zweiten Art kann man mit dem Verhalten der Parabel $y = x^{\frac{4}{3}}$ im Nullpunkt vergleichen, die dort unendlich große Krümmung hat, aber qualitativ nicht von der gewöhnlichen Parabel $y = x^2$ verschieden ist.

Es gilt nun, wenn man auch die Oskulationen höherer, allerdings endlicher Ordnung berücksichtigt, der einfache Satz — im folgenden „Spitzenregel“ genannt —, daß genau die Oskulationsstellen ungerader Ordnung des Streifens durch die BT. in Rückkehrspitzen verwandelt werden.

Betrachten wir nun im x -Raum ein Flächenstück F mit einem in bezug auf (S) parabolischen Punkt P . Die Oskulationsrichtung in P sei d , L der durch P laufende parabolische Kurvenzug; dann haben wir zunächst ein Paradoxon. Einmal geht durch das Bild Π von P sicher eine L abbildende singuläre Linie des Bildes Φ von F . Man erwartet, daß alle auf Φ liegenden durch Π gehenden Streifen eine Singularität in Π haben. Diese Streifen sind aber Bilder von Streifen auf F , die durch P gehen. Von diesen Streifen können nun nach dem vorigen nur diejenigen singuläre Bilder haben, die in P die Richtung d haben. Es muß also doch nicht-singuläre Streifen auf Φ durch Π geben, und offenbar kommen dafür nur solche Streifen in Betracht, deren Kurve die singuläre Linie durch Π nicht schneidet, sondern berührt.

In der Tat löst sich das Paradoxon durch den Satz (im folgenden „Drehsatz“ genannt):

Alle von d verschiedenen Richtungen in F durch P werden auf eine und dieselbe Richtung in Φ durch Π abgebildet.

Diese Richtung ist die Richtung der singulären Linie von Φ , wenn eine solche überhaupt auftritt und die Singularität nicht in einer Kegelspitze besteht; denn außer in diesem Fall hat L nach § 2 eine von d verschiedene Richtung.

§ 4.

Die heuristischen Betrachtungen aus § 3 sollen nun formal bewiesen werden, wobei wir in der Regel von der Darstellung (A) der BT. ausgehen.

Bekanntlich gibt es eine wichtige Gattung von BT., die sich dieser Darstellung entziehen; es sind diejenigen, die man nicht aus Flächenkomplexen, sondern aus Kurvenkomplexen gewinnt. Diese Gattung ist im folgenden außer Acht gelassen.

Seien $x = x(t), \dots, q = q(t)$ die Gleichungen eines Streifens l im x -Raum. Dann ist $\dot{z} = p\dot{x} + q\dot{y}$ ($\dot{x} = \frac{dx(t)}{dt}$ usw.).

Wir wollen die Gleichungen des transformierten Streifens λ im ξ -Raum berechnen, insbesondere die Ableitungen $\dot{\xi}, \dot{\eta}, \dot{\zeta}, \dot{\pi}, \dot{\nu}$. Hierzu müssen wir (A) nach t ableiten. Das gibt fünf Gleichungen $\dot{1}, \dot{2}, \dot{3}, \dot{4}, \dot{5}$, die in den zehn Ableitungen $\dot{x}, \dots, \dot{q}; \dot{\xi}, \dots, \dot{\nu}$ homogen linear sind mit Koeffizienten, die von den 2×5 Elementkoordinaten $x, \dots, q; \xi, \dots, \nu$ abhängen. Da wir die BT. selbst an der betrachteten Stelle als nichtausgeartet annehmen, müssen die Gleichungen bei gegebenen $\dot{x}, \dot{y}, \dot{z}, \dot{p}, \dot{q}$ eindeutig nach $\dot{\xi}, \dots, \dot{\nu}$ auflösbar sein. Eine Singularität des Bildstreifens λ kann nach § 3 nur darin bestehen, daß sich $\dot{\xi} = \dot{\eta} = \dot{\zeta} = 0$ ergibt (Stillstand von Π).

Nun kommen aber $\dot{\pi}, \dot{\nu}$ in den Gleichungen $\dot{1}, \dot{2}, \dot{3}$. gar nicht vor, daher muß sich aus diesen Gleichungen allein schon $\dot{\xi} = \dot{\eta} = \dot{\zeta} = 0$ ergeben. Anders ausgedrückt: $\dot{1}, \dot{2}, \dot{3}$. müssen erfüllt sein, wenn man bei der Differentiation der linken Seiten von 1., 2., 3. nach t die ξ, η, ζ als Konstanten betrachtet. Deuten wir diese Art der Differentiation durch Klammern $\{ \dots \}$ an, so haben wir die Gleichungen:

$$\begin{aligned} 1. \quad \{f\}' &= 0 = f_x \dot{x} + f_y \dot{y} + f_z \dot{z} = (f_x + p f_z) \dot{x} + (f_y + q f_z) \dot{y}, \\ (B) \quad 2. \quad \{f_x + p f_z\}' &= 0, \\ 3. \quad \{f_y + q f_z\}' &= 0. \end{aligned}$$

(B) gibt das Kriterium dafür, daß eine Stelle eines Streifens durch die BT. singulär gemacht wird. (B) bedeutet aber nichts anderes, als daß der Streifen l im x -Raum im betrachteten Punkt die Komplexfläche $f(x, y, z; \xi, \eta, \zeta) = 0$ oskuliert in mindestens erster Ordnung. Die Oskulation ist genau erster Ordnung, wenn nicht gleichzeitig auch noch $\{f_x + p f_z\}'' = 0, \{f_y + q f_z\}'' = 0$; sonst ist sie mindestens zweiter Ordnung. Im ersten Fall ergibt die doppelte Differentiation von (A) 1., 2., 3., wenn man nachher $\dot{\xi} = \dot{\eta} = \dot{\zeta} = 0$ einsetzt, offenbar, daß nicht gleichzeitig $\ddot{\xi} = \ddot{\eta} = \ddot{\zeta} = 0$. Dagegen ist $\ddot{\xi} = \ddot{\eta} = \ddot{\zeta} = 0$ im zweiten Fall. Wie aus den Anfangsgründen der Flächentheorie bekannt, bedeutet die Aussage: „ $\dot{\xi} = \dot{\eta} = \dot{\zeta} = 0$; nicht gleichzeitig $\ddot{\xi} = \ddot{\eta} = \ddot{\zeta} = 0$ “, daß die Kurve $\xi = \xi(t), \eta = \eta(t), \zeta = \zeta(t)$ im betrachteten Punkt eine Spitze hat. Damit ist die Spitzenregel für den einfachsten Fall bewiesen. Der Beweis für Oskulationen beliebig hoher Ordnung soll hier nicht ausgeführt werden; wenn man (A) 1., 2., 3. weiterdifferenziert und das Verschwinden der niederen Ableitungen beachtet, gewinnt man leicht eine geeignete Rekursionsformel für den allgemeinen Fall.

Wir betrachten nunmehr ein Flächenstück F im x -Raum in einem Punkt P . Um das Vorige anwenden zu können, denken wir uns das Stück aus irgend ∞^1 Streifen aufgebaut, die in allen Richtungen durch das Element P laufen.

Für alle diese Streifen gelten in P die Streifenrelationen

$$(C) \quad \begin{aligned} 1. \quad \dot{z} &= p \dot{x} + q \dot{y}, \\ 2. \quad \dot{p} &= r \dot{x} + s \dot{y}, \\ 3. \quad \dot{q} &= s \dot{x} + t \dot{y}, \end{aligned}$$

wobei r, s, t die in der üblichen Weise bezeichneten zweiten Ableitungen von z auf der Fläche nach x, y sind.

Auf der Bildfläche Φ von F müssen in Bild Π von P die entsprechenden Formeln gelten:

$$(C') \quad \begin{aligned} 1. \quad \dot{\xi} &= \pi \dot{\xi} + \kappa \dot{\eta}, \\ 2. \quad \dot{\pi} &= \rho \dot{\xi} + \sigma \dot{\eta}, \\ 3. \quad \dot{\kappa} &= \sigma \dot{\xi} + \tau \dot{\eta}. \end{aligned}$$

Um zu verfolgen, wie ρ, σ, τ sich aus $x, y, z; p, q; r, s, t$ berechnen, gehen wir auf die fünf Gleichungen (A) 1., 2., 3., 4., 5. zurück. Mittels (C) kann man $\dot{z}, \dot{p}, \dot{q}$ aus jenen Gleichungen eliminieren. Dann sind die Koeffizienten von \dot{x}, \dot{y} linear abhängig von r, s, t ; die Koeffizienten von $\dot{\xi}, \dots, \dot{\kappa}$ hängen nur von den Elementen Π, P ab, die Gleichungen müssen wegen des regulären Verhaltens der Transformation in P, Π nach $\dot{\xi}, \dots, \dot{\kappa}$ auflösbar sein. Dies gibt ein Schema

$$(D) \quad \begin{aligned} 1. \quad \dot{\xi} &= a_1 \dot{x} + b_1 \dot{y}, \\ 2. \quad \dot{\eta} &= a_2 \dot{x} + b_2 \dot{y}, \\ 3. \quad \dot{\zeta} &= a_3 \dot{x} + b_3 \dot{y}, \\ 4. \quad \dot{\pi} &= a_4 \dot{x} + b_4 \dot{y}, \\ 5. \quad \dot{\kappa} &= a_5 \dot{x} + b_5 \dot{y}. \end{aligned}$$

Die a_i, b_i ($i = 1, \dots, 5$) hängen linear von r, s, t ab. Ist

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = \Delta(r, s, t) \neq 0,$$

so lassen sich (D) 1., 2. nach \dot{x}, \dot{y} auflösen:

$$\begin{aligned} \dot{x} &= \alpha_1 \dot{\xi} + \beta_1 \dot{\eta}, \\ \dot{y} &= \alpha_2 \dot{\xi} + \beta_2 \dot{\eta}. \end{aligned}$$

Die α, β sind gebrochene Funktionen von r, s, t mit linearen Zählern und dem Nenner Δ , der in r, s, t quadratisch ist. Die Gleichungen (D) 4., 5. erhalten somit die Gestalt

$$\begin{aligned} \dot{\pi} &= (a_4 \alpha_1 + b_4 \alpha_2) \dot{\xi} + (a_4 \beta_1 + b_4 \beta_2) \dot{\eta}, \\ \dot{\kappa} &= (a_5 \alpha_1 + b_5 \alpha_2) \dot{\xi} + (a_5 \beta_1 + b_5 \beta_2) \dot{\eta}. \end{aligned}$$

Hieraus folgt aber (wegen der Willkürlichkeit von $\dot{x}, \dot{y}, \dot{z}, \dot{p}, \dot{q}$ bzw. von $\dot{\xi}, \dot{\eta}, \dot{\zeta}, \dot{p}, \dot{x}$) nach (C')

$$(E) \quad \begin{aligned} 1. & \quad a_4 \alpha_1 + b_4 \alpha_2 = \rho, \\ 2. & \quad a_4 \beta_1 + b_4 \beta_2 = a_5 \alpha_1 + b_5 \alpha_2 = \sigma, \\ 3. & \quad a_5 \beta_1 + b_5 \beta_2 = \tau. \end{aligned}$$

ρ, σ, τ sind also gebrochene Funktionen von r, s, t mit dem quadratischen gemeinsamen Nenner Δ und quadratischen Zählern; die in (E) 2. enthaltene Gleichung zwischen den a, b muß identisch in r, s, t gelten. Daß dies der Fall ist, liegt bekanntlich im Wesen jeder BT.

Die Abbildung wird dann und nur dann singularär, wenn $\Delta = 0$. Wir müssen diese Gleichung als Oskulationsgleichung deuten. In der Tat besagen (D) 1. 2. im Falle $\Delta = 0$: Es gibt in P eine Flächenrichtung $\dot{x}_0 : \dot{y}_0$, so daß

$$\begin{aligned} a_1 \dot{x}_0 + b_1 \dot{y}_0 &= 0; & \underline{\dot{\xi}} &= 0; & \underline{\dot{\zeta}} &= \pi \dot{\xi} + \kappa \dot{\eta} = 0. \\ a_2 \dot{x}_0 + b_2 \dot{y}_0 &= 0; & \underline{\dot{\eta}} &= 0; & & \end{aligned}$$

Jeder Streifen der Fläche, der in P die Richtung $\dot{x}_0 : \dot{y}_0$ hat, oskuliert also dort die durch das Element P gehende Komplexfläche. P ist parabolisch, die Oskulationsrichtung d ist durch $\dot{x}_0 : \dot{y}_0$ gegeben. Unsere Annahme, daß es nur eine Oskulationsrichtung in P gibt, besagt, daß in P nicht $a_1 = b_1 = a_2 = b_2 = 0$; in der Tat würde nur dann $\dot{x}_0 : \dot{y}_0$ unbestimmt werden.

Ist $\dot{x} : \dot{y} \neq \dot{x}_0 : \dot{y}_0$ in P , so ergeben (D) 1. 2. nicht mehr $\dot{\xi} = \dot{\eta} = 0$; aber offenbar ist dann wegen $\Delta = 0$ das Verhältnis $\dot{\xi} : \dot{\eta}$ unabhängig von \dot{x}, \dot{y} . Damit ist der Drehsatz bewiesen: Alle Richtungen durch P , die von d verschieden sind, werden auf eine und dieselbe Richtung durch Π abgebildet; insbesondere ist diese die Richtung der singularären Linie durch Π , die ja als Bild der parabolischen Kurve $\Delta = 0$ durch P nur auftritt, wenn diese Kurve in P nicht die Oskulationsrichtung d hat.

Wie aus dem Früheren klar ist, kann die singularäre Linie durch Π entweder eine Rückkehrkante oder eine unsichtbare singularäre Linie sein. Beispiele beider Typen erhalten wir, indem wir einerseits die Neilsche Parabel $y = x^{\frac{2}{3}}$, andererseits die Parabel $y = x^{\frac{4}{3}}$ durch Parallelverschiebung aus ihren Ebenen heraus zu Zylinderflächen machen.

Es zeigt sich nun, daß der Fall der unsichtbaren singularären Linie dann und nur dann eintritt, wenn $\Delta = 0$ auf F Doppelkurve bzw. $2n$ -fache Kurve ist, d. h. wenn Δ zu beiden Seiten der Kurve $\Delta = 0$ auf F gleiches Vorzeichen hat.

Der Beweis soll hier nicht ausgeführt werden. Er beruht darauf, daß genau n -fache Oskulation von F mit der Komplexfläche durch P stattfindet, wenn Δ in P genau von n -ter Ordnung verschwindet.

„Im allgemeinen“ wird also die parabolische Kurve durch die BT. in eine Rückkehrkante verwandelt. Außerdem können, abgesehen von Kegelspitzen und unsichtbaren singulären Linien, noch kompliziertere Singularitäten durch die BT. erzeugt werden, wenn die parabolische Kurve der Ausgangsfläche z. B. Doppelpunkte oder Spitzen hat (siehe § 12). Endlich treten isolierte singuläre Punkte als Bilder isolierter parabolischer Punkte auf (§ 14). Da die Determinante Δ in der Umgebung einer isolierten Nullstelle nicht das Vorzeichen wechseln kann, sind isolierte Singularitäten immer unsichtbar.

§ 5.

Im vorigen Abschnitt ist die Transformation der zweiten Ableitungen ihrem formalen Charakter nach erst sehr oberflächlich beschrieben worden.

ϱ, σ, τ sind gebrochene quadratische Funktionen von r, s, t . Da man auch r, s, t in derselben Weise aus ϱ, σ, τ erhält, indem man ξ -Raum und x -Raum vertauscht, so muß die Transformation der zweiten Ableitungen eine *Cremonatransformation* sein.

Die genauere Betrachtung zeigt nun, daß diese Cremonatransformation von folgendem einfachen Charakter ist: Man setze $r = \frac{r_1}{r_5}, s = \frac{r_2}{r_5}, t = \frac{r_3}{r_5}$, $rt - s^2 = \frac{r_4}{r_5}$; analog $\varrho = \frac{\varrho_1}{\varrho_5}, \sigma = \frac{\varrho_2}{\varrho_5}, \tau = \frac{\varrho_3}{\varrho_5}, \varrho\tau - \sigma^2 = \frac{\varrho_4}{\varrho_5}$, so daß in den r_i, ϱ_i die Relationen gelten:

$$r_1 r_3 - r_2^2 - r_4 r_5 = 0; \quad \varrho_1 \varrho_3 - \varrho_2^2 - \varrho_4 \varrho_5 = 0.$$

Dann erhalten die Transformationsgleichungen der zweiten Ableitungen den Charakter homogener linearer Transformationen in den r_i, ϱ_i :

$$(G) \quad \varrho_i = \sum_{k=1}^5 c_i^k r_k \quad (i = 1, \dots, 5);$$

$|c_i^k| \neq 0$ an allen regulären Stellen der BT.

Dieses Resultat ist von Engel und seinem Schüler Spitz⁵⁾ gewonnen worden. Man kann es offenbar durch Ausführung der in § 4 angedeuteten Rechnungen bestätigen. Der eigentliche Grund dafür wird vielleicht aus der folgenden geometrischen Deutung klarer.

$\dot{x}, \dot{y}, \dot{p}, \dot{q}$ seien irgendwelche projektiven homogenen Koordinaten eines dreidimensionalen Raumes (R); $\dot{\xi}, \dot{\eta}, \dot{\pi}, \dot{\kappa}$ entsprechende Koordinaten eines andern Raumes (P). Die BT. bestimmt (für $\dot{z} = p\dot{x} + q\dot{y}, \dot{\zeta} = \pi\dot{\xi} + \kappa\dot{\eta}$) mittels (A) 2., 3., 4., 5. eine projektive Abbildung (R) \rightarrow (P). Eine Ge-

⁵⁾ Engel, Die höheren Differentialquotienten, Leipz. Ber. 1894 u. 1902. Spitz, Zur Theorie der Elemente höherer Ordnung in der Ebene und im Raum, Diss. Greifswald 1912.

rade (r) in (R) kann stets durch zwei Gleichungen folgender Gestalt gegeben werden:

$$\begin{aligned} r_5 \dot{p} - r_1 \dot{x} - r_2 \dot{y} &= 0, \\ r_5 \dot{q} - r'_2 \dot{x} - r_3 \dot{y} &= 0. \end{aligned}$$

Bestimmt man ein \dot{r}_4 durch die Gleichung

$$r_1 r_3 - r_2 r'_2 - r_4 r_5 = 0,$$

so erkennt man leicht, daß die $r_1, r_2, r'_2, r_3, r_4, r_5$ *Kleinsche Linienkoordinaten* der Geraden (r) sind; d. h. bei der Abbildung (R) \rightarrow (P) geht (r) in eine Gerade (ϱ) über, die durch die Gleichungen

$$\begin{aligned} \varrho_5 \dot{\pi} - \varrho_1 \dot{\xi} - \varrho_2 \dot{\eta} &= 0, \\ \varrho_5 \dot{\kappa} - \varrho'_2 \dot{\xi} - \varrho_3 \dot{\eta} &= 0 \end{aligned}$$

gegeben wird, wobei die ϱ_i sich homogen linear aus den r_i transformieren; wenn man noch ϱ_4 durch die Gleichung einführt,

$$\varrho_1 \varrho_3 - \varrho_2 \varrho'_2 - \varrho_4 \varrho_5 = 0,$$

so drückt sich auch ϱ_4 durch die r_i linear aus, desgl. umgekehrt die r_i durch

$$\varrho_1, \varrho_2, \varrho'_2, \varrho_3, \varrho_4, \varrho_5.$$

Eine homogene lineare Gleichung zwischen den r_i bestimmt einen *linearen Geradenkomplex* in (R). Ein solcher wird durch (R) \rightarrow (P) auf einen linearen Geradenkomplex in (P) abgebildet. Bringt man das in Zusammenhang mit der Bedeutung der $\dot{x}, \dots, \dot{q}; \dot{\xi}, \dots, \dot{\kappa}$ im x -Raum und im ξ -Raum, so kann man sagen:

I. Einem System zweiter Ableitungen r, s, t im x -Raum entspricht im (R)-Raum eine Gerade (r) des Komplexes $r_2 = r'_2$ mittels der Gleichungen

$$r = \frac{r_1}{r_5}, \quad s = \frac{r_2}{r_5} = \frac{r'_2}{r_5}, \quad t = \frac{r_3}{r_5}; \quad r t - s^2 = \frac{r_4}{r_5}.$$

II. Jede projektive Abbildung (R) \rightarrow (P), die einer BT. entspricht, führt den Komplex $r_2 = r'_2$ in den Komplex $\varrho_2 = \varrho'_2$ über, und der aus (r) transformierten Geraden (ϱ) entspricht dabei das aus r, s, t transformierte System zweiter Ableitungen ϱ, σ, τ im ξ -Raum mittels der Gleichungen

$$\varrho = \frac{\varrho_1}{\varrho_5}, \quad \sigma = \frac{\varrho_2}{\varrho_5} = \frac{\varrho'_2}{\varrho_5}, \quad \tau = \frac{\varrho_3}{\varrho_5}, \quad \varrho \tau - \sigma^2 = \frac{\varrho_4}{\varrho_5}.$$

Damit ist der formale Charakter der BT. in den zweiten Ableitungen, insbesondere der Sinn des Systems (G), soweit aufgeklärt, wie im folgenden erforderlich.

Offenbar muß sich die Gleichung $\Delta = 0$ aus § 4, die die singular werden- den Elemente des x -Raumes gibt, in die Form $\varrho_5 = 0$ bringen lassen. Wenn die BT. keine erweiterte Punkttransformation ist, muß $\Delta = 0$ auch nichtsinguläre Elemente ($r_5 \neq 0$) des x -Raumes enthalten; wie aus §§ 3, 4 und dem vorigen hervorgeht, ist $\Delta = 0$ eine Monge-Ampèresche Gleichung im x -Raum. Nur für erweiterte Punkttransformationen kann also das System (G) die Gleichung $\varrho_5 = r_5$ enthalten; für erweiterte Punkttransformationen muß diese Gleichung aber auch wirklich gelten, weil durch eine Punkttransformation, soweit sie nicht ausgeartet ist, stetig gekrümmte Flächen in stetig gekrümmte Flächen übergehen.

Daraus, daß $rt - s^2$ mit r, s, t gleichberechtigt in den Transformationsgleichungen auftritt, folgt, daß $rt - s^2$ in der Nähe der hier behandelten Singularitäten schwächer unendlich wird, als ein allgemeiner in r, s, t quadratischer Ausdruck; nämlich von gleicher Ordnung wie ein in r, s, t linearer Ausdruck.

Zum Schluß seien der Deutlichkeit halber die Hauptformeln der vorigen Paragraphen noch einmal für den besonders einfachen Fall der Legendre-Transformation aufgestellt, die im folgenden hauptsächlich angewandt wird.

$$1. f(x, y, z; \xi, \eta, \zeta) = z + \zeta - x\xi - y\eta = 0;$$

$$2. -\xi + p = 0;$$

$$3. -\eta + q = 0;$$

$$4. -x + \pi = 0;$$

$$5. -y + \kappa = 0;$$

(A₀)

$$1. \dot{z} - p\dot{x} - q\dot{y} + \dot{\zeta} - \pi\dot{\xi} - \kappa\dot{\eta} = 0;$$

$$2. -\dot{\xi} + \dot{p} = 0;$$

$$3. -\dot{\eta} + \dot{q} = 0;$$

$$4. -\dot{x} + \dot{\pi} = 0;$$

$$5. -\dot{y} + \dot{\kappa} = 0.$$

$$1. \dot{z} - p\dot{x} - q\dot{y} = 0;$$

(B₀)

$$2. \dot{p} = 0;$$

$$3. \dot{q} = 0.$$

$$1. \dot{\xi} = r\dot{x} + s\dot{y};$$

$$2. \dot{\eta} = s\dot{x} + t\dot{y};$$

(D₀)

$$3. \dot{\zeta} = (xr + ys)\dot{x} + (xs + yt)\dot{y},$$

$$4. \dot{\pi} = \dot{x},$$

$$5. \dot{\kappa} = \dot{y};$$

$$\Delta = rt - s^2;$$

$$\dot{x} = \frac{t}{rt - s^2} \dot{\xi} - \frac{s}{rt - s^2} \dot{\eta};$$

$$\dot{y} = -\frac{s}{rt - s^2} \dot{\xi} + \frac{r}{rt - s^2} \dot{\eta}$$

$$(E_0) \quad \varrho = \frac{t}{rt - s^2}, \quad \sigma = \frac{-s}{rt - s^2}, \quad \tau = \frac{r}{rt - s^2}.$$

$$\varrho_1 = r_3,$$

$$\varrho_2 = -r_2,$$

$$(G_0) \quad \varrho_3 = r_1,$$

$$\varrho_4 = r_5,$$

$$\varrho_5 = r_4.$$

Zweiter Teil.

§ 6.

Wir benutzen die Ergebnisse des ersten Teils, um das *Cauchysche Problem* der Integration partieller Differentialgleichungen zweiter Ordnung in solchen Fällen in einem erweiterten Sinne zu lösen, in denen man es bisher als unlösbar bezeichnete, und in denen es auch im gewöhnlichen Sinne unlösbar bleibt.

Es sei die partielle Differentialgleichung zweiter Ordnung

$$\Phi(\xi, \eta, \zeta; \pi, \kappa; \varrho, \sigma, \tau) = 0$$

vorgelegt, sowie ein analytischer⁶⁾ Streifen erster Ordnung λ :

$$\xi = \xi(t), \quad \eta = \eta(t), \quad \zeta = \zeta(t), \quad \pi = \pi(t), \quad \kappa = \kappa(t); \quad \dot{\zeta} = \pi \dot{\xi} + \kappa \dot{\eta}.$$

Es soll eine Integralfläche $\zeta = \zeta(\xi, \eta)$ der DFGL.⁷⁾ durch den gegebenen Streifen gelegt werden; darin besteht bekanntlich das Cauchysche Problem. Die drei partiellen Ableitungen ϱ, σ, τ haben auf dem Streifen den drei Gleichungen zu genügen:

$$(H) \quad \begin{aligned} 1. & \quad \Phi = 0, \\ 2. & \quad \dot{\pi} - \varrho \dot{\xi} - \sigma \dot{\eta} = 0, \\ 3. & \quad \dot{\kappa} - \sigma \dot{\xi} - \tau \dot{\eta} = 0. \end{aligned}$$

Drei Gleichungen mit drei Unbekannten können 1. abzählbar viele, 2. kontinuierlich unendlich viele, 3. gar keine Lösungen haben. Tritt in (H) der erste Fall ein, so hat bekanntlich das Cauchy-Problem so viele Lösungen, wie (H) Lösungssysteme ϱ, σ, τ hat; zu jedem dieser Lösungssysteme gehört genau ein Integral, das in λ die betreffenden ϱ, σ, τ aufweist. Im zweiten Fall ist λ charakteristischer Streifen der DFGL., und die Lösung ist nur bis auf eine willkürliche Funktion einer Veränderlichen bestimmt.

⁶⁾ Zu dieser Annahme vgl. die Einleitung.

⁷⁾ Der partiellen Differentialgleichung zweiter Ordnung.

Im dritten Fall hat das Problem keine Lösung im gewöhnlichen Sinn. *Diesen Fall wollen wir nun mittels der Ergebnisse des ersten Teils auf den ersten Fall zurückführen.*

Ersetzt man nämlich ϱ, σ, τ gemäß § 5 durch $\varrho_1, \dots, \varrho_5$, so transformiert sich $\Phi(\xi, \dots, \kappa; \varrho, \sigma, \tau) = 0$ in eine in den ϱ_i homogene Gleichung

$$\Psi(\xi, \dots, \kappa; \varrho_1, \varrho_2, \varrho_3, \varrho_4, \varrho_5) = 0,$$

und das Cauchy-Problem wird nunmehr durch die vier *homogenen* Gleichungen gegeben:

$$(J) \quad \begin{aligned} 1. & \Psi = 0, \\ 2. & \varrho_1 \varrho_3 - \varrho_2^2 - \varrho_4 \varrho_5 = 0, \\ 3. & \varrho_5 \dot{x} - \varrho_1 \dot{\xi} - \varrho_2 \dot{\eta} = 0, \\ 4. & \varrho_5 \dot{z} - \varrho_2 \dot{\xi} - \varrho_3 \dot{\eta} = 0. \end{aligned}$$

Vier homogene Gleichungen mit fünf Unbekannten haben immer Lösungssysteme, in denen nicht alle Unbekannten verschwinden. In den beiden ersten Fällen von (H) hat (J) Lösungen mit $\varrho_5 \neq 0$; im dritten Fall ist offenbar stets $\varrho_5 = 0$.⁸⁾

Angenommen nun, wir haben eine BT.

$$\xi \rightarrow x,$$

so daß auf dem aus λ transformierten Streifen l das aus (J) transformierte System (J') eine Lösung mit $r_5 \neq 0$ hat⁹⁾. Dann kann man zur gewöhnlichen inhomogenen Form (H') von (J') übergehen, indem man durch r_5 (bzw. durch eine Potenz von r_5) überall dividiert und r_4 eliminiert. (H') verwirklicht also nicht mehr den dritten, sondern den ersten oder zweiten Fall des transformierten Cauchy-Problems. Offenbar kann aber

⁸⁾ Wegen des nichtlinearen Charakters von (J) 2. ist folgende Fehlerquelle zu vermeiden: $\varrho_5 = 0$ löst meistens (J) auch in solchen Fällen, wo $\varrho_5 \neq 0$ die allein sinnvollen Lösungen liefert. In der Tat sind (J) 2., 3., 4. erfüllt durch $\varrho_5 = 0$; $\varrho_1 = \dot{\eta}^2$, $\varrho_2 = -\dot{\xi} \dot{\eta}$, $\varrho_3 = \dot{\xi}^2$. Wenn in (J) 1. explizit ϱ_4 vorkommt, kann man jene Werte von $\varrho_1, \varrho_2, \varrho_3, \varrho_5$ in (J) 1. eintragen und (J) 1. durch passende Wahl von ϱ_4 erfüllen.

Dieses System ist aber für das Folgende unbrauchbar, weil es, wie man am Beispiel der Legendre-Transformation leicht sieht, nicht invariant gegen BT. ist; das transformierte System wird die aus (J) transformierten Gleichungen im allgemeinen nicht lösen. Etwas Invariantes erhält man erst, wenn man zu (J) noch die Gleichungen

$$5. \varrho_4 \dot{\xi} - \varrho_3 \dot{x} + \varrho_2 \dot{z} = 0,$$

$$6. \varrho_4 \dot{\eta} + \varrho_2 \dot{z} - \varrho_1 \dot{z} = 0$$

hinzunimmt, die für $\varrho_5 \neq 0$ aus (J) 2., 3., 4. folgen, für $\varrho_5 = 0$ aber nicht; die Deutung der ϱ_i als Linienkoordinaten klärt diese Erscheinung auf und zeigt auch, warum man zur Ergänzung von (J) gerade Gleichungen von der Form 5., 6. braucht.

⁹⁾ Bezeichnungen $x, \dots, q; r, s, t; r_1, \dots, r_5$ wie im ersten Teil.

nur der erste Fall verwirklicht sein. Denn der zweite, der die charakteristischen Streifen kennzeichnet, ist gegen BT. bekanntlich invariant. Da er in (H) nicht eintrat, kann er auch in (H') nicht eintreten. Zu (H') gehören also abzählbar viele Integrale $z = z_*(x, y)$. Indem wir nun durch die zu $\xi \rightarrow x$ inverse BT. $x \rightarrow \xi$ wieder in den ursprünglichen ξ -Raum zurückkehren, verwandeln wir die Flächen $z = z_*(x, y)$ in Flächen $\zeta = \zeta_*(\xi, \eta)$. Diese Flächen erfüllen die ursprüngliche DFGL. $\Phi = 0$ und gehen durch den ursprünglichen Streifen λ . Wir bezeichnen sie daher als die Lösungen des gegebenen Cauchy-Problems; von der BT. $\xi \rightarrow x$ sagen wir, sie „regularisiert“ das Problem.

Damit hätten wir den dritten Fall des Cauchy-Problems in dem Sinne gelöst, daß wir durch den Streifen eine Integralfläche legen können, die dort singular wird nach Art des in §§ 3, 4 Besprochenen. Dabei bleiben aber noch drei Fragen, von denen in diesem Abschnitt nur die beiden ersten gestellt und beantwortet werden sollen.

1. Ist die Methode immer anwendbar? D. h. gibt es im dritten Fall des Cauchy-Problems immer regularisierende BT.?

2. Ist die Methode eindeutig? D. h. ist das Integral, das durch irgendeine Regularisierung gewonnen wurde, unabhängig von der zu dieser Regularisierung benutzten BT.?

Beide Fragen sind zu bejahen.

Um im allgemeinen dritten Fall eine regularisierende BT. in einfacher Weise zu finden, versuche man zunächst, durch Legendre-Transformationen zu regularisieren; nach (G_0) , § 5 gelingt das nur dann nicht, wenn aus (J) nicht nur $\varrho_5 = 0$, sondern auch $\varrho_4 = 0$ folgt. In diesem Fall gibt es aber sicher irgendeine lineare Kombination $\alpha_1 \varrho_1 + 2\alpha_2 \varrho_2 + \alpha_3 \varrho_3$ mit konstanten reellen $\alpha_1, \alpha_2, \alpha_3$, die längs eines beliebig großen Stückes von λ nicht verschwindet. Macht man nun die Punkttransformation $\xi \rightarrow \bar{\xi}$:

$$\bar{\xi} = \xi, \quad \bar{\eta} = \eta, \quad \bar{\zeta} = \zeta + \frac{\alpha_3}{2} \xi^2 - \alpha_2 \xi \eta + \frac{\alpha_1}{2} \eta^2,$$

so erhält man für die Ableitungen:

$$\begin{aligned} \bar{\pi} &= \pi + \alpha_3 \xi - \alpha_2 \eta, & \bar{x} &= x - \alpha_2 \xi + \alpha_1 \eta; \\ \bar{\varrho} &= \varrho + \alpha_3, & \bar{\sigma} &= \sigma - \alpha_2, & \bar{\tau} &= \tau + \alpha_1. \end{aligned}$$

Die drei letzten Gleichungen geben in homogener Schreibung (weil nach § 5 $\bar{\varrho}_5 = \varrho_5$):

$$\begin{aligned} \bar{\varrho}_1 &= \varrho_1 + \alpha_3 \varrho_5, & \bar{\varrho}_2 &= \varrho_2 - \alpha_2 \varrho_5, & \bar{\varrho}_3 &= \varrho_3 + \alpha_1 \varrho_5; \\ \bar{\varrho}_4 &= \varrho_4 + \alpha_1 \varrho_1 + 2\alpha_2 \varrho_2 + \alpha_3 \varrho_3 + (\alpha_1 \alpha_3 - \alpha_2^2) \varrho_5. \end{aligned}$$

Für das transformierte Cauchy-Problem im $\bar{\xi}$ -Raum ist auf $\bar{\lambda}$, dem Bild von λ , wegen $\varrho_4 = \varrho_5 = 0$: $\bar{\varrho}_4 = \alpha_1 \varrho_1 + 2\alpha_2 \varrho_2 + \alpha_3 \varrho_3 \neq 0$. Man kann

also dieses (wegen $\bar{\varrho}_5 = \varrho_5$, natürlich noch singuläre) Cauchy-Problem nunmehr durch Legendre-Transformation ($\bar{\xi} \rightarrow x$) regularisieren.

Die aus $\xi \rightarrow \bar{\xi}$, $\bar{\xi} \rightarrow x$ zusammengesetzte Transformation $\xi \rightarrow x$ ist offenbar eine BT.; sie regularisiert das gegebene Problem.

Man erkennt dabei, daß jedes Cauchy-Problem dritter Art sogar durch „fast jede“ BT. regularisiert wird.

Um die zweite — scheinbar schwerere — Frage zu bejahen, genügt eine einfache Betrachtung von (J). Alle Integrale, die mittels beliebiger Regularisierung durch λ gelegt sind, erfüllen (J) in λ mit $\varrho_5 = 0$; das System der übrigen ϱ_i ist durch (J) auf einen festen Vorrat beschränkt, der nichts mit Regularisierung zu tun hat. Sei (ϱ) ein System aus diesem Vorrat. J_1, J_2 seien zwei durch verschiedene Regularisierungen gewonnene Integrale, deren homogene zweite Ableitungen in λ beide die Werte (ϱ) haben. Dann ist zu zeigen, daß J_1 mit J_2 identisch ist; das ist aber klar; denn eine beliebige Regularisierung führt (ϱ) auf dem Bilde l von λ in ein — von J_1, J_2 völlig unabhängiges — System (r) mit $r_5 \neq 0$ über; J_1, J_2 gehen über in zwei Integralfächen einer und derselben DFGL. durch einen und denselben Streifen l und haben in l dieselben endlichen zweiten Ableitungen r, s, t . Wie oben bewiesen, ist l nicht charakteristisch. Nach dem Eindeutigkeitssatz von Cauchy-Kowalewska ist daher J_1 mit J_2 identisch, w. z. b. w.

§ 7.

Der Leser erprobe die im vorigen gegebene Methode an der einfachen DFGL. $\sigma = 0$. Als Streifen dritter Art ergeben sich nur Streifen λ , auf denen $\dot{\xi} = 0$, $\dot{\pi} \neq 0$, oder $\dot{\eta} = 0$, $\dot{\kappa} \neq 0$. Man kann durch Legendre-Transformation regularisieren. Aber man erhält dadurch im ξ -Raum gar keine Fläche, sondern einen (nach Lie) „ausgearteten Elementverein“, der ganz auf der Kurve des Streifens λ liegt. Man erkennt leicht, daß auch für solche ausgearteten Integrale der in § 6 bewiesene Satz gilt, daß sie unabhängig von der Art der Regularisierung herauskommen.

Die in § 6 angekündigte dritte Frage ist nun: Wann artet der dritte Fall aus? D. h. welche Streifen dritter Art geben ausgeartete Elementvereine als Integrale?

Sei λ die Trägerkurve eines solchen ausgearteten Integrals; es enthält nur Streifen, die auf λ liegen. Alle diese Streifen müssen nun zweiter oder dritter Art sein. Denn zu einem Streifen erster Art würde nicht das ausgeartete Integral, sondern eine gewöhnliche, durch den Streifen regulär hindurchgehende Integralfäche gehören; auch zweiter Art können nicht alle Streifen des Integrals sein, sondern offenbar sind „fast alle“ dritter Art.

Nehmen wir umgekehrt an, ein Streifen dritter Art gehöre zu einer Schar Σ von Streifen auf der Kurve λ , die ebenfalls dritter Art sind. Jede regularisierende BT. $\xi \rightarrow x$ führt Σ in eine Fläche S über; man erkennt nun leicht, daß S der transformierten DFGL. genügt; denn dritter Art ist ein Streifen auf der Kurve λ offenbar genau dann, wenn wir in (J) die gegebene DFGL. (J) 1. durch die Gleichung $\varrho_3 = 0$ ersetzen können. Die Bildfläche S von Σ erfüllt die aus $\varrho_3 = 0$ transformierte DFGL. Also auch die aus (J) 1. transformierte.

Dafür, daß ein Streifen dritter Art ein ausgeartetes Integral ergibt, ist notwendig und hinreichend, daß er zu einem Büschel von Streifen dritter Art gehört, die mit ihm auf derselben Trägerkurve liegen.

§ 8.

Im nicht-ausgearteten dritten Fall liefert die Methode von § 6 eine *Parameterdarstellung des Integrals*, die den Namen „Regularisierung“ erst rechtfertigt. Das in λ singuläre Integral $\zeta = \zeta(\xi, \eta)$ geht durch $\xi \rightarrow x$ ja über in eine Fläche $z = z(x, y)$, die in der Nähe des Bildes von λ regulär ist. Die BT. $\xi \rightarrow x$ liefert nun ξ, η, ζ allgemein als reguläre Funktionen von x, y, z, p, q ; setzt man hier $z = z(x, y)$, $p = \frac{\partial z(x, y)}{\partial x}$, $q = \frac{\partial z(x, y)}{\partial y}$ ein, so ergeben sich drei Gleichungen $\xi = \varphi_1(x, y)$, $\eta = \varphi_2(x, y)$, $\zeta = \varphi_3(x, y)$, mit regulären $\varphi_1(x, y)$, $\varphi_2(x, y)$, $\varphi_3(x, y)$; diese Gleichungen stellen offenbar die Fläche $\zeta = \zeta(\xi, \eta)$ dar, bezogen auf x, y als Parameter. Hierdurch wird es erleichtert, die Fläche $\zeta = \zeta(\xi, \eta)$ in der Nähe des singulären Streifens zu diskutieren, insbesondere den Verlauf der Charakteristiken zu verfolgen.

Dieses und weitere allgemeine Erscheinungen, die mit der Regularisierung zusammenhängen, meist Anwendungen des Drehsatzes und der Spitzenregel aus §§ 3, 4, sollen erst im folgenden Teil am konkreten Fall der Biegungsgleichung zur Sprache kommen; die systematische Reihenfolge wird dadurch verwischt, dafür das Verständnis hoffentlich erleichtert.

Die Diskussion des allgemeinen Falls soll hier abgebrochen werden, um die Arbeit nicht mit abstrakten Einzelfragen zu überladen. Folgendes wäre naturgemäß anzuschließen:

1. Behandlung der Kegelspitzen als singulärer Vorgaben.
2. Betrachtung singulärer Charakteristiken¹⁰⁾.
3. Einteilung der DFGL. nach dem Verhalten ihrer Streifen dritter Art. (Z. B. liefert jeder Streifen dritter Art ausgeartete Integrale bei

¹⁰⁾ Man erkennt leicht, daß $\varrho_3 = 0$ auch auf Streifen *zweiter* Art, d. h. auf charakteristischen Streifen, erfüllbar ist, aber natürlich noch keine Integralfäche eindeutig bestimmt.

allen DFGL. der Form: $\alpha_1 \rho + \alpha_2 \sigma + \alpha_3 \tau + \alpha_4 = 0$, wo $\alpha_1, \alpha_2, \alpha_3$ nicht von π, κ abhängen; dazu gehört das Beispiel $\sigma = 0$ aus § 7.)

4. Verallgemeinerung des Vorstehenden auf Differentialgleichungen höherer als zweiter Ordnung; $\rho_5 = 0$ ist dabei als *Vorgabe auf dem Anfangsstreifen* (zweiter oder höherer Ordnung) aufzufassen; das nächstliegende, bei *endlichen zweiten* Ableitungen erst die *dritten oder höheren* unendlich werden zu lassen, stößt auf Schwierigkeiten; denn nach einem bekannten Satz von Bäcklund¹¹⁾ gibt es die verallgemeinerten BT. nicht, die dieses Problem regularisieren würden.

Dritter Teil.

§ 9.

Eine Fläche kann „in sich“ regulär sein, d. h. ein reguläres, beliebig fortsetzbares Linienelement haben, und kann doch als Gebilde im Raum Singularitäten aufweisen. Die Kegel und Torsen sind die einfachsten Beispiele dafür. Als weitere Beispiele sind alle Dreh- und Schraubflächen konstanter Gaußscher Krümmung — außer den Kugeln — zu nennen, die z. B. Darboux¹²⁾ ausführlich beschreibt.

Nach den Ergebnissen des Früheren können wir jetzt das Auftreten und Aussehen singulärer Ränder auf reellen Flächen regulären Linienelements allgemein erörtern. Dabei werden wir Sätze erhalten, die selbst für den Sonderfall der Torsen und jener Dreh- und Schraubflächen vermutlich neu sind, aber viel mehr umfassen als diesen. Wir beschränken uns auf die einfachsten Fragen, insbesondere bleibt der Fall der Kegelspitzen bis auf einen kurzen Hinweis (§ 13) außer Betracht.

Bekanntlich¹³⁾ wird das Biegungsproblem, genauer das Problem, eine Fläche des gegebenen Linienelements ds^2 im Euklidischen Raum (E) (ξ, η, ζ) zu finden, nicht in der in ξ, η, ζ symmetrischen Form behandelt, in der es sich zuerst darbietet; sondern man bestimmt zunächst eine Koordinate, z. B. ζ als Integral $\zeta = \zeta(u, v)$ einer DFGL. vom Monge-Ampèreschen Typ, in bezug auf die unabhängigen Veränderlichen u, v , die irgendwelche regulären Parameter des Linienelements sind. Dadurch ist die gesuchte Fläche bis auf Bewegungen in der (ξ, η) -Ebene schon bestimmt, und die Berechnung der beiden Koordinaten $\xi = \xi(u, v)$, $\eta = \eta(u, v)$ erfordert nur noch die Integration gewöhnlicher Differentialgleichungen.

Das Cauchy-Problem in $\zeta(u, v)$ ist zunächst ohne geometrische Bedeutung, weil abhängig von der Wahl von u, v . Es ist aber folgendem

¹¹⁾ Math. Annalen 9, S. 297.

¹²⁾ Théorie des Surfaces 1, 1. Buch, Kap. 9.

¹³⁾ Zum folgenden Bericht vgl. Darboux, loc. cit. 3, 7. Buch, Kap. 5.

geometrischen Problem äquivalent: In ds^2 sei eine Kurve λ gegeben¹⁴⁾, und in (E) eine Raumkurve L . Durch L ist eine Fläche F des Linienelements ds^2 zu legen, so daß L , in F gemessen, mit λ übereinstimmt¹⁵⁾.

Die drei Fälle des Cauchy-Problems (§ 6) entsprechen dabei folgendem:

1. Fall. Die Krümmung von L und die geodätische Krümmung von λ in ds^2 sind in entsprechenden Punkten verschieden. Es gibt genau zwei Flächen der verlangten Art.

2. Fall. } Diese Krümmungen sind gleich.
3. Fall. }

2. Fall. Die Torsion $\frac{1}{\tau}$ auf L und die Gaußsche Krümmung K von ds^2 auf λ erfüllen in entsprechenden Punkten die Ennepersche Gleichung

$$\frac{1}{\tau^2} + K = 0.$$

L ist Asymptotenlinie auf F , F ist bei festem L noch verbiegbar.

3. Fall. Die Ennepersche Gleichung gilt nicht.

Eine allgemeinere Erscheinung wird sein, daß auf der Vorgabe (λ, L) Bögen erster Art mit Punkten dritter oder zweiter Art abwechseln. Darauf wird in § 14 eingegangen. Zunächst untersuchen wir den dritten Fall.

§ 10.

In diesem Abschnitt wird die Biegungsgleichung aufgestellt, das geometrische (λ, L) -Problem auf die übliche Form des Cauchy-Problems gebracht, dessen dritter Fall aufgesucht, und festgestellt, wann er ausartet.

Da keine allgemeinen Formeln, sondern geometrische Ergebnisse erstrebt werden, können wir über die Parameter u, v von ds^2 im Hinblick auf λ , und über die kartesischen Koordinaten ξ, η, ζ von (E) im Hinblick auf L willkürlich verfügen. Die Parameter u, v von ds^2 nehmen wir als „Gaußsche Koordinaten“: $u = 0$ ist die Gleichung von λ ; v ist Bogenlänge auf $u = 0$; $v = \text{konst.}$ sind geodätische Linien senkrecht $u = 0$; $u = \text{konst.}$ sind geodätisch parallel, also senkrecht $v = \text{konst.}$; u ist die Bogenlänge auf $v = \text{konst.}$ Dann hat ds^2 bekanntlich die Gestalt:

$$(I) \quad \begin{aligned} 1. \quad & ds^2 = du^2 + G(u, v) dv^2, \\ 2. \quad & G(0, v) = 1; \quad G_v(0, v) = 0; \quad G_u(0, v) = -\frac{2}{e_g}, \\ 3. \quad & G^2 K = \frac{1}{4} G_u^2 - \frac{1}{2} G G_{uu}. \end{aligned}$$

¹⁴⁾ Etwa Anfangspunkt und -richtung, und von da ab die geodätische Krümmung als Funktion der Bogenlänge.

¹⁵⁾ Offenbar ist die Zuordnung eines Punktpaares von λ, L , sowie die Zuordnung der Richtungssinne in diesem Paar allein willkürlich, alles übrige dadurch bestimmt, daß die von dort aus gemessenen Bogenlängen auf λ und L gleich sind.

Dabei ist $\frac{1}{\varrho_\nu}$ die geodätische Krümmung von λ , K die Gaußsche Krümmung von ds^2 .

Die Biegungsgleichung für $\zeta(u, v)$ erhält man nach Darboux¹³⁾ folgendermaßen: Wir haben $ds^2 = d\xi^2 + d\eta^2 + d\zeta^2$, also

$$(II) \quad ds_0^2 = ds^2 - d\zeta^2 = d\xi^2 + d\eta^2.$$

Auf der gesuchten Fläche ist $\zeta = \zeta(u, v)$; $d\zeta = \pi du + \kappa dv$; ¹⁶⁾ ds_0^2 wird dadurch auf u, v als Parameter bezogen:

$$(III) \quad \begin{aligned} ds_0^2 &= (1 - \pi^2) du^2 - 2\pi\kappa du dv + (G - \kappa^2) dv^2 \\ &= E_0 du^2 + 2F_0 du dv + G_0 dv^2. \end{aligned}$$

$K_0(E_0, F_0, G_0)$, die Gaußsche Krümmung von ds_0^2 , muß wegen (II) verschwinden. Da E_0, F_0, G_0 von π, κ abhängen, gibt das eine Gleichung für ζ ; das ist die Biegungsgleichung. Sie heißt explizit:

$$(IV) \quad \begin{aligned} 1. \quad & a_1 \varrho + 2a_2 \sigma + a_4 (\varrho \tau - \sigma^2) + a_5 = 0, \\ 2. \quad & a_1 = \frac{1}{2} (G G_u \pi - G_v \kappa), \\ 3. \quad & a_2 = \frac{1}{2} G_u \kappa, \\ 4. \quad & a_4 = G, \\ 5. \quad & a_5 = \frac{1}{2} G_{uu} [G(1 - \pi^2) - \kappa^2] - \frac{1}{4} G_u^2 (1 - \pi^2). \end{aligned}$$

Unser Cauchy-Problem bez. λ vereinfacht sich dadurch, daß λ durch $u = 0$ gegeben ist, also $\dot{u} = 0$. Indem wir $\dot{v} = 1$ nehmen, erhalten wir nach § 6 das Cauchy-Problem in der Gestalt

$$(V) \quad \begin{aligned} 1. \quad & a_1 \varrho_1 + 2a_2 \varrho_2 + a_4 \varrho_4 + a_5 \varrho_5 = 0, \\ 2. \quad & \varrho_1 \varrho_3 - \varrho_2^2 - \varrho_4 \varrho_5 = 0, \\ 3. \quad & -\varrho_2 + \dot{\pi} \varrho_5 = 0, \\ 4. \quad & -\varrho_3 + \dot{\kappa} \varrho_5 = 0, \\ 5. \quad & -\dot{\kappa} \varrho_1 + \dot{\pi} \varrho_2 + \varrho_4 = 0, \text{ } ^8) \\ 6. \quad & -\dot{\kappa} \varrho_2 + \dot{\pi} \varrho_3 = 0, \text{ } ^8) \\ 7. \quad & \dot{\zeta} - \kappa = 0. \end{aligned}$$

(V) 1., 3., 4., 5. sind in den ϱ_i linear; der dritte Fall wird also nach der elementaren Theorie linearer Gleichungen (entsprechendes gilt für alle Monge-Ampèreschen DFGL.) geliefert durch

¹⁶⁾ Anders als in den ersten beiden Teilen der Arbeit setzen wir von jetzt ab $\zeta_u = \pi$, $\zeta_v = \kappa$, $\zeta_{uu} = \varrho$, $\zeta_{uv} = \sigma$, $\zeta_{vv} = \tau$. Die homogenen $\varrho_1, \dots, \varrho_5$ entstehen dann aus ϱ, σ, τ nach § 5.

$$\begin{vmatrix} a_1 & 2a_2 & 0 & a_4 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ -\dot{\kappa} & \dot{\pi} & 0 & 1 \end{vmatrix} = 0; \quad \begin{vmatrix} a_1 & 2a_2 & 0 & a_5 \\ 0 & -1 & 0 & \dot{\pi} \\ 0 & 0 & -1 & \dot{\kappa} \\ -\dot{\kappa} & \dot{\pi} & 0 & 0 \end{vmatrix} \neq 0.$$

Indem man das ausrechnet und die Werte der a_i auf λ einsetzt, erhält man schließlich als Formeln für den dritten Fall

$$(VI) \quad \begin{aligned} 1. \quad \dot{\kappa} &= \ddot{\zeta} = \frac{1}{e_g} \cdot \pi, \\ 2. \quad \dot{\pi}^2 + \frac{2}{e_g} \kappa \dot{\pi} + K(1 - \pi^2 - \kappa^2) + \frac{1}{e_g^3} \kappa^2 &\neq 0, \\ 3. \quad e_3 &= e_2 = e_3 = 0, \\ 4. \quad e_4 - \frac{1}{e_g} \pi e_1 &= 0. \end{aligned}$$

Der zweite Fall würde eintreten, wenn in (VI) 2. nicht $\neq 0$, sondern $= 0$ stände¹⁷⁾.

Um Anschluß an §§ 1—8 zu gewinnen, haben wir u, v, ζ als Koordinaten eines „ u -Raumes“ aufzufassen, der dem ξ -Raum aus §§ 1—8 entspricht. $\dot{u}(t) = 0, \dot{v}(t) = 1, \zeta(t)$ bestimmen die Kurve λ im u -Raum, auch $\kappa(t)$ ist wegen (V) 7. schon durch die Kurve λ festgelegt. Die verschiedenen Streifen dieser Kurve werden also durch verschiedene Wahl der Funktion $\pi(t)$ festgelegt.

Nun erkennen wir, wann unser Problem dritter Art ausartet. Dies ist nach § 7 genau dann der Fall, wenn auf der Kurve λ verschiedene Streifen dritter Art liegen; d. h. wenn (VI) bei festem $\zeta = \kappa, \dot{\zeta} = \dot{\kappa}$ durch mehr als eine Funktion $\pi(t)$ erfüllt wird.

Die notwendige und hinreichende Bedingung dafür ist nach (VI) 1. $\frac{1}{e_g} = 0$; d. h. λ geodätisch. Von der zugehörigen Raumkurve L in (E) wissen wir aus § 9, daß ihre Krümmung verschwindet. L ist also (wenn wir von gewissen imaginären Kurven absehen), notwendig eine gerade Linie.

Das Problem, eine geodätische Linie auf ds^2 durch eine Gerade auf F in (E) zu verwirklichen, längs der die Ennepersche Gleichung nicht gilt, spielt unter den singulären Biegungsproblemen eine Sonderrolle; es ist durch Regularisierung nicht lösbar.

¹⁷⁾ Von hier aus kann die geometrische Interpretation des Cauchy-Problems in § 9 leicht bewiesen werden. Für $\pi = 1$ ist $\dot{\zeta}$ identisch mit der Krümmung $\frac{1}{e_L}$ von L in (E) , und (VI) 1. wird $\frac{1}{e_L} = \frac{1}{e_g}$. Für $\kappa = \kappa = 0$ geht die linke Seite von (VI) 2. über in $\frac{1}{\pi^2} + K$. Die Relationen müssen aber ihrer Invarianz wegen auch bei beliebigen π, κ gelten, womit alles gezeigt ist.

Hier drängt sich die Frage auf, ob eine andere Methode nicht dennoch eine „Lösung“ des Problems liefern könnte, die in (λ, L) vielleicht stärker singularär würde, als nach Art von §§ 1—5. Auf den allgemeinen Sinn dieser Frage wird am Schluß der Arbeit eingegangen.

§ 11.

Daß λ geodätisch ist, $\frac{1}{e_g} \equiv 0$, schließen wir im folgenden aus. Ferner beschränken wir unsere Betrachtung zunächst auf ein kleines Stück von λ mit $\frac{1}{e_g} \neq 0$, das also keinen geodätischen Wendepunkt in ds^2 hat, sondern gleichsinnig gekrümmt ist. λ soll im folgenden nicht die ganze Kurve, sondern nur dies kleine Stück bezeichnen; L bezeichne das zugehörige Kurvenstück in (E) ; ebenso wird später F nur ein kleines Stück der Lösungsfläche bedeuten, das an (λ, L) anschließt.

Wir wissen aus §§ 6, 7, daß unter unseren Voraussetzungen eine nicht-ausgeartete Lösung $\zeta = \zeta(u, v)$ des Problems im u -Raum existiert. Daß sie, wie im regulären Fall, durch passende ξ, η zu einer Lösungsfläche F durch L in (E) ergänzt werden kann, können wir noch nicht sagen.

Um das klarstellen und die Geometrie von F nahe (λ, L) untersuchen zu können, brauchen wir eine *explizite* regularisierende Darstellung von $\zeta = \zeta(u, v)$ nahe λ (vgl. § 8).

Versuchen wir sie durch Legendre-Transformation zu erhalten. Das ist nach § 6 möglich, wenn (VI) mit $e_k \neq 0$ erfüllt ist; wenn also nach (VI) 4.: $\pi \neq 0$ auf λ . Das dürfen wir voraussetzen; denn durch Drehung des Koordinatensystems in (E) können wir, wie man leicht sieht, über π auf λ weitgehend verfügen.

Wir bilden nun durch Legendre-Transformation $u \rightarrow \dot{x}$ den u -Raum auf einen x -Raum ab, dessen Koordinaten wir wie in §§ 1—8 mit $x, y, z; p, q; r, s, t; r_1, \dots, r_5$ bezeichnen. l sei das Bild von λ . l ist also durch $\dot{x} = \dot{x}, \dot{y} = \dot{y}; \dot{p} = \dot{u} = 0, \dot{q} = \dot{v} = 1$ bestimmt. (IV) 1. und (VI) werden, inhomogen geschrieben:

$$\begin{aligned}
 & 1. \quad a_1 t - 2a_2 s + a_5(rt - s^2) + a_k = 0, \\
 & 2. \quad \dot{y} = \frac{1}{e_g} x. \\
 \text{(VII)} \quad & 3. \quad \dot{x}^2 + \frac{2}{e_g} y \dot{x} + K(1 - x^2 - y^2) + \frac{1}{e_g^2} y^2 \neq 0, \\
 & 4. \quad r = s = r t - s^2 = 0, \\
 & 5. \quad t = \frac{e_g}{x} = \frac{1}{\dot{y}}.
 \end{aligned}$$

$z = z(x, y)$ sei die Lösungsfläche von (VII), also das Bild von $\zeta = \zeta(u, v)$ im u -Raum.

Aus (VII) 4. folgt, daß l auf $z(x, y)$ parabolische Kurve im gewöhnlichen Sinne ist, was natürlich auch aus §§ 2, 3 hervorgeht.

Nach §§ 3, 4 wird die Feststellung wichtig sein, ob l einfache oder mehrfache parabolische Kurve auf $z(x, y)$ ist. Zunächst aber wollen wir, um den Drehsatz aus §§ 3, 4 anwenden zu können, sehen, wie auf l in $z(x, y)$ die Oskulationsrichtung zur Richtung von l und zu den charakteristischen Richtungen von (VII) 1. liegt. Die Charakteristiken auf $z(x, y)$ sind ja die Bilder der Charakteristiken von $\zeta(u, v)$; diese aber entsprechen nach Darboux¹³⁾ den Asymptotenlinien der gesuchten Fläche F in (E) . Die Charakteristiken von $z(x, y)$ haben also geometrische Bedeutung.

Die Oskulationsrichtung d in einem Punkt von l auf $z(x, y)$ ist nach § 2 die Richtung, in der die zugehörige Normalkrümmung auf $z(x, y)$ verschwindet. Nun ist in l nach (VII) 4. $r=0$; d ist also durch $\dot{y}=0$ gegeben. Längs l ist $\dot{y} \neq 0$ nach (VII) 2., weil n. V. $\frac{1}{e_g} \neq 0$, $x = \pi \neq 0$.

l hat nirgends Oskulationsrichtung.

Die charakteristischen Richtungen $x':y'$ von (VII) 1. auf $z(x, y)$ genügen bekanntlich der Gleichung

$$(VIII) \quad (a_1 + a_3 r)x'^2 + 2(a_2 + a_5 s)x'y' + a_5 t y'^2 = 0.$$

Wäre d charakteristisch, so würde aus (VIII) folgen $a_1 + a_3 r = 0$, also nach (VII) 4. $a_1 = 0$, also $\frac{1}{e_g} \pi = 0$ auf λ , was nicht der Fall ist.

Die Oskulationsrichtung ist nicht charakteristisch.

Nach (VIII) fallen die charakteristischen Richtungen auf $z(x, y)$ in l genau dann zusammen, wenn $\left| \begin{array}{cc} a_1 + a_3 r & a_2 + a_5 s \\ a_2 + a_5 s & a_5 t \end{array} \right| = 0$ auf l . Das ist, wie die Rechnung ergibt, gleichbedeutend mit

$$K(1 - x^2 - y^2) = 0.$$

Über $1 - x^2 - y^2 = 1 - \pi^2 - \kappa^2$ können wir durch Koordinatendrehung in (E) so verfügen, daß $0 < 1 - x^2 - y^2 < 1$.

Die charakteristischen Richtungen fallen in genau den Punkten von l zusammen, in deren Originalpunkten auf λ ds^2 verschwindende Gaußsche Krümmung hat.

Nach diesen Vorbereitungen konstruieren wir die gesuchte Fläche F in (E) . Nach der Integrationstheorie gewöhnlicher Differentialgleichungen läßt sich (II), (III)

$$d\xi^2 + d\eta^2 = ds_0^2 = E_0 du^2 + 2F_0 du dv + G_0 dv^2,$$

wegen $K_0(E_0, F_0, G_0) = 0$, durch Funktionen $\xi = \xi(u, v)$, $\eta = \eta(u, v)$ lösen, die überall da in der (u, v) -Ebene regulär sind, wo E_0, F_0, G_0

regulär sind und $E_0 G_0 - F_0^2 \neq 0$. Unser Integral $\zeta = \zeta(u, v)$ bewirkt für E_0, F_0, G_0 beide Bedingungen in der Nähe von λ , aber nicht in λ selbst; λ ist Rand des Definitionsbereichs von $\xi(u, v), \eta(u, v)$.

$$\xi = \xi(u, v),$$

$$\eta = \eta(u, v),$$

$$\zeta = \zeta(u, v)$$

bestimmen jedenfalls in (E) eine Fläche F des Linienelements (I) in einem Bereich dieses Linienelements, der an λ anschließt.

Nun muß aber noch gezeigt werden, daß λ , der Rand von F in (E) mit L zusammenfällt bzw. mit einer Kurve, die aus L durch Bewegung in der (ξ, η) -Ebene entsteht. Ich weiß dafür leider keinen einfacheren Beweis als folgenden:

L_0 sei die Projektion von L auf die (ξ, η) -Ebene; $s_0, \frac{1}{e_0}$ Bogenlänge und Krümmung von L_0 ; s die Bogenlänge von L . Dann ist die Gestalt von L auf dem Umweg über L_0 bestimmbar durch die Funktionen $\frac{1}{e_0}(s_0), s(s_0)$; $s(s_0)$ mißt nämlich die Steigung von L über L_0 , womit L festgelegt ist, wenn L_0 es ist; L_0 wird durch $\frac{1}{e_0}(s_0)$ festgelegt.

Andererseits sei $\frac{1}{e_c}(v)$ die geodätische Krümmung der Linien $u = c = \text{konst.}$ bezüglich $ds_0^2(u, v)$. Angenommen, für $c \rightarrow 0$ konvergiert $\frac{1}{e_c}$ gleichmäßig gegen eine Funktion $\frac{1}{e_0^*}(v)$. Dann mißt $\frac{1}{e_0^*}(v)$ die Krümmung der Projektion des Randes von F auf die (ξ, η) -Ebene. Damit dieser Rand die gewünschte Gestalt habe, ist notwendig $\frac{1}{e_0^*}(v) = \frac{1}{e_0^*}(s_0)$ für $v = s_0$. Ebenso muß notwendig $\frac{ds}{ds_0}$ in $u = 0$ den Wert haben, der durch (L, L_0) definierten Funktion $s(s_0)$ entspricht. Das Eintreffen dieser beiden Bedingungen ist aber auch hinreichend, denn die Gestalt des Randes von F ist durch $s(s_0), \frac{1}{e_0^*}(v)$ völlig bestimmt, sie ist also die richtige, wenn $s(s_0), \frac{1}{e_0^*}(v)$ die richtigen Funktionen sind; das muß nun gezeigt werden.

Für $\frac{ds}{ds_0}$ ist das klar, denn es ist in $u = 0$

$$\frac{ds}{ds_0} = \frac{\sqrt{G}}{\sqrt{G - \kappa^2}} = \frac{1}{\sqrt{1 - \kappa^2}}.$$

Diese Funktion hängt nur von den vorgegebenen Streifengrößen ab, hat also von selber die richtigen Werte.

Für $\frac{1}{\varrho_c}(v)$ findet man nach den üblichen Regeln:

$$\begin{aligned} \frac{1}{\varrho_c(v)} &= -\frac{1}{\sqrt{E_0 G_0 - F_0^2} G_0^{\frac{3}{2}}} \left[\frac{1}{2} F_0 G_{0v} - G_0 F_{0v} + \frac{1}{2} G_0 G_{0u} \right] \\ &= \frac{1}{\sqrt{G(1-\pi^2) - \kappa^2 (G - \kappa^2)^{\frac{3}{2}}}} \left[-G\pi\tau - \frac{1}{2} G G_u + \frac{\kappa}{2} (\kappa G_u + \pi G_v) \right]. \end{aligned}$$

Das einzige Glied dieses Ausdrucks, das sich in der Nähe von λ möglicherweise nicht regulär verhält, ist $G\pi\tau$. Wir benutzen an dieser Stelle unsere

Regularisierung; durch $u \rightarrow x$ haben wir ja $\tau = \frac{r}{r^2 - s^2}$; es genügt, das Verhalten dieses Ausdrucks in der Nähe von l auf $z(x, y)$ zu untersuchen.

Auf l ist nach (VII) 4.: $r = s = 0$, also $\tau = \frac{dr}{t dr - 2s ds}$, wo dr, ds Zuwüchse in einer nicht an l tangentialen Richtung mit $z(x, y)$ sind. In § 12 werden wir zeigen, daß l nicht parabolische *Doppelkurve* auf $z(x, y)$ ist, und damit $dr \neq 0$ nachweisen. Indem wir dies Resultat vorwegnehmen, haben wir $\lim \tau = \frac{1}{t} = \frac{\pi}{\varrho_g}$:

$$\lim_{c \rightarrow 0} \frac{1}{\varrho_c(v)} = \frac{1}{\varrho_0^*(v)} = \frac{1}{\varrho_g} \frac{\sqrt{1 - \pi^2 - \kappa^2}}{(1 - \kappa^2)^{\frac{3}{2}}}.$$

Das ist aber die „richtige“ Funktion, wie man nach den Regeln des Projizierens erkennt. ($\arccos \sqrt{1 - \kappa^2}$ bzw. $\arccos \sqrt{1 - \pi^2 - \kappa^2}$ sind die Winkel der (ξ, η) -Ebene mit der Tangente der Randkurve von F bzw. der Normalen von F in der Randkurve.)

Damit ist die Existenz der gesuchten Fläche gesichert.

Wir beweisen nun: Statt u, v kann man neue Parameter α, β auf ds^2 einführen, so daß F dargestellt wird durch

$$\begin{aligned} \xi &= \varphi_1(\alpha, \beta), \\ \eta &= \varphi_2(\alpha, \beta), \\ \zeta &= \varphi_3(\alpha, \beta), \end{aligned}$$

wo $\varphi_1, \varphi_2, \varphi_3$ in λ regulär von α, β abhängen; die Singularität auf F in $\lambda = L$ wird nur dadurch bewirkt, daß die Determinante von ds^2 bezüglich α, β in $\lambda = L$ verschwindet; die Determinante tritt bekanntlich in den Nenner der Krümmungsgrößen von F . Die Singularität in L entspricht also in dieser gesuchten Darstellung genau den Pollinien analytischer Funktionen mehrerer Veränderlicher. Jedes Netz $\alpha = \text{konst.}, \beta = \text{konst.}$ der verlangten Art nennen wir im voraus ein „regularisierendes“ Netz von F ; wir beweisen dann sogar:

1. Die Krümmungslinien sind auf F ein regularisierendes Netz.

2. Die Asymptotenlinien sind auf F ein regularisierendes Netz, wenn $K \neq 0$ in λ .

3. Beide Scharen der Asymptotenlinien und eine Schar Krümmungslinien haben $\lambda = L$ zur Einhüllenden.

3. folgt leicht aus dem Drehsatz § 4. Wie schon bemerkt, entsprechen die Asymptotenrichtungen auf F den charakteristischen Richtungen auf $\zeta(u, v)$ und auf $z(x, y)$. Wir haben bewiesen, daß diese letzten von der Oskulationsrichtung in l verschieden sind. Ihnen entspricht also nach dem Drehsatz die Richtung von λ auf $\zeta(u, v)$ und auf F , womit der Satz für die Asymptotenlinien bewiesen ist. Die Krümmungslinien halbieren die Winkel der Asymptotenlinien; eine Schar muß also λ berühren, eine auf λ senkrecht stehen.

Um 2. zu beweisen, gehen wir von § 8 aus. Die Fläche $\zeta = \zeta(u, v)$ im u -Raum wird in λ mittels $u \rightarrow x$ in regulärer Parameterform dargestellt:

$$(IX) \quad \begin{aligned} u &= f_1(x, y), \\ v &= f_2(x, y), \\ \zeta &= f_3(x, y). \end{aligned}$$

Wir nehmen nun an, $K \neq 0$ in λ, l . Dann fallen die charakteristischen Richtungen auf $z(x, y)$ in l nicht zusammen, wie wir bewiesen haben. Sind also α, β passende Parameter des Charakteristikennetzes auf $z(x, y)$, so kann man in l x, y durch α, β umkehrbar regulär ausdrücken: $x = x(\alpha, \beta)$, $y = y(\alpha, \beta)$. (IX) geht über in

$$(X) \quad \begin{aligned} 1. \quad &u = g_1(\alpha, \beta), \\ 2. \quad &v = g_2(\alpha, \beta), \\ 3. \quad &\zeta = g_3(\alpha, \beta), \end{aligned}$$

wo g_1, g_2, g_3 in λ regulär sind. Eine Koordinate in (E) , ζ , hängt nach (X) 3. regulär von α, β ab; also aus Symmetriegründen auch ξ, η , wenn man schlimmstenfalls α, β durch passende Funktionen $\bar{\alpha}(\alpha), \bar{\beta}(\beta)$ ersetzt; denn $\alpha = \text{konst.}$, $\beta = \text{konst.}$ ist, wie wiederholt bemerkt, das Netz der Asymptotenlinien auf F , also etwas von den Koordinaten Unabhängiges.

1. ist offenbar völlig entsprechend 2. bewiesen, wenn man von dem Bildnetz $\alpha' = \text{konst.}$ $\beta' = \text{konst.}$ der Krümmungslinien von F auf $z(x, y)$ weiß, daß seine Richtungen in l nicht zusammenfallen. Da die Krümmungsrichtungen auf F nicht zusammenfallen, könnten es nach dem Drehsatz ihre Bildrichtungen d_1, d_2 auf $z(x, y)$ in l nur tun, wenn sie außerdem beide in die Oskulationsrichtung d fielen. Die Krümmungsrichtungen auf F liegen harmonisch zu den Asymptotenrichtungen. Also liegen d_1, d_2 auf $z(x, y)$ harmonisch zu den charakteristischen Richtungen. Fielen daher d_1, d_2 mit d zusammen, so müßte auch eine charakteristische Richtung mit d zusammenfallen, was nicht der Fall ist.

§ 12.

Wir untersuchen jetzt, ob λ auf $\zeta(u, v)$ im u -Raum stets Rückkehrkante ist, oder auch „unsichtbar“ sein kann (§ 3). Das Ergebnis haben wir auf die Geometrie von F in (E) anzuwenden, wodurch einige weitere Sätze herauskommen werden.

Wenn λ „unsichtbar“ auf $\zeta(u, v)$ sein soll, muß nach § 4 l auf $z(x, y)$ $2n$ -fache parabolische Kurve sein, also mindestens Doppelkurve; d. h. außer $rt - s^2$ müssen auf l auch $\frac{\partial}{\partial x}(rt - s^2)$, $\frac{\partial}{\partial y}(rt - s^2)$ verschwinden. Wir beweisen nun, daß das nie eintritt, wenn nur λ , wie wir voraussetzen, gleichsinnig geodätisch gekrümmt ist.

Zur Abkürzung setzen wir für die dritten Ableitungen von $z(x, y)$:

$$r_x = \alpha_1; \quad r_y = s_x = \alpha_2; \quad s_y = t_x = \alpha_3, \quad t_y = \alpha_4;$$

also:

$$\frac{\partial}{\partial x}(rt - s^2) = \alpha_1 t + r \alpha_3 - 2s \alpha_2.$$

Längs l wird das:

$$\frac{\partial}{\partial x}(rt - s^2) = \alpha_1 t.$$

Auf l ist $t \neq 0$ nach (VII) 5.; wir haben also nur $\alpha_1 \neq 0$ auf l nachzuweisen (damit wird offenbar auch die in § 11 gemachte Annahme $dr \neq 0$ bewiesen sein). Indem wir (VII) 1. nach x ableiten und $r = s = \dot{r} = \dot{s} = 0$ auf l benutzen, erhalten wir für $\alpha_1, \alpha_2, \alpha_3$ auf l drei lineare Gleichungen:

$$\begin{aligned} \text{(XI)} \quad & 1. \quad \dot{x} \alpha_1 + \dot{y} \alpha_2 = 0 (= \dot{r}), \\ & 2. \quad \dot{x} \alpha_2 + \dot{y} \alpha_3 = 0 (= \dot{s}), \\ & 3. \quad \alpha_3 t \alpha_1 - 2\alpha_2 \alpha_3 + \alpha_1 \alpha_3 = -t(a_1)_x - (a_4)_x. \end{aligned}$$

Die Determinante D der α_i in (XI) ist, wie aus der allgemeinen Theorie der DFGL. bekannt, die linke Seite der charakteristischen Gleichung von (VII) 1.; daher ist $D \neq 0$ auf l ; denn l ist nichtcharakteristisch auf $z(x, y)$, weil λ nichtcharakteristisch auf $\zeta(u, v)$ ist (§ 6). Also:

$$\alpha_1 = \frac{-\dot{y}^2}{\alpha_3 t \dot{y}^2 + 2\alpha_2 \dot{x} \dot{y} + \alpha_1 \dot{x}^2} [t(a_1)_x + (a_4)_x].$$

Um $(a_1)_x, (a_4)_x$ zu berechnen, ist eine Bemerkung nützlich, die auch später gebraucht wird. Es ist allgemein für $i = 1, \dots, 5$:

$$(a_i)_x = \frac{\partial a_i}{\partial u} \cdot \frac{\partial u}{\partial x} + \frac{\partial a_i}{\partial v} \cdot \frac{\partial v}{\partial x} + \frac{\partial a_i}{\partial \pi} \cdot \frac{\partial \pi}{\partial x} + \frac{\partial a_i}{\partial \kappa} \cdot \frac{\partial \kappa}{\partial x}.$$

Nun ist aber $\frac{\partial \pi}{\partial x} = \frac{\partial x}{\partial x} = 1$, $\frac{\partial \kappa}{\partial x} = \frac{\partial y}{\partial x} = 0$, und auf l auch $\frac{\partial u}{\partial x} = \frac{\partial p}{\partial x} = r = 0$,

$\frac{\partial v}{\partial x} = \frac{\partial q}{\partial x} = s = 0$; also einfach:

$$(a_i)_x = \frac{\partial}{\partial \pi} a_i(u, v, \pi, \kappa).$$

Demnach findet man aus (IV): $(a_4)_x = 0$, $(a_1)_x = -\frac{1}{e_g}$. Also nach (VII) wegen $t = \frac{1}{y} = -\frac{1}{a_1}$:

$$(XII) \quad \alpha_1 = \frac{1}{e_g} \cdot \frac{1}{a_5 - 2a_2 \dot{x} - \dot{x}^2} = \frac{1}{e_g} \cdot \frac{1}{N}.$$

Zunächst folgt daraus $\alpha_1 \neq 0$, was zu beweisen war. λ ist Rückkehrkante auf $\zeta(u, v)$. Projiziert man daher im u -Raum $\zeta = \zeta(u, v)$ auf die (u, v) -Ebene, so gibt λ ($u = 0$) eine Randkurve λ_0 der Projektion. Eine Seite von λ_0 bleibt unbedeckt, eine wird doppelt bedeckt. Daraus folgt:

Die Fläche F in (E) hat in $\lambda = L$ eine Rückkehrkante; beide in $\lambda = L$ zusammenstoßenden Blätter von F verwirklichen bez. ds^2 nur eine und dieselbe Seite von λ .

Das ist die Verallgemeinerung einer bei den Torsen bekannten Erscheinung. Sei etwa Λ_0 ein gleichsinnig gekrümmtes Kurvenstück einer Ebene Φ_0 . Φ sei eine Torse mit der Rückkehrkante Λ . Λ sei in Φ gemessen kongruent Λ_0 in Φ_0 . Dann bildet Φ den Teil von Φ_0 zweimal isometrisch ab, der auf der Außenseite¹⁸⁾ von Λ_0 liegt. Der Innenseite von Λ_0 auf Φ_0 entspricht kein (reeller) Punkt von Φ .

Daß die Torse immer gerade auf der Außenseite ihrer Rückkehrkurve liegt (in der Torse gemessen), ist wiederum Sonderfall eines allgemeinen Satzes. Er lautet (natürlich im Reellen gemeint):

1. Ist auf $\lambda = L$: $K \geq 0$, so liegen beide Blätter von F auf der Außenseite von λ (in F gemessen).
2. Ist auf $\lambda = L$: $K < 0$, $\frac{1}{r^2} + K > 0$, so liegen beide Blätter von F auf der Außenseite von λ .
3. Ist auf $\lambda = L$: $K < 0$, $\frac{1}{r^2} + K < 0$, so liegen beide Blätter von F auf der Innenseite von λ .

Unanschaulich, aber kürzer kann man sagen: F liegt auf der Außenseite von λ , wenn $\frac{1}{r^2} + K > 0$, auf der Innenseite, wenn $\frac{1}{r^2} + K < 0$. $\frac{1}{r^2} + K = 0$ fällt aus, denn das ist die Ennepersche Gleichung, die ja n. V. (§ 9) auf $\lambda = L$ nicht gelten soll.

Wir beweisen den Satz in letztgenannter Form. Er folgt unmittelbar aus (XII). Es ist nämlich $\alpha_1 = r_x = u_{xx}$. In λ ist aber $u = 0$, $u_x = r = 0$.

¹⁸⁾ „Außenseite“ einer gleichsinnig geodätisch gekrümmten Kurve in einer Fläche nennen wir die Seite, die von den geodätischen Tangenten der Kurve überstrichen wird; die andere also Innenseite. Diese Unterscheidung wird gewöhnlich durch „konvex“, „konkav“ bezeichnet, deren Bedeutungen aber bei verschiedenen Autoren miteinander vertauscht werden.

Daher in der Nähe von λ :

$$(XIII) \quad \text{sign } u = \text{sign } \alpha_1.$$

Bekanntlich bezeichnet stets

$$(XIV) \quad u > 0 \text{ die } \begin{cases} \text{Außenseite} \\ \text{Innenseite} \end{cases} \text{ von } u = 0, \text{ wenn } \begin{cases} \frac{1}{\rho_g} < 0 \\ \frac{1}{\rho_g} > 0 \end{cases} \text{ auf } u = 0.$$

Damit ist alles auf das Vorzeichen von α_1 in (XII) zurückgeführt. Nun ist aber für $\pi \rightarrow 0$, $\varkappa \rightarrow 0$:

$$N \rightarrow -\frac{1}{\tau^2} - K.$$

Also ist wegen der Invarianz der hier erörterten Fragen auch für beliebige π , \varkappa :

$$(XV) \quad \begin{array}{ll} N > 0 & \text{für } \frac{1}{\tau^2} + K < 0, \\ N < 0 & \text{für } \frac{1}{\tau^2} + K > 0. \end{array}$$

(XII), (XIII), (XIV), (XV) beweisen alles Verlangte.

Die bekannten Drehflächen positiver und negativer fester Gaußscher Krümmung geben Beispiele des eben bewiesenen Satzes. Die auftretenden singulären Streifen sind *ebene* Streifen, und in der Tat liegen die Flächen $K > 0$ auf der Außenseite, die $K < 0$ auf der Innenseite ihrer Ränder. Gleichzeitig erklärt unsere Betrachtung die zunächst überraschende Erscheinung, daß längs der singulären Streifen immer zwei Exemplare jener Drehflächen aneinanderkleben. Sie sind eben die beiden Blätter eines und desselben Integrals $\zeta(u, v)$ der Biegungsgleichung.

Wenden wir unsere Ergebnisse auf *stetige* Biegungsvorgänge an. Denken wir uns z. B. die oben erwähnte Ebene Φ_0 stetig in die Torse Φ verbogen, so daß $\Lambda_0 \rightarrow \Lambda$. Dann bekommt Φ_0 offenbar auf der Innenseite von Λ_0 einen Riß, der bei der Verbiegung vorrückt; in der Endlage $\Phi_0 \sim \Phi$ ist die ganze Innenseite von $\Lambda_0 \sim \Lambda$ weggerissen.

Ganz dasselbe tritt bei stetiger Verbiegung beliebig gekrümmter Flächenstücke F auf. Dem Biegungsvorgang entspricht eine Folge von Integralen $\zeta(u, v)$ bzw. $z(x, y)$ von (IV) 1. bzw. (VII) 1. Der stetigen Folge der parabolischen Kurven auf den $z(x, y)$ entspricht ein vorrückender Riß auf den F .

Nur kurz sei angedeutet, was eintritt, wenn man, anders als bisher, *Wendepunkte* auf λ zuläßt. Um in diesem Fall durch Legendre-Transformationen regularisieren zu können, muß man vorher v durch $\bar{v}(v)$ ersetzen, so daß für die neuen Koordinaten im Wendepunkt von λ :

$\frac{\partial G}{\partial v} \neq 0$; $\alpha_1 \neq 0$. Regularisiert man dann durch $u \rightarrow x$, so erhält man wieder eine Fläche mit einer parabolischen Kurve l als Bild von λ ; dem Wendepunkt von λ entspricht ein Punkt auf l , durch den außer l noch ein zweiter parabolischer Kurvenzug von $z(x, y)$ läuft; man zeigt das, indem man das Vorzeichen von $rt - s^2$ zu beiden Seiten von λ verfolgt. Die Lösungsfläche in (E) hat demnach zwei Rückkehrkanten, die im Wendepunkt von λ zusammenstoßen.

Der Leser betrachte diese Verhältnisse am Beispiel der Torsen, deren Rückkehrkante einen Wendepunkt hat. Die Tangente der Rückkehrkante im Wendepunkt ist dann selbst zweite Rückkehrkante der Torse, die Ebene wird in der Nähe des Wendepunktes vierblättrig isometrisch abgebildet usw.

Durch stetige Verbiegung kann man eine Flächenkurve mit geodätischem Wendepunkt nicht in ihrer Gesamtlänge zur singulären Linie machen; die Fläche bekommt bei der Verbiegung einen Riß, der durch jenen Wendepunkt hindurchgeht.

§ 13.

Wir beschränken uns jetzt wieder auf gleichsinnig geodätisch gekrümmte λ .

Die beiden Hauptkrümmungen von F in (E) werden ∞ und 0 , wie man überlegen oder nachrechnen mag. Man kann aber, wie jetzt gezeigt werden soll, „reduzierte“ Hauptkrümmungen auf F nahe λ einführen, die aus den ursprünglichen durch Multiplikation mit einfachen Funktionen hervorgehen, die in λ 0 bzw. ∞ werden; diese reduzierten Hauptkrümmungen sind in der Nähe der Rückkehrkante endlich und stetig und erlauben in die Gestalt der Fläche ebensolche Einsicht wie die gewöhnlichen Hauptkrümmungen anderswo.

P sei ein Punkt nahe $\lambda = L$ auf F . s sei die in F gemessene Entfernung der Rückkehrkante von P . $\frac{1}{R_1}$ sei die größere, $\frac{1}{R_2}$ die kleinere Hauptkrümmung von F in P .¹⁹⁾ Dann beweisen wir die merkwürdige Formel:

$$(XVI) \quad \lim_{s \rightarrow 0} \frac{2s}{R_1^2} = \varrho_g \left(\frac{1}{r^2} + K \right); \quad \lim_{s \rightarrow 0} \frac{1}{2s R_2^2} = \frac{K^2}{\varrho_g \left(\frac{1}{r^2} + K \right)}.$$

Beweis. Wenn man in $z(x, y)$ auf irgendeinem regulären Weg in einen Punkt von l geht, ist, wie leicht nachzurechnen, stets

$$\lim \frac{pt^2}{(rt - s^2)^2} = -\frac{1}{2\alpha_1};$$

¹⁹⁾ Nabelpunkte gibt es auf F nahe $\lambda = L$ nicht.

ferner trivialerweise:

$$\lim \frac{ps^2}{(rt-s^2)^2} = \lim \frac{pr^2}{(rt-s^2)^2} = 0.$$

Dementsprechend ist auf F , wenn man auf irgendeinem regulären Weg in einen Punkt von λ geht (also $u \rightarrow 0$):

$$\lim_{u \rightarrow 0} u \varrho^2 = -\frac{1}{2\alpha_1},$$

$$\lim_{u \rightarrow 0} u \sigma^2 = \lim_{u \rightarrow 0} u \tau^2 = 0.$$

u ist nach Definition von (I) die geodätische Entfernung des auf F wandernden Punktes von der Kurve λ ($u=0$). Wenn also $u = u(t)$, $v = v(t)$ der betrachtete Weg in F nach λ ist, hat man:

$$(XVII) \quad \lim_{s \rightarrow 0} \sqrt{s} (\varrho \dot{u}^2 + 2\sigma \dot{u} \dot{v} + \tau \dot{v}^2) = \sqrt{\frac{1}{-2\alpha_1}} \lim \dot{u}^2.$$

Aus (XVII) leiten wir (XVI) durch den Schluß ab, den wir schon mehrfach anwandten; wir drücken beide Seiten von (XVII) für $\pi = \kappa = 0$ durch invariante geometrische Größen aus; die so entstehende Gleichung gilt dann für beliebige π, κ .

Sei t die Bogenlänge des betrachteten Weges auf F , also $\dot{u}^2 + \dot{v}^2 G = 1$. Dann wird (XVII) für $\pi = \kappa = 0$, wie leicht nach (XII) zu verifizieren:

$$(XVI^*) \quad \lim_{s \rightarrow 0} \sqrt{2s} \frac{1}{R} = \dot{u}_0^2 \sqrt{\varrho_g \left(\frac{1}{r^2} + K \right)} \quad (\dot{u}_0 = \lim_{u \rightarrow 0} \dot{u} \text{ gesetzt}).$$

Dabei ist $\frac{1}{R}$ die Normalkrümmung von F im wandernden Punkt in der Wegrichtung.

In (XVI*) ist aber (XVI) enthalten. Das Maximum von $\left| \frac{1}{R} \right|$ ist ja die größere Hauptkrümmung; dies erhält man für $\dot{u}_0 = \pm 1$, $\dot{v}_0 = 0$. Damit ist die erste Formel von (XVI) bewiesen, die zweite folgt aus ihr wegen $\frac{1}{R_1 R_2} = K$.

Ferner ist der Satz aus § 11 erneut bewiesen, daß die eine Schar Krümmungslinien λ senkrecht trifft; denn $\dot{u}_0 = \pm 1$, $\dot{v}_0 = 0$ ist, wie eben gezeigt, Krümmungsrichtung in u ; $\dot{u}_0 = \pm 1$, $\dot{v}_0 = 0$ steht aber nach (I) auf $u = \text{konst.}$ senkrecht. Nun wissen wir auch (was aus § 11 nicht folgt), daß jene Schar zur größeren Hauptkrümmung gehört; das hätte man aber auch mit einfacheren Mitteln beweisen können.

Bei veränderlichem \dot{u}_0, \dot{v}_0 ist (XVI*) das Analogon zu der bekannten Eulerschen Formel $\frac{1}{R} = \frac{\cos^2 \varphi}{R_1} + \frac{\sin^2 \varphi}{R_2}$.

$\frac{1}{r}$, die Windung der Rückkehrkante, ist das einzige, was auf den rechten Seiten von (XVI) nicht von der inneren Geometrie von ds, λ , sondern von der räumlichen Gestalt der Rückkehrkante abhängt. Man kann daher eine anschauliche Anwendung von (XVI) auf stetige Biegungsvorgänge machen, indem man fragt, wie sich F ändert, wenn man L „verwindet“, d. h. die Windung von L bei gleichbleibender Bogenlänge und Krümmung stetig ändert; nach § 9 bleibt L dabei Rückkehrkante von F , wenn nur die Windung nicht den Enneperschen Wert hat. Nach (XVI) biegt F bei wachsender Windung von L immer schärfer vom Streifen L ab. Besonders schön sieht man das an Modellen von Torsen. Verwindet man deren Rückkehrkante stärker, so spreiten die beiden Blätter der Torse sich auseinander, umgekehrt falten sie sich ja zu einem Stück Doppellebene zusammen, wenn man die Windung der Rückkehrkante zum Verschwinden bringt.

(XVI) schlägt ferner eine Brücke zur Theorie der Kegelspitzen, deren systematische Betrachtung in dieser Arbeit unterblieben ist. Man betrachte zunächst eine Folge von Torsen mit Rückkehrkanten, deren isometrische Bilder in der Ebene konzentrische Kreise von abnehmendem Radius sind; läßt man die Windung dieser Rückkehrkanten verkehrt proportional zum Kreisradius wachsen, so konvergieren die Torsen bei verschwindendem Radius gegen einen Kegel; die Windung der Torsen geht in die Seitenkrümmung des Kegels über. Dasselbe gelingt nun im allgemeinen Fall, wie hier nur angedeutet sei; man hat die Folge der Vorgaben λ, L bei gleichbleibendem ds^2 genau wie bei den Torsen zu wählen. Die Konvergenz der Flächenfolge beweist man auf dem Umweg über die zugehörigen $z(x, y)$. Man erhält dann aus (XVI) folgende Sätze:

1. *Die Krümmungslinien auf F in der Umgebung einer Kegelspitze S verhalten sich wie die Meridiane und Breitenkreise gewöhnlicher Drehkegel.*

2. *Zu den Meridianen gehört die kleinere Hauptkrümmung $\frac{1}{R_2}$ auf F . Bezeichnet s die Entfernung eines Punktes auf F von S , $\frac{1}{T}$ die Seitenkrümmung des Kegels s , so ist*

$$(XVI^{**}) \quad \lim_{s \rightarrow 0} \frac{1}{s R_2} = \sqrt{2} K T.$$

Man sieht, daß *Wendelinien* $\left(\frac{1}{T} = 0, T = \infty\right)$ des Vorgabekegels eine Ausnahmerolle spielen im Falle $K \neq 0$, im Falle $K = 0$ dagegen nicht. Die wiederholt bemerkte Analogie beliebiger Flächen zu den abwickelbaren besteht also an dieser Stelle nicht.

Die in § 10 bewiesene Sonderstellung geodätischer Vorgaben ($\varrho_g \equiv \infty$) kommt in (XVI) wieder zum Ausdruck, ebenso die Sonderstellung der Wendepunkte ($\varrho_g = \infty$).

§ 14.

Wir gehen nun zu dem Problem über, das am Schluß von § 9 angedeutet wurde: Auf einem Vorgabestreifen (λ, L) liegt ein Punkt dritter Art auf einem Bogen sonst erster Art. Dann kann man nach Analogie zur Physik fragen, ob die so vorgegebene Singularität sich fortpflanzt oder nicht, d. h. ob durch den Punkt dritter Art eine Rückkehrkante der Lösungsfläche läuft, oder ob der Punkt (bei geeigneter Wahl der Vorgabe) *isolierte* Singularität auf der Lösungsfläche sein kann; wir stellen also nun losgelöst von der Vorstellungsweise des Cauchy-Problems die Frage, ob Flächen regulären Linienelements *isolierte* singuläre Punkte haben können. Dabei sind natürlich nicht Kegelspitzen gemeint (die als entartete Kurvenstreifen aufzufassen sind), sondern Punkte, in denen die Krümmungsgrößen unstetig werden, die Tangentialebene sich dagegen stetig verhält.

Die Antwort lautet:

Punkte $K < 0$ können isolierte Singularitäten sein, Punkte $K > 0$ dagegen nicht.

Der Beweis dieser Einzelfrage soll ausführlich dargestellt werden. Er läßt sich nämlich von der Biegungsgleichung auf Monge-Ampèresche Gleichungen viel allgemeineren Typs übertragen; die Unterscheidung $K \geq 0$ entspricht der Unterscheidung, ob die Charakteristiken der Fläche in der Umgebung der fraglichen Punkte reell oder imaginär sind. Der Fall $K = 0$ bzw. zusammenfallender Charakteristiken läßt sich nicht so leicht erledigen und bleibt deshalb hier unerörtert.

Wir geben den Beweis gleich für die allgemeinere Form des Problems (unter möglicher Beibehaltung der Bezeichnungen), damit das Wesentliche nicht durch die Besonderheiten der Biegungsgleichung verdeckt wird. Der Leser überzeuge sich, daß alle Voraussetzungen, die im folgenden gemacht werden, für die Biegungsgleichung erfüllt sind.

Es sei die DFGL. vorgelegt:

$$(XVIII) \quad \begin{aligned} 1. \quad & a_1 \rho + 2 a_2 \sigma + a_3 \tau + a_4 (\rho \tau - \sigma^2) + a_5 = 0, \\ 2. \quad & a_1 t - 2 a_2 s + a_3 r + a_4 (r t - s^2) + a_5 = 0, \end{aligned}$$

wo 2. aus 1. durch die Legendre-Transformation $u \rightarrow x$ hervorgeht und nur der Deutlichkeit halber noch einmal angeführt ist. Wir untersuchen, ob durch ein Flächenelement $\Pi_{(u, v, \xi; \pi, \kappa)}$ des u -Raums ein Integral von (XVIII) 1. geht, das in Π einen isolierten singulären Punkt hat. Von den a_i setzen wir in Π voraus:

- (XIX)
1. $\frac{\partial a_4}{\partial \pi} = \frac{\partial a_4}{\partial \kappa} = 0, \quad a_4 \neq 0,$
 2. nicht $a_1 = a_2 = a_3 = 0,$
 3. a_1, a_2, a_3 linear in $\pi, \kappa,$
 4. $\frac{\partial a_i}{\partial \xi} = 0 \quad (i = 1, \dots, 5).$

Das Element P sei das Bild von Π mittels $u \rightarrow x$. Dann ist P isolierter parabolischer Punkt eines Integrals $z = z(xy)$ von (XVIII) 2, $\Delta = rt - s^2$ muß also, auf $z(xy)$ betrachtet, eine isolierte Nullstelle in P haben. Wir haben zu zeigen, daß das eintreten kann, wenn auf $z(xy)$ die Charakteristiken von (XVIII) 2. in P reell sind, dagegen nicht, wenn sie dort imaginär sind.

Die Oskulationsrichtung in P kann nicht unbestimmt sein. Denn dies würde in homogener Schreibung bedeuten $r_1 = r_2 = r_3 = 0$. Außerdem wäre aber $\Delta = 0$, also $r_4 = 0$. Dies gäbe nach (XVIII) 2., (XIX) 1. $r_5 = 0$; das geht nicht, weil nicht alle r_i gleichzeitig verschwinden können.

Durch Drehung der (x, y) -Ebene, der bekanntlich bei $x \rightarrow u$ eine Drehung der (u, v) -Ebene entspricht, können wir also erreichen, daß in Π $r = s = 0, t \neq 0$. (XIX) ist gegen solche Drehungen invariant; wir dürfen daher der Kürze halber annehmen, daß die Koordinatensysteme in u, x von vornherein die gewünschte Orientierung haben. (XIX) 2. ist dann offenbar zu ersetzen durch $a_1 \neq 0$. Daß P isolierter parabolischer Punkt ist, schreibt sich jetzt einfach (wenn wir die vierten Ableitungen von z mit

$\beta_1 = \alpha_{1x}, \beta_2 = \alpha_{2x} = \alpha_{1y}, \beta_3 = \alpha_{3x} = \alpha_{2y}, \beta_4 = \alpha_{4x} = \alpha_{3y}, \beta_5 = \alpha_{4y}$ bezeichnen):

- (XX)
1. $r = s = 0, \quad a_1 \neq 0, \quad a_4 \neq 0, \quad t = -\frac{a_4}{a_1} \neq 0.$
 2. $\Delta_x = \Delta_y = \alpha_1 = \alpha_2 = 0,$
 3. $\Delta_{xx}\Delta_{yy} - \Delta_{xy}^2 = D = (\beta_1\beta_3 - \beta_2^2)t^2 - 2\beta_1t\alpha_3^2 \geq 0.$

Um α_3 und eine Gleichung zwischen $\beta_1, \beta_2, \beta_3$ zu erhalten, leiten wir (XVIII) 2. zweimal nach x ab; die auftretenden totalen Ableitungen $(a_i)_x, (a_i)_{xx}$ kann man dann, wie in § 12, durch die partiellen a_{ix}, a_{ixx} ersetzen, wie der Leser überlegen möge. Man erhält wegen (XIX), (XX):

- (XXI)
1. $a_1\alpha_3 + a_{1x}t = 0,$
 2. $a_1\beta_3 - 2a_2\beta_2 + (a_3 + a_3t)\beta_1 = -2a_3a_{1x}.$

Dies gibt nach einiger Umformung:

(XXII)
$$D = -t^2 \left[\frac{a_1 a_3 - a_4 a_5}{a_1^2} \beta_1^2 - 2 \frac{a_2}{a_1} \beta_1 \beta_2 + \beta_2^2 \right].$$

D ist eine quadratische Form in β_1, β_2 ; ihre Determinante ist bis auf einen positiven Faktor die Determinante der charakteristischen Gleichung von (XVIII) 2. in P . Sind also die Charakteristiken von (XVIII) 2. in P auf $z(xy)$ reell und voneinander verschieden, so ist D indefinit; durch geeignete Wahl von β_1, β_2 erreicht man daher $D > 0$ in P . $\beta_3, \beta_4, \beta_5$ werden dann in P durch (XXI) 2. und die andern beiden zweiten Ableitungen von (XVIII) 2. gegeben. Alle Integrale von (XVIII) 2., die in P diese $p, q; rst; \alpha_1 \dots \alpha_4; \beta_1 \dots \beta_5$ haben, besitzen P zum isolierten parabolischen Punkt. Sie sind willkürlich von den fünften Ableitungen an. Ihre Bilder im u -Raum haben in Π eine isolierte Singularität. Damit ist der erste Teil unseres Satzes bewiesen.

Sind die Charakteristiken in P aber imaginär, so ist D negativ definit. Wäre P dann isolierter parabolischer Punkt auf $z(xy)$, so müßte in P jedenfalls $\beta_1 = \beta_2 = 0$ sein, [und $D = 0$. In hinreichender Nähe von P müßte aber D positiv sein, wie aus den Elementen der Extremumstheorie folgt.

Es läßt sich nun beweisen, daß D um P in Wirklichkeit negativ ist, wenn man die Integralfläche als analytisch regulär in P voraussetzt; da das Integral in P elliptischen Charakter hat (imaginäre Charakteristiken), so ist diese Einschränkung nicht unnatürlich.

Nach dieser Annahme können nicht alle höheren Ableitungen von β_1, β_2 nach x verschwinden, wenn P isolierter parabolischer Punkt sein soll. Es gibt also ein $n \geq 1$, so daß

$$\beta_1^{(k)} = \beta_2^{(k)} = 0 \quad \text{für } k \leq n-1 \quad \left(\beta_1^{(k)} = \frac{\partial^{(k)} \beta_1}{\partial x^k} \text{ usw.} \right),$$

aber nicht

$$\beta_1^{(n)} = \beta_2^{(n)} = 0.$$

Eine nicht schwierige, aber etwas umständliche Rechnung zeigt nun, daß dann in P :

$$D^{(k)} = 0 \quad \text{für } k \leq 2n-1,$$

$$D^{(2n)} = -2^n t^2 \left[\frac{a_1 a_3 - a_4 a_5}{a_1^2} (\beta_1^{(n)})^2 - 2 \frac{a_2}{a_1} \beta_1^{(n)} \beta_2^{(n)} + (\beta_2^{(n)})^2 \right].$$

Bis auf den positiven Faktor 2^n (der als Polynomkoeffizient hereinkommt) ist $D^{(2n)}$ bez. $\beta_1^{(n)}, \beta_2^{(n)}$ dieselbe quadratische Form wie D bez. β_1, β_2 . Also $D^{(2n)} < 0$, da nicht $\beta_1^{(n)} = \beta_2^{(n)} = 0$.

Hieraus folgt nach dem Mittelwertsatz, daß D auf $z(x, y)$ zu beiden Seiten von P negativ ist in der Flächenrichtung $\dot{y} = 0$; P ist also nicht isolierter parabolischer Punkt, was zu beweisen war.

Schlußbemerkung.

Die in dieser Arbeit gegebene Erweiterung des Integralbegriffs liegt in folgender Richtung:

Die Cauchysche Definition geht von der Vorgabe aus und erzeugt das Integral durch Differentiationen in dieser. Hier dagegen treten Integrale auf, die man in der Vorgabe gar nicht differenzieren kann, die aber mit hinreichender Stetigkeit gegen die Vorgabe konvergieren. Wenn man — ganz abgesehen von unserer Regularisierungsmethode — überhaupt als Lösung des Cauchyschen Problems jedes Integral ansieht, das in irgendeinem näher festzusetzenden Sinn gegen die Vorgabe konvergiert, versagt der Cauchy-Kowalewskasche Eindeigkeitssatz; vielleicht gehen durch Streifen erster Art außer den analytischen Integralen des Eindeigkeitssatzes noch unendlich viele Integrale jener erweiterten Bedeutung; natürlich müßten solche Integrale in der Nähe der Vorgabe wesentlich stärker singularär werden, als die in dieser Arbeit untersuchten.

Für *gewöhnliche* Differentialgleichungen ist jene erweiterte Integraldefinition bekanntlich bereits untersucht; wenn die Lipschitz-Bedingung erfüllt ist, dann ist das übliche Integral auch bei der erweiterten Konkurrenz das einzige eines regulären Anfangswertproblems. Die entsprechende Untersuchung für *partielle* Differentialgleichungen steht meines Wissens noch aus.

(Eingegangen am 4. 3. 1927.)

Über die Summen durch den Zufall bestimmter unabhängiger Größen.

Von

A. Kolmogoroff in Moskau.

§ 1.

Bezeichnungen.

$\mathfrak{B}(e)$ = die Wahrscheinlichkeit des Ereignisses e .

$\mathfrak{B}_{e_1}(e_2)$ = die relative Wahrscheinlichkeit von e_2 in bezug auf e_1 .

$\mathfrak{D}(y)$ = die mathematische Erwartung der Größe y .

$\mathfrak{D}_e(y)$ = die mathematische Erwartung der Größe y in bezug auf die Hypothese e .

$\mathfrak{S}_e(y) = \mathfrak{B}(e) \mathfrak{D}_e(y)$.

Die Rechnungsregeln für das letzte Symbol sind analog denjenigen für über Mengen erstreckte Integrale. Zum Beispiel gilt, wenn die Ereignisse e_1 und e_2 einander ausschließen, die Formel

$$\mathfrak{S}_{e_1+e_2}(y) = \mathfrak{S}_{e_1}(y) + \mathfrak{S}_{e_2}(y).$$

Bezeichnen wir noch mit 1 ein notwendig eintretendes Ereignis, so gilt offenbar

$$\mathfrak{D}(y) = \mathfrak{S}_1(y).$$

Für die endliche Folge von Größen

$$y_1 \ y_2 \ \dots \ y_n$$

setzen wir

$$s_k = \sum_{i=1}^k y_i, \quad s = s_n;$$

$$v = \text{Max } |s_k|;$$

$$z_k = y_k - \mathfrak{D}(y_k);$$

$$t_k = \sum_{i=1}^k z_i, \quad t = t_n;$$

$$u = \text{Max } |t_k|;$$

$$M = \text{Max } |z_k|;$$

$$D^2 = \mathfrak{D}(t^2).$$

Dabei ist vorausgesetzt, daß $\mathfrak{D}(y_k)$ und $\mathfrak{D}(z_k^2)$ für jedes k existieren. Offenbar ist

$$\mathfrak{D}(z_k) = 0;$$

$$t_k = s_k - \mathfrak{D}(s_k).$$

Im Falle der gegenseitigen Unabhängigkeit der Größen y_k gilt außerdem

$$\mathfrak{D}(t_k^2) = \sum_{i=1}^k \mathfrak{D}(z_i^2).$$

§ 2.

Einige Sätze über finite Summen.

Im folgenden setzen wir die Größen y_k gegenseitig unabhängig voraus.

Satz I. R sei eine positive Zahl. Dann ist

$$\mathfrak{B}(u \geq R) \leq \frac{D^2}{R^2}.$$

Satz II. e sei eine positive, m eine natürliche Zahl und

$$R = eD + M.$$

Dann ist

$$\mathfrak{B}(u \geq mR) \leq \frac{1}{e^{2m}}.$$

Satz III. R sei eine positive Zahl. Dann ist

$$\mathfrak{B}(v \leq R) \leq 4 \frac{(M+R)^2}{D^2}.$$

Satz IV.

$$\mathfrak{B}(t > 0) \geq \frac{1}{16} \frac{D-M}{D+M}.$$

Satz V.

$$\mathfrak{B}(t > -3M) \geq \frac{1}{48}.$$

Satz VI.

$$\mathfrak{B}(|s| > R) \geq \frac{1}{48} \left[1 - 4 \frac{(4M+R)^2}{D^2} \right].$$

Beweis des Satzes I. Wir bezeichnen mit e_k den Fall, wo folgende Ungleichungen erfüllt sind:

$$|t_i| < R, \quad i < k;$$

$$|t_k| \geq R.$$

Offenbar ist

$$\mathfrak{B}(u \geq R) = \sum_{k=1}^n \mathfrak{B}(e_k).$$

Da die Größen z_i für $i > k$ in der Definition des Falles e_k nicht auftreten sind sie auch in der Hypothese e_k voneinander und von t_k unabhängig. Deshalb ist

$$\mathfrak{S}_{e_k}(t^2) = \mathfrak{S}_{e_k}(t_k^2) + \sum_{i=k+1}^n \mathfrak{S}_{e_k}(z_i^2) \geq \mathfrak{B}(e_k) \mathfrak{D}_{e_k}(t_k^2) \geq \mathfrak{B}(e_k) R^2.$$

Folglich:

$$D^2 = \mathfrak{S}_1(t^2) \geq \sum_{k=1}^n \mathfrak{S}_{e_k}(t^2) \geq \mathfrak{B}(u \geq R) R^2.$$

Aus der letzten Ungleichung folgt unser Satz unmittelbar.

Zusatz. Die bewiesene Ungleichung verwandelt sich in eine Gleichung im folgenden Falle:

$$n = 1,$$

$$\mathfrak{B}(z_1 = R) = \frac{D^2}{2R^2}, \quad \mathfrak{B}(z_1 = -R) = \frac{D^2}{2R^2}, \quad \mathfrak{B}(z_1 = 0) = 1 - \frac{D^2}{R^2}.$$

Beweis des Satzes II. Wir bezeichnen mit $e_{i,k}$ den Fall, wo folgende Ungleichungen erfüllt sind:

$$\begin{aligned} |t_j| &< iR, & j < k; \\ |t_k| &\geq iR. \end{aligned}$$

Wir setzen ferner

$$e_i = \sum_{k=1}^n e_{i,k}.$$

Im Falle $e_{i,k}$ haben wir:

$$|t_k| \leq iR + M.$$

Offenbar ist

$$\mathfrak{B}_{e_{i,k}}(e_{i+1}) = \mathfrak{B}_{e_{i,k}}[u \geq (i+1)R] \leq \mathfrak{B}_{e_{i,k}} \left[\text{Max} \left| \sum_{j=k+1}^p z_j \right|_{p=k+1}^n \geq eD \right].$$

Hier ist aber der Satz I anzuwenden, da die Größen z_j für $j > k$ in der Definition des Falles $e_{i,k}$ nicht auftreten. Somit erhalten wir:

$$\mathfrak{B}_{e_{i,k}}(e_{i+1}) \leq \frac{D^2}{e^2 D^2} = \frac{1}{e^2},$$

und folglich:

$$\mathfrak{B}_{e_i}(e_{i+1}) \leq \frac{1}{e^2},$$

$$\mathfrak{B}(u \geq mR) = \mathfrak{B}(e_m) = \mathfrak{B}_{e_{m-1}}(e_m) \mathfrak{B}_{e_{m-2}}(e_{m-1}) \dots \mathfrak{B}_{e_0}(e_1) \leq \frac{1}{e^{2m}}.$$

Damit ist aber unser Satz bewiesen.

Beweis des Satzes III. Wir bezeichnen mit \mathfrak{f}_k den Fall, wo $|s_i| \leq R$ für jedes $i \leq k$ gilt. Ferner setzen wir

$$\mathfrak{f} = \mathfrak{f}_n = (v \leq R);$$

$$e_k = \mathfrak{f}_{k-1} - \mathfrak{f}_k.$$

Offenbar ist

$$(1) \quad \mathfrak{B}(\mathfrak{f}) + \sum_{k=1}^n \mathfrak{B}(e_k) = 1.$$

Wir setzen

$$a_k = \mathfrak{D}_{\mathfrak{f}_k}(t_k).$$

Dann erhalten wir im Falle \mathfrak{f}_k :

$$|t_k - a_k| = |s_k - \mathfrak{D}_{\mathfrak{f}_k}(s_k)| \leq 2R;$$

$$|a_k - a_{k-1}| = |\mathfrak{D}_{\mathfrak{f}_k}(z_k)| \leq M.$$

Nun betrachten wir den Ausdruck:

$$(2) \quad \mathfrak{F}_{\mathfrak{f}_{k-1}}(t_k - a_k)^2 = \mathfrak{F}_{\mathfrak{f}_k}(t_k - a_k)^2 + \mathfrak{F}_{e_k}[t_{k-1} - a_{k-1} - (a_k - a_{k-1}) + z_k]^2 \\ \leq \mathfrak{F}_{\mathfrak{f}_k}(t_k - a_k)^2 + \mathfrak{B}(e_k) 4(R + M)^2.$$

Da die Größe z_k in die Definition des Falles \mathfrak{f}_{k-1} eingeht, haben wir andererseits:

$$(3) \quad \mathfrak{F}_{\mathfrak{f}_{k-1}}(t_k - a_k)^2 = \mathfrak{F}_{\mathfrak{f}_{k-1}}[(t_{k-1} - a_{k-1}) - (a_k - a_{k-1}) + z_k]^2 \\ = \mathfrak{F}_{\mathfrak{f}_{k-1}}[(t_{k-1} - a_{k-1})^2 + (a_k - a_{k-1})^2 + z_k^2] \\ \geq \mathfrak{F}_{\mathfrak{f}_{k-1}}(t_{k-1} - a_{k-1})^2 + \mathfrak{B}(\mathfrak{f}) \mathfrak{D}(z_k^2).$$

Vergleichen wir (2) und (3), so ergibt sich

$$\mathfrak{F}_{\mathfrak{f}_{k-1}}(t_{k-1} - a_{k-1})^2 + \mathfrak{B}(\mathfrak{f}) \mathfrak{D}(z_k^2) \\ \leq \mathfrak{F}_{\mathfrak{f}_k}(t_k - a_k)^2 + \mathfrak{B}(e_k) 4(R + M)^2.$$

Setzen wir in der erhaltenen Ungleichung $k = 0, 1, 2, \dots, n$ und addieren, so ergibt sich wegen (1):

$$\mathfrak{B}(\mathfrak{f}) D^2 \leq \mathfrak{F}_{\mathfrak{f}_n}(t_n - a_n)^2 + \sum_{k=1}^n \mathfrak{B}(e_k) 4(R + M)^2 \\ \leq \mathfrak{B}(\mathfrak{f}) 4R^2 + \sum_{k=1}^n \mathfrak{B}(e_k) 4(R + M)^2 \leq 4(R + M)^2,$$

was offenbar unseren Satz beweist.

Zusatz. Insbesondere können wir $y'_k = z_k$ setzen; dann wird $v' = u$. Deshalb gilt unser Satz auch, wenn v durch u ersetzt wird.

Beweis des Satzes IV. Wir bezeichnen mit e_m den Fall, wo

$$mR < t \leq (m+1)R,$$

$$R = 8D + M.$$

Ferner setzen wir:

$$\tilde{f}_m = \sum_{k=m}^{\infty} e_k.$$

Offenbar ist

$$\tilde{f}_0 = (t > 0).$$

Wegen Satz II erhalten wir

$$\mathfrak{D}(\tilde{f}_m) \leq \frac{1}{8^{2m}};$$

$$\begin{aligned} \mathfrak{S}_{\tilde{f}_0}(t) &= \sum_{m=0}^{\infty} \mathfrak{S}_{e_m}(t) \leq \sum_{m=0}^{\infty} (m+1) R \mathfrak{B}(e_m) \\ &= R \sum_{m=0}^{\infty} \mathfrak{B}(\tilde{f}_m) \leq R \left[\mathfrak{B}(\tilde{f}_0) + \sum_{m=1}^{\infty} \frac{1}{8^{2m}} \right] \\ &\leq R \left[\mathfrak{B}(\tilde{f}_0) + \frac{1}{32} \right]. \end{aligned}$$

Andererseits ist leicht zu beweisen, daß

$$\mathfrak{S}_{\tilde{f}_0}(t) \geq \frac{1}{2} D.$$

So ergibt sich

$$\begin{aligned} \frac{1}{2} D &\leq R \left[\mathfrak{B}(\tilde{f}_0) + \frac{1}{32} \right] = (8D + M) \left[\mathfrak{B}(\tilde{f}_0) + \frac{1}{32} \right]; \\ \mathfrak{B}(\tilde{f}_0) &\leq \frac{D}{2(8D + M)} - \frac{1}{32} \leq \frac{1}{16} \frac{D - M}{D + M}, \end{aligned}$$

w. z. b. w.

Beweis des Satzes V. Es sei

$$D \geq \frac{3}{2} M,$$

so ist wegen Satz IV:

$$\mathfrak{B}(t > 0) \geq \frac{1}{16} \frac{D - M}{D + M} \geq \frac{1}{48}.$$

Es sei

$$D < \frac{3}{2} M,$$

so ist wegen Satz I:

$$\mathfrak{B}(|t| \leq 3M) \leq \frac{D^2}{9M^2} < \frac{1}{4}.$$

In beiden Fällen ist unser Satz erfüllt.

Beweis des Satzes VI. Bezeichnen wir durch w

$$\text{Max} \left| s - \sum_{i=k}^n z_i \right|,$$

so erhalten wir wegen Satz III:

$$\mathfrak{B}(e) = \mathfrak{B}(w > R + 3M) \geq 1 - 4 \frac{(4M + R)^2}{D^2},$$

und durch die Anwendung des Satzes V auf die Summen

$$\sum_{i=k}^n z_i$$

ist leicht zu beweisen, daß

$$\mathfrak{B}_e(s > R) \leq \frac{1}{48}.$$

Daraus aber folgt unser Satz unmittelbar.

§ 3.

Konvergenz von Reihen.

Wir beweisen hier wieder die notwendige und hinreichende Bedingung für die Konvergenz einer Reihe von unabhängigen durch den Zufall bestimmten Größen, die wir in einer früheren gemeinsamen Arbeit mit A. Khintchine festgestellt haben¹⁾.

Betrachten wir zwei Folgen durch den Zufall bestimmter Größen:

$$y_1 \ y_2 \ \dots \ y_n \ \dots, \\ \bar{y}_1 \ \bar{y}_2 \ \dots \ \bar{y}_n \ \dots.$$

Wir wollen diese Folgen im Falle der Konvergenz der Reihe

$$(4) \quad \sum_{n=1}^{\infty} \mathfrak{B}(y_n \neq \bar{y}_n)$$

äquivalent nennen²⁾.

Im folgenden setzen wir die Größen jeder Reihe voneinander unabhängig, aber die entsprechenden Größen der beiden Reihen, im allgemeinen, voneinander abhängig voraus.

Satz VII. Die Wahrscheinlichkeit P der Konvergenz der Reihe

$$(5) \quad \sum_{n=1}^{\infty} y_n$$

ist gleich Eins, wenn eine der Folge $[y_n]$ äquivalente Folge $[\bar{y}_n]$ existiert, für welche die Reihen

$$(6) \quad \sum_{n=1}^{\infty} \mathfrak{D}(\bar{y}_n)$$

und

$$(7) \quad \sum_{n=1}^{\infty} \mathfrak{D}(\bar{z}_n^2)$$

konvergieren; P ist gleich Null, wenn es keine solche äquivalente Folge gibt.

Beweis der ersten Hälfte des Satzes. Wegen Satz I haben wir für jedes positive η :

$$(8) \quad \mathfrak{B} \left[\text{Max} \left| \sum_{k=n}^p \bar{z}_k \right|_{p=n}^N \geq \eta \right] \leq \frac{1}{\eta^2} \sum_{k=n}^N \mathfrak{D}(\bar{z}_k^2).$$

¹⁾ Rec. Math. de Moscou 32, 2 (1925), S. 668.

²⁾ Diese Definition gehört A. Khintchine.

Da die Reihe (7) konvergiert, gibt es für jedes positive ε ein so großes m , daß, wenn $m \leq n \leq N$ ist, der Ausdruck (8) kleiner als ε sein wird. Nach Definition³⁾ bedeutet das, daß die Wahrscheinlichkeit der Konvergenz der Reihe

$$(9) \quad \sum_{n=1}^{\infty} \bar{z}_n$$

gleich Eins ist. Da die Reihe (6) konvergiert, so gilt dasselbe für die Reihe

$$(10) \quad \sum_{n=1}^{\infty} \bar{y}_n.$$

Da endlich die Wahrscheinlichkeiten der Konvergenz zweier äquivalenter Reihen einander gleich sind, ist auch $P = 1$.

Beweis der zweiten Hälfte des Satzes. Wir setzen

$$(11) \quad \begin{aligned} \bar{y}_n &= y_n, & \text{wenn } |y_n| \leq 1; \\ \bar{y}_n &= 0, & \text{wenn } |y_n| > 1. \end{aligned}$$

Wir betrachten zuerst den Fall der Äquivalenz der Folgen $[y_n]$ und $[\bar{y}_n]$ oder, was dasselbe ist, der Konvergenz der Reihe

$$(12) \quad \sum_{n=1}^{\infty} \mathfrak{B} [|y_n| > 1].$$

In diesem Falle muß eine der Reihen (6) oder (7) divergieren. Wenn die Reihe (7) divergiert, erhalten wir, wegen Satz III, für jedes η und jedes n

$$\mathfrak{B} \left[\text{Max} \left| \sum_{k=n}^p \bar{y}_k \right|_{p=n}^N \leq \eta \right] \leq 4 \frac{(\eta + 2)^2}{\sum_{k=n}^N \mathfrak{D}(z_k^2)} \rightarrow 0$$

für $N \rightarrow \infty$. Das bedeutet, daß die Wahrscheinlichkeit der Konvergenz der Reihe (10) gleich Null ist. Wegen der Äquivalenz der Folgen $[y_n]$ und $[\bar{y}_n]$ folgt daraus, daß $P = 0$ ist. Wenn aber die Reihe (7) konvergiert, so ist, nach der bewiesenen ersten Hälfte des Satzes, die Wahrscheinlichkeit der Konvergenz der Reihe (9) gleich Eins, und wegen der Divergenz der Reihe (6) gilt auch $P = 1$.

Nun betrachten wir den Fall der Divergenz der Reihe (12). In diesem Fall erhalten wir unmittelbar, daß für jedes n

$$\mathfrak{B} [\text{Max} |y_k|_{k=n}^N \leq 1] = \prod_{k=n}^N [1 - \mathfrak{B} (|y_k| > 1)] \rightarrow 0$$

für $N \rightarrow \infty$, und daß folglich $P = 0$ ist. Damit ist unser Satz in allen möglichen Fällen bewiesen.

³⁾ Wir definieren

$$P = \lim_{\eta \rightarrow 0} \lim_{n \rightarrow \infty} \lim_{N \rightarrow \infty} \mathfrak{B} \left[\text{Max} \left| \sum_{k=n}^p y_k \right|_{p=n}^N < \eta \right].$$

Zusatz. Wenn eine den Bedingungen des Satzes genügende Folge $\{\bar{y}_n\}$ überhaupt existiert, so bestimmen insbesondere die Formeln (11) eine solche Folge.

§ 4.

Verallgemeinertes Gesetz der großen Zahlen.

Wir betrachten ein System zufälliger Größen

$$\begin{array}{cccc} y_{11} & y_{12} & \cdots & y_{1m_1} \\ y_{21} & y_{22} & \cdots & y_{2m_2} \\ \cdot & \cdot & \cdot & \cdot \\ y_{n1} & y_{n2} & \cdots & y_{nm_n} \\ \cdot & \cdot & \cdot & \cdot \end{array}$$

welches kurz mit $\|y_{nk}\|$ bezeichnet wird. Wir setzen die Größen jeder einzelnen Zeile voneinander unabhängig, die Größen verschiedener Zeilen aber im allgemeinen voneinander abhängig voraus.

Wir sagen, daß die Mittelwerte

$$\sigma_n = \frac{y_{n1} + y_{n2} + \cdots + y_{nm_n}}{m_n}$$

stabil sind, wenn es eine solche Zahlenfolge

$$d_1 \ d_2 \ \dots \ d_n \ \dots$$

gibt, daß für jedes positive η

$$\mathfrak{B}(|\sigma_n - d_n| \geq \eta) \rightarrow 0$$

für $n \rightarrow \infty$. Wir wollen nun eine notwendige und hinreichende Bedingung für eine solche Stabilität der Mittelwerte σ_n geben.

Wir werden die Systeme $\|y_{nk}\|$ und $\|\bar{y}_{nk}\|$ äquivalent nennen, wenn

$$m_n = \bar{m}_n$$

und

$$\mathfrak{B}(\sigma_n \neq \bar{\sigma}_n) \rightarrow 0$$

für $n \rightarrow \infty$. Es ist offenbar, daß äquivalente Systeme die Stabilität der Mittelwerte entweder zugleich besitzen oder zugleich nicht besitzen.

Satz VIII. Die notwendige und hinreichende Bedingung für die Stabilität der Mittelwerte σ_n besteht in der Existenz eines mit $\|y_{nk}\|$ äquivalenten Systems $\|\bar{y}_{nk}\|$, für welches

$$(13) \quad \frac{1}{m_n^2} \sum_{k=1}^{m_n} \mathfrak{D}(\bar{z}_{nk}^2) \rightarrow 0$$

für $n \rightarrow \infty$.

Beweis, daß die Bedingung hinreichend ist. Wegen Satz I haben wir für jedes positive η

$$\mathfrak{B} [|\bar{\sigma}_n - \mathfrak{D}(\bar{\sigma}_n)| \geq \eta] = \mathfrak{B} \left[\left| \sum_{k=1}^{m_n} \bar{z}_{nk} \right| \geq m_n \eta \right] \leq \frac{1}{m_n^2 \eta^2} \sum_{k=1}^{m_n} \mathfrak{D}(\bar{z}_{nk}^2).$$

Somit folgt aus (13), daß

$$\mathfrak{B} [|\bar{\sigma}_n - \mathfrak{D}(\bar{\sigma}_n)| \geq \eta] \rightarrow 0,$$

und aus der Äquivalenz der Systeme $\|y_{nk}\|$ und $\|\bar{y}_{nk}\|$, daß

$$(14) \quad \mathfrak{B} [|\sigma_n - \mathfrak{D}(\bar{\sigma}_n)| \geq \eta] \rightarrow 0.$$

Also ist unsere Bedingung hinreichend.

Beweis der Notwendigkeit der Bedingung. Wir nehmen an, daß die Mittelwerte σ_n stabil sind. Für jedes y_{nk} existiert eine Konstante y_{nk} , die folgenden Ungleichungen genügt:

$$(15) \quad \begin{aligned} \mathfrak{B} (y_{nk} > f_{nk}) &\leq \frac{1}{2}, \\ \mathfrak{B} (y_{nk} < f_{nk}) &\leq \frac{1}{2}. \end{aligned}$$

Es ist leicht zu beweisen, daß für jedes positive η

$$(16) \quad \sum_{k=1}^{m_n} \mathfrak{B} (|y_{nk} - f_{nk}| \geq m_n \eta) \rightarrow 0$$

für $n \rightarrow \infty$.

Wir setzen nun

$$(17) \quad \begin{cases} \bar{y}_{nk} = y_{nk}, & \text{wenn } |y_{nk} - f_{nk}| \leq m_n; \\ \bar{y}_{nk} = f_{nk}, & \text{wenn } |y_{nk} - f_{nk}| > m_n; \\ \bar{y}_{nk}(\eta) = y_{nk}, & \text{wenn } |y_{nk} - f_{nk}| \leq m_n \eta; \\ \bar{y}_{nk}(\eta) = f_{nk}, & \text{wenn } |y_{nk} - f_{nk}| > m_n \eta. \end{cases}$$

Wegen (16) sind die Systeme $\|y_{nk}\|$, $\|\bar{y}_{nk}\|$ und $\|\bar{y}_{nk}(\eta)\|$ äquivalent. Also sind auch die Mittelwerte $\bar{\sigma}_n$ und $\bar{\sigma}_n(\eta)$ stabil. Bemerken wir, daß

$$|\bar{z}_{nk}(\eta)| \leq 2 m_n$$

ist, so ergibt sich wegen Satz VI:

$$\begin{aligned} \mathfrak{B} [|\bar{\sigma}_n(\eta) - d_n| > \eta] &= \mathfrak{B} \left[\left| \sum_{k=1}^{m_n} \bar{y}_{nk}(\eta) - m_n d_n \right| > m_n \eta \right] \\ &\geq \frac{1}{48} \left[1 - \frac{324 m_n^2 \eta^2}{\sum_{k=1}^{m_n} \mathfrak{D}[\bar{z}_{nk}^2(\eta)]} \right]. \end{aligned}$$

Für jedes positive η konvergiert die linke Seite dieser Ungleichung gegen 0 für $n \rightarrow \infty$. Folglich erhalten wir

$$(18) \quad \limsup \frac{1}{m_n^2} \sum_{k=1}^{m_n} [\bar{z}_{nk}^2(\eta)] \leq 324 \eta^2.$$

Für $\eta \leq 1$ gilt aber

$$|\mathfrak{D}(\bar{z}_{nk}^2) - \mathfrak{D}[\bar{z}_{nk}^2(\eta)]| \leq 8m_n^2 \mathfrak{B}(|y_{nk} - f_{nk}| > m_n \eta)$$

und wegen (16) auch

$$\frac{1}{m_n^2} \sum_{k=1}^{m_n} |\mathfrak{D}(\bar{z}_{nk}^2) - \mathfrak{D}[\bar{z}_{nk}^2(\eta)]| \rightarrow 0.$$

Deswegen folgt aus (18) unsere Bedingung (13).

Zusatz. Wenn ein den Bedingungen des Satzes genügendes System $\|\bar{y}_{nk}\|$ überhaupt existiert, so bestimmen insbesondere die Formeln (15) und (17) ein solches System.

Wir sagen, daß die Mittelwerte σ_n normale Stabilität besitzen, wenn für jede positive η

$$\mathfrak{B}(|\sigma_n - \mathfrak{D}(\sigma_n)| \geq \eta) \rightarrow 0$$

für $n \rightarrow \infty$.

Satz IX. Die notwendige und hinreichende Bedingung für die normale Stabilität der Mittelwerte σ_n besteht in der Existenz eines äquivalenten Systems $\|y_{nk}\|$, für welches (13) und

$$\mathfrak{D}(\sigma_n) - \mathfrak{D}(\bar{\sigma}_n) \rightarrow 0$$

gilt.

Dieser Satz folgt unmittelbar aus der Formel (14), welche für jedes stabile System festgestellt ist.

Zusatz. Im Falle der normalen Stabilität können wir statt (15) und (17) auch

$$\begin{aligned} \bar{y}_{nk} &= y_{nk}, & \text{wenn } |z_{nk}| \leq m_n, \\ \bar{y}_{nk} &= \mathfrak{D}(y_{nk}), & \text{wenn } |z_{nk}| > m_n, \end{aligned}$$

setzen.

§ 5.

Gesetz der großen Zahlen.

Die Bedingung der Stabilität der Mittelwerte

$$\sigma_n = \frac{y_1 + y_2 + \dots + y_n}{n},$$

wo die zufälligen Größen

$$y_1 \ y_2 \ \dots \ y_n \ \dots$$

unabhängig sind, wird als ein Spezialfall der Sätze VIII und IX erhalten, wenn

$$m_n = n,$$

$$y_{nk} = y_k$$

gesetzt wird. Gewöhnlich formuliert man die normale Stabilität solcher Mittelwerte als *Gesetz der großen Zahlen*. In diesem Falle gilt folgender Satz:

Satz X. Damit für jedes positive η

$$(19) \quad \mathfrak{B} [|\sigma_n - \mathfrak{D}(\sigma_n)| > \eta]$$

für $n \rightarrow \infty$ gegen Null konvergiert, ist die Existenz eines solchen Systems $\|\bar{y}_{nk}\|$ notwendig und hinreichend, daß

0. $\bar{m}_n = n$,
1. $\sum_{k=1}^n \mathfrak{B}(\bar{y}_{nk} \neq y_k) \rightarrow 0$,
2. $\frac{1}{n} \sum_{k=1}^n [\mathfrak{D}(\bar{y}_{nk}) - \mathfrak{D}(y_k)] \rightarrow 0$,
3. $\frac{1}{n^2} \sum_{k=1}^n \mathfrak{D}(\bar{z}_{nk}^2) \rightarrow 0$.

Zusatz. Wenn es ein den vorigen Bedingungen genügendes System $\|\bar{y}_{nk}\|$ gibt, so bestimmen die folgenden Formeln ein solches System:

$$\begin{aligned} \bar{y}_{nk} &= y_k, & \text{wenn } |z_k| \leq n; \\ \bar{y}_{nk} &= \mathfrak{D}(y_k), & \text{wenn } |z_k| > n. \end{aligned}$$

Satz XI. Damit (19) gegen Null konvergiert, sind auch die folgenden Bedingungen notwendig und hinreichend:

1. $\sum_{k=1}^n \mathfrak{B}(|z_k| > n) \rightarrow 0$,
2. $\frac{1}{n} \sum_{k=1}^n \mathfrak{D}_{(|z_k| \leq n)}(z_k) \rightarrow 0$,
3. $\frac{1}{n^2} \sum_{k=1}^n \mathfrak{D}_{(|z_k| \leq n)}(z_k^2) \rightarrow 0$.

Moskau, 24. Dezember 1926.

Berichtigung

zu der Abhandlung von H. Brinkmeier: Über das Maß der Bestimmtheit des Wachstums einer ganzen transzendenten Funktion durch die absoluten Beträge der Koeffizienten ihrer Potenzreihe, *Math. Annalen* 96, S. 108—118.

Herr Brinkmeier hat das Hauptergebnis seiner Arbeit, das durch die Formeln (8) und (10) gegeben ist, ausführlich bewiesen, dagegen den Beweis der durch die Formeln (8a), (10a) und (10b) wiedergegebenen weitergehenden Behauptungen lediglich in einer Schlußbemerkung skizziert. Wie ich aus einem Brief von Herrn M. Plancherel entnehme, hat ihn Herr R. Jungen (Zürich) darauf aufmerksam gemacht, daß der Beweis von (10b), wie ihn Brinkmeier andeutet, in den Fällen versagt, wo die Exponenten μ_1, μ_2, \dots sämtlich verschwinden oder der erste nichtverschwindende unter ihnen *negativ* ist. Der durch (8a) gegebene Satz wird durch diese Einwendungen nicht betroffen. Da Brinkmeier inzwischen das Opfer eines Unfalles geworden ist, läßt es sich nicht mehr feststellen, ob sein Beweis für (10a), über den er keine Aufzeichnungen hinterlassen hat, durch den Einwand von Herrn Jungen ebenfalls berührt wird.

Die Frage, ob es ganze Funktionen gibt, die (10a) oder (10b) befriedigen, ist also offen.

Kiel, im September 1927.

O. Toeplitz.

Über analytische Abbildungen einer Klasse von vierdimensionalen Gebieten.

Von

N. Kritikos in München¹⁾.

1. Die vorliegende Arbeit soll einen Beitrag zum Problem der Abbildung von vierdimensionalen Bereichen mittels zweier analytischer Funktionen zweier komplexer Veränderlicher liefern. Sie beantwortet, mit Hilfe der von Herrn Carathéodory²⁾ für dieses Problem erdachten Methoden, die Frage nach den Abbildungen gewisser spezieller Gebiete, zu deren Erklärung wir jetzt übergehen wollen.

2. Im vierdimensionalen Raum der reellen Veränderlichen

$$x_1, x_2, y_1, y_2,$$

die zu den Verbindungen

$$x = x_1 + i x_2, \quad y = y_1 + i y_2$$

zusammengefaßt werden, heiße Kreisgebiet mit dem Mittelpunkt $x = a$, $y = b$ eine beschränkte, offene, zusammenhängende Punktmenge, die mit jedem Punkt

$$x = \xi, \quad y = \eta$$

auch alle Punkte

$$x = a + \xi u, \quad y = b + \eta u$$

mit dem komplexen Parameter u für

$$|u| \leq 1$$

¹⁾ Fellow of the International Education Board.

²⁾ C. Carathéodory, Über das Schwarzsche Lemma bei analytischen Funktionen von zwei komplexen Veränderlichen, *Math. Annalen* 97 (1926), S. 76—98. Diese Arbeit enthält nur den ersten Teil der Untersuchungen von Herrn Carathéodory, ich durfte aber auch andere, noch in Veröffentlichung begriffene Teile benutzen. Für dieses freundliche Entgegenkommen und die Anregung, die mir Herr Carathéodory zuteil werden ließ, möchte ich ihm an dieser Stelle meinen Dank aussprechen.

enthält. Diese Punkte füllen im Falle $(a, b) \neq (\xi, \eta)$ eine Kreisfläche aus, die auf der durch (a, b) und (ξ, η) gehenden, eindeutig bestimmten charakteristischen Ebene

$$\eta(x - a) - \xi(y - b) = 0$$

liegt. Ein Kreisgebiet mit dem Mittelpunkt (a, b) ist also ein Gebiet, aus welchem jede durch (a, b) gehende charakteristische Ebene des Raumes eine (offene) Kreisfläche herauschneidet. Das Kreisgebiet mit seiner Begrenzung werde in üblicher Weise ein abgeschlossenes Kreisgebiet genannt. Ferner verstehe man unter charakteristischem Kreise eine in einer charakteristischen Ebene des Raumes enthaltene Kreislinie, und unter analytischer Abbildung eines Gebietes G_1 des Raumes der x, y auf ein zweites G_2 eine umkehrbar eindeutige, durch zwei in G_1 reguläre analytische Funktionen von x, y vermittelte Beziehung der Punkte von G_1 zu denen von G_2 .

3. Unter den speziellen Kreisgebieten, deren analytische Abbildungen bisher allein näher untersucht worden sind, sind der Dizylinder

$$|x - a| < 1, \quad |y - b| < 1$$

und die Hyperkugel

$$|x - a|^2 + |y - b|^2 < 1$$

zu nennen. Neben denselben kann man sich als einfachste und wichtige Kreisgebiete die folgenden denken.

Auf zwei verschiedenen charakteristischen Ebenen durch den Punkt (a, b) nehme man je einen Kreis mit nicht verschwindendem Radius und dem Zentrum (a, b) . Die konvexe Hülle dieser beiden Kreise, d. i. der kleinste konvexe Bereich im vierdimensionalen Raum, der beide Kreise enthält, ist dann ein abgeschlossenes Kreisgebiet mit dem Mittelpunkt (a, b) . Man sieht es am einfachsten ein, indem man eine affine analytische Transformation des Raumes, d. h. eine ganze lineare Transformation von x, y vornimmt, die die obigen Kreise in die beiden

$$|x| = 1, \quad y = 0$$

bzw.

$$x = 0, \quad |y| = 1$$

überführt. Eine solche Transformation führt offenbar Kreisgebiete mit dem Mittelpunkte (a, b) in solche mit dem Mittelpunkte $(0, 0)$ über und umgekehrt. Außerdem muß sie, als lineare Beziehung, die konvexe Hülle der beiden ersten Kreise in die konvexe Hülle der beiden letzten transformieren. Letztere ist aber der Bereich

$$(1) \quad \bar{K}: \quad |x| + |y| \leq 1,$$

und dies ist ersichtlich ein abgeschlossenes Kreisgebiet mit dem Mittelpunkt $(0, 0)$. Das zugehörige (offene) Kreisgebiet wollen wir mit K be-

zeichnen, also

$$(2) \quad K: \quad |x| + |y| < 1.$$

Es kann als die Normalform der oben erklärten Kreisgebiete angesehen werden.

Was wir nun behandeln wollen, ist die Frage der analytischen Abbildungen dieser Kreisgebiete, welche die konvexe Hülle zweier charakteristischer Kreise bilden — also im wesentlichen des Kreisgebietes K — auf beliebige Kreisgebiete.

4. Ein Teil der aufgeworfenen Frage wird durch folgenden Satz erledigt:

Satz 1. *Jede analytische Abbildung des Kreisgebietes*

$$|x| + |y| < 1$$

auf ein beliebiges Kreisgebiet G mit dem Mittelpunkt (a, b) , bei welcher die Mittelpunkte der beiden Gebiete einander entsprechen, ist eine Affinität:

$$X - a = px + qy,$$

$$Y - b = rx + sy.$$

Dieser Satz ist zum Teil in einem Ergebnis von Herrn Reinhardt³⁾ enthalten. In seinem obigen Umfang ist er durch Herrn Carathéodory bewiesen worden. Wir geben in den §§ 16—17 diesen Beweis in einer teilweise vereinfachten Form wieder.

5. Durch Satz 1 kennen wir alle diejenigen analytischen Abbildungen von K auf Kreisgebiete, die die Mittelpunkte ineinander überführen. Es fragt sich jetzt, ob es analytische Abbildungen gibt, die das nicht tun. Gibt es insbesondere analytische Abbildungen von K auf sich selbst, die den Mittelpunkt verrücken?

Wir werden zeigen, daß diese Fragen zu verneinen sind, in Bestätigung einer dahingehenden Vermutung von Herrn Carathéodory. Das in der Hauptsache zu Beweisende formulieren wir im

Satz 2. *Bei jeder analytischen Abbildung des Kreisgebietes $|x| + |y| < 1$ auf sich selbst bleibt der Mittelpunkt fest.*

Aus diesem letzten Satz in Verbindung mit Satz 1 folgt leicht das Korollar. *Die einzigen analytischen Abbildungen von $|x| + |y| < 1$ auf sich selbst werden durch die Drehungen gegeben:*

$$(3) \quad \begin{aligned} X &= e^{i\alpha} x, \\ Y &= e^{i\beta} y \end{aligned}$$

³⁾ K. Reinhardt, Über Abbildungen durch analytische Funktionen zweier Veränderlicher, Math. Annalen 83 (1921), S. 243, Satz 1.

oder

$$\begin{aligned} X &= e^{i\alpha} y, \\ Y &= e^{i\beta} x \end{aligned}$$

(4)

mit reellem α und β .

Dieses Verhalten von K steht in einem hervorstechenden Gegensatz zu demjenigen des Dizylinders und der Hyperkugel, welche beide bekanntlich analytische Transformationen in sich zulassen, durch die ein beliebiger ihrer Punkte in den Mittelpunkt gebracht werden kann.

6. Andererseits ergibt sich aus Satz 2 der

Satz 3. *Bei jeder analytischen Abbildung von $|x| + |y| < 1$ auf ein Kreisgebiet entsprechen die Mittelpunkte einander und die Abbildung ist eine Affinität.*

Beweis. Gäbe es nämlich eine analytische Abbildung A von K auf ein Kreisgebiet G mit dem Mittelpunkt (a, b) , bei welcher die Mittelpunkte nicht ineinander übergangen, so bilde man weiter G auf sich selbst ab durch die Abbildung

$$B: \quad x' - a = (x - a)e^{i\theta}, \quad y' - b = (y - b)e^{i\theta},$$

in der θ einen festen reellen Wert zwischen 0 und 2π hat. Die zusammengesetzte analytische Abbildung ABA^{-1} bildete dann K auf sich selbst ab, ohne den Mittelpunkt festzulassen, was gegen Satz 2 verstößt. Das weitere folgt dann aus Satz 1.

7. Zusammenfassend werden wir also folgendes sagen können. *Die Kreisgebiete, die zusammen mit ihrer Begrenzung die konvexe Hülle zweier konzentrischer, aber verschiedenen charakteristischen Ebenen angehörender Kreise ausmachen, bilden unter den Kreisgebieten eine Klasse für sich; sie können auf keine anderen Kreisgebiete analytisch abgebildet werden, und die einzigen analytischen Abbildungen, die sie aufeinander oder auf sich selbst abbilden, sind die Affinitäten, welche die Paare der die Gebiete festlegenden Kreise ineinander überführen.*

8. Was nun den Beweis von Satz 2 betrifft, so konnten wir ihn unter der Annahme, daß die Abbildungsfunktionen noch auf dem Rande von K regulär sind und dort eine von Null verschiedene Funktionaldeterminante besitzen, sehr kurz und mit geläufigen funktionentheoretischen Mitteln führen⁴⁾. Ohne diese sehr einschränkende Annahme ist er uns aber nur mit Heranziehung der neuartigen Begriffsbildungen und Ergebnisse von Herrn Carathéodory gelungen. Wir wollen daher in den §§ 9—15 kurz zusammenstellen, was wir daraus brauchen werden.

⁴⁾ N. Kritikos, Über analytische Abbildungen des Gebietes $|x| + |y| < 1$ auf sich, Bull. de la Soc. Math. de Grèce 8 (1927), S. 42—45.

9. Es sei G ein beschränktes Gebiet des vierdimensionalen Raums der x, y . Wir denken uns in G die Carathéodorysche Metrik durch die Distanzfunktion $D_G(A, B)$ erklärt⁵⁾ und fassen einen Punkt $P_0(x_0, y_0)$ von G ins Auge. Es sei

$$P_t: \quad x = \varphi(t), \quad y = \psi(t) \quad (0 \leq t \leq 1)$$

ein reguläres Kurvenstück, das von P_0 ausgeht; dividiert man den für hinreichend kleine positive t gebildeten Vektor

$$x - x_0 = \varphi(t) - \varphi(0), \quad y - y_0 = \psi(t) - \psi(0)$$

durch $D_G(P_t, P_0)$ und geht man zur Grenze $t = 0^+$ über, so erhält man einen Vektor

$$(5) \quad \begin{aligned} \xi - x_0 &= \lim_{t \rightarrow 0^+} \frac{\varphi(t) - \varphi(0)}{D_G(P_t, P_0)}, \\ \eta - y_0 &= \lim_{t \rightarrow 0^+} \frac{\psi(t) - \psi(0)}{D_G(P_t, P_0)}, \end{aligned}$$

der in die Richtung der Tangente an der Kurve im Punkte P_0 fällt und den wir uns von diesem Punkt aus aufgetragen denken. Jetzt variere man die Kurve, so daß sie alle möglichen Richtungen durch P_0 annimmt. Der Vektor (5) beschreibt dann eine Punktmenge $\Gamma_G(x_0, y_0)$, die Herr Carathéodory als die Indikatrix des Gebiets G im Punkt P_0 bezeichnet. Von derselben beweist er, daß sie ein abgeschlossenes, konvexes Kreisgebiet mit dem Mittelpunkt P_0 ist. Ist G ein Kreisgebiet, so wird außerdem gezeigt, daß seine Indikatrix im Mittelpunkte mit der konvexen Hülle von G zusammenfällt.

10. Wegen der Invarianz der Distanzfunktion bei analytischen Abbildungen⁶⁾ gilt nun folgendes.

Das Gebiet G möge auf G^* durch

$$x^* = f(x, y), \quad y^* = g(x, y)$$

analytisch abgebildet sein, der Punkt $P_0(x_0, y_0)$ gehe in $P_0^*(x_0^*, y_0^*)$ über. Dann hängen die Indikatrices $\Gamma_G(x_0, y_0)$ und $\Gamma_{G^*}(x_0^*, y_0^*)$ affin zusammen durch die Gleichungen

$$\begin{aligned} \xi^* - x_0^* &= p(\xi - x_0) + q(\eta - y_0), \\ \eta^* - y_0^* &= r(\xi - x_0) + s(\eta - y_0), \end{aligned}$$

wobei die Transformationskoeffizienten die Werte

$$\begin{aligned} p &= f_x(x_0, y_0), & q &= f_y(x_0, y_0), \\ r &= g_x(x_0, y_0), & s &= g_y(x_0, y_0) \end{aligned}$$

haben.

⁵⁾ § 4, S. 78 der in Fußnote ²⁾ zitierten Carathéodoryschen Abhandlung.

⁶⁾ Satz 1, S. 81 der Carathéodoryschen Abhandlung.

11. Um also Satz 2 zu beweisen, genügt es zu zeigen, daß die Indikatriz von K im Mittelpunkt und in einem beliebigen andern seiner Punkte keine solche affine Transformation der einen in die andere gestatten.

Nun ist, nach § 9, die Indikatrix eines konvexen Kreisgebietes in seinem Mittelpunkt identisch mit dem abgeschlossenen Kreisgebiet selbst. Also kennen wir die Indikatrix von K im Mittelpunkt (d. i. \bar{K} nach unserer Bezeichnung (1)), und es handelt sich weiter darum, ausreichenden Aufschluß über die Indikatrix von K in einem beliebigen andern Punkt zu erlangen, um den Nachweis der Unmöglichkeit eines affinen Zusammenhangs zwischen beiden führen zu können.

12. Die Lösung der letzten Aufgabe wird durch den Begriff der metrischen Ebene ermöglicht. Darunter wird folgendes verstanden.

Es seien

$$x = f(u), \quad y = g(u)$$

zwei im Einheitskreis $|u| < 1$ reguläre analytische Funktionen, die z. B. eine umkehrbar eindeutige Zuordnung der Punkte eines beschränkten Stückes F einer charakteristischen Fläche im vierdimensionalen Raum der x, y zu den Punkten des Einheitskreises $|u| < 1$ vermitteln. Die Gesamtheit der analytischen Funktionen von x, y , die in den Punkten von F regulär und absolut genommen < 1 sind, erzeugt, genau wie bei einem vierdimensionalen Gebiet, eine Metrik auf F vermöge der analog zu definierenden Distanz $D_F(P_1, P_2)$ zweier Punkte P_1 und P_2 von F . Wie Herr Carathéodory zeigt, ist diese Distanz gleich der nichteuclidischen Entfernung $E(u_1, u_2)$ der entsprechenden Punkte im Kreis $|u| < 1$:

$$(6) \quad D_F(P_1, P_2) = E(u_1, u_2).$$

Man setze jetzt voraus, F liege innerhalb des Gebietes G . Dann gilt offenbar, für zwei beliebige Punkte von F ,

$$D_F(P_1, P_2) \geq D_G(P_1, P_2). \quad 7)$$

Besteht nun in dieser Relation das Gleichheitszeichen durchweg für alle Punktepaare, so heißt F nach Herrn Carathéodory eine metrische Ebene von G . Die Kenntnis einer metrischen Ebene liefert uns die Metrik von G wenigstens für die Punkte dieses Gebietes, die auf dem zweidimensionalen Flächenstück liegen.

13. Eine Methode, um metrische Ebenen zu konstruieren, gewinnt Herr Carathéodory durch folgende Betrachtungen.

Es werde mit Z der Dizylinder

$$|x| < 1, \quad |y| < 1$$

7) Vgl. Satz 2, S. 81 der Carathéodoryschen Abhandlung.

bezeichnet, und es sei F ein in Z gelegenes charakteristisches Flächenstück, gegeben durch die Gleichung

$$y = f(x),$$

worin $f(x)$ eine für $|x| < 1$ reguläre analytische Funktion mit einem absoluten Betrag < 1 darstellt. P_1 und P_2 seien zwei Punkte von F , die den Punkten x_1 bzw. x_2 entsprechen mögen. Nach dem Schwarzschen Lemma in der Formulierung von Herrn Pick⁸⁾ gilt

$$E(x_1, x_2) \geq E(f(x_1), f(x_2)).$$

Also ist⁹⁾

$$D_Z(P_1, P_2) = E(x_1, x_2).$$

Andererseits haben wir nach (6)

$$D_F(P_1, P_2) = E(x_1, x_2),$$

folglich gilt für zwei beliebige Punkte von F

$$D_F(P_1, P_2) = D_Z(P_1, P_2),$$

d. h. aber: F ist eine metrische Ebene von Z .

14. Dieses Resultat läßt sich jetzt folgendermaßen verwerten, um, unter gewissen Voraussetzungen, metrische Ebenen auch für andere Gebiete G nachzuweisen.

Es liege das obige Flächenstück F in G und G sei eine Teilmenge von Z . Dann gilt für zwei beliebige Punkte von F

$$D_F(P_1, P_2) \geq D_G(P_1, P_2) \geq D_Z(P_1, P_2)^{10)}.$$

Andererseits ist, wie wir soeben gesehen haben,

$$D_F(P_1, P_2) = D_Z(P_1, P_2).$$

Folglich besteht die Gleichung

$$D_F(P_1, P_2) = D_G(P_1, P_2),$$

und F ist eine metrische Ebene nicht nur von Z , sondern auch von G .

15. Um also metrische Ebenen eines Gebietes G zu finden, kann man versuchen, durch Ausführung einer geeigneten analytischen Abbildung A von G , dieses Gebiet in ein solches G^* zu transformieren, das in einem Dizylinder Z liegt und dessen Begrenzung einen gemeinsamen Teil T mit der Begrenzung von Z hat. Läßt sich dann, was nicht immer der Fall zu sein braucht, ein charakteristisches Flächenstück F^* so konstruieren,

⁸⁾ G. Pick, Über eine Eigenschaft der konformen Abbildung kreisförmiger Bereiche, *Math. Annalen* 77 (1916), S. 1–6.

⁹⁾ Satz 5, S. 83 der Carathéodoryschen Abhandlung.

¹⁰⁾ Siehe Zitat ⁷⁾.

daß es selbst in G^* , sein Rand aber in T liegt, so ist, nach dem Vorhergehenden, F^* eine metrische Ebene von G^* , folglich das durch A^{-1} gewonnene Bild F von F^* eine solche von G .

16. Wir bringen nun zunächst den Beweis von Satz 1. Er stützt sich auf den

Hilfssatz. *Es seien $\varphi(x, y)$ und $\psi(x, y)$ regulär in K . Zweitens sei*

$$|\varphi(x, y)| + |\psi(x, y)| < 1$$

für alle Punkte von K . Drittens sollen in der Umgebung von $(0, 0)$ die Entwicklungen

$$(7) \quad \begin{aligned} \varphi(x, y) &= x + \sum_{m+n \geq 2} a_{mn} x^m y^n, \\ \psi(x, y) &= y + \sum_{m+n \geq 2} b_{mn} x^m y^n \end{aligned}$$

gelten. Dann ist

$$\varphi(x, y) \equiv x, \quad \psi(x, y) \equiv y.$$

Beweis. Es sei $0 \leq \alpha \leq 1$ und u eine komplexe Variable mit einem absoluten Betrag < 1 . Die Punkte $(\alpha u, (1 - \alpha)u)$ und $(\alpha u, (\alpha - 1)u)$ liegen dann sämtlich in K , daher sind die Funktionen

$$\Phi(u) = \varphi(\alpha u, (1 - \alpha)u) + \psi(\alpha u, (1 - \alpha)u),$$

$$\Psi(u) = \varphi(\alpha u, (\alpha - 1)u) - \psi(\alpha u, (\alpha - 1)u)$$

nach unseren Voraussetzungen für $|u| < 1$ regulär und absolut genommen < 1 . Ferner ist

$$\Phi(0) = \Psi(0) = 0,$$

$$\Phi'(0) = \Psi'(0) = 1.$$

Also gilt, nach dem Schwarzschen Lemma,

$$\Phi(u) \equiv \Psi(u) \equiv u$$

für alle α aus dem Intervall $(0, 1)$. Daraus ergeben sich, unter Berücksichtigung der Entwicklungen (7), die Gleichungen

$$\sum_{m=0}^k (a_{m, k-m} + b_{m, k-m}) \alpha^m (1 - \alpha)^{k-m} = 0,$$

$$\sum_{m=0}^k (a_{m, k-m} - b_{m, k-m}) \alpha^m (\alpha - 1)^{k-m} = 0$$

für $k = 2, 3, \dots$ und alle α des Intervalls $(0, 1)$. Folglich muß

$$a_{mn} + b_{mn} = 0,$$

$$a_{mn} - b_{mn} = 0$$

oder

$$a_{mn} = b_{mn} = 0$$

für alle in Betracht kommenden Indizes sein, womit die Behauptung bewiesen ist.

17. Beweis von Satz 1. Es liege jetzt eine analytische Abbildung

$$(8) \quad X = f(x, y), \quad Y = g(x, y)$$

von K auf ein Kreisgebiet G vor, als dessen Mittelpunkt wir ersichtlich den Punkt $(0, 0)$ annehmen dürfen. Die Abbildung soll den Mittelpunkt festlassen. Wir transformieren G durch eine analytische Affinität

$$(9) \quad \begin{aligned} X^* &= AX + BY, \\ Y^* &= CX + DY \end{aligned}$$

in ein Kreisgebiet G^* derart, daß die Funktionen

$$(10) \quad \begin{aligned} \varphi(x, y) &= Af(x, y) + Bg(x, y), \\ \psi(x, y) &= Cf(x, y) + Dg(x, y), \end{aligned}$$

welche die analytische Abbildung von K auf G^* vermitteln, den Bedingungen genügen:

$$\begin{aligned} \varphi'_x(0, 0) &= 1, & \varphi'_y(0, 0) &= 0, \\ \psi'_x(0, 0) &= 0, & \psi'_y(0, 0) &= 1. \end{aligned}$$

Nach § 10 sind dann die Indikatrizten $\Gamma_K(0, 0)$ und $\Gamma_{G^*}(0, 0)$ miteinander identisch; andererseits fällt jede derselben nach § 9 mit der konvexen Hülle des zugehörigen Kreisgebietes zusammen. Folglich ist die konvexe Hülle von G^* identisch mit \bar{K} , und mithin ist G^* in K enthalten. Die Funktionen (10) erfüllen demnach alle Voraussetzungen des Hilfssatzes, und wir können schließen, daß die Abbildung von K auf G^* die Identität ist, daß also K mit G^* zusammenfällt. Die Abbildung (8) von K auf G ist somit invers zur Affinität (9), also selbst eine analytische Affinität. Damit ist Satz 1 bewiesen.

18. Wir gehen jetzt zur Bestimmung von metrischen Ebenen des Kreisgebietes K über.

Durch die analytische Transformation

$$(11) \quad x' = x + y, \quad y' = x - y$$

wird K in ein Gebiet K' übergeführt, das offenbar im Dizylinder

$$Z': \quad |x'| < 1, \quad |y'| < 1$$

enthalten ist. Die Punkte

$$x = \gamma e^{i\theta}, \quad y = (1 - \gamma) e^{i\theta},$$

mit reellem γ aus dem Intervalle $0 \leq \gamma \leq 1$ und beliebigem reellem θ , gehören der Begrenzung von K an und werden in die Punkte

$$x' = e^{i\theta}, \quad y' = (2\gamma - 1) e^{i\theta}$$

transformiert, die sowohl Begrenzungspunkte von K' als auch von Z' sind. Die Punktmenge

$$(12) \quad T': \quad x' = e^{i\theta}, \quad y' = (2\gamma - 1)e^{i\theta},$$

für alle reellen θ und jedes γ aus $0 \leq \gamma \leq 1$, ist also ein Teil des Durchschnitts der beiden Begrenzungen von K' und Z' ; auf ihr ist das Verhältnis $y':x'$ reell und außerdem

$$-1 \leq \frac{y'}{x'} \leq 1.$$

19. Es handelt sich jetzt nach § 15 darum, eine analytische Funktion

$$y' = f(x')$$

zu finden, die für $|x'| < 1$ regulär ist und im vierdimensionalen Raum durch ein Flächenstück innerhalb K' dargestellt wird, dessen Rand in T' liegt. Dazu ist notwendig, daß die Funktion

$$\varphi(x') = \frac{f(x')}{x'}$$

für $|x'| < 1$ regulär sei bis auf den Punkt $x' = 0$, wo sie einen Pol erster Ordnung haben kann, und daß sie reelle Grenzwerte aus dem Intervall $(-1, 1)$ annehme, wenn x' gegen einen Punkt der Kreisperipherie $|x'| = 1$ strebt. Diese Bedingung ist erfüllt, wenn man $\varphi(x')$ entweder als eine reelle Konstante aus dem Intervall $(-1, 1)$ oder als eine Funktion definiert, die den Einheitskreis $|x'| < 1$ auf die längs eines Stückes dieses reellen Intervalls aufgeschlitzte komplexe Zahlenebene konform abbildet und dabei den Punkt $x' = 0$ in den unendlich fernen Punkt überführt.

Sind ϱ und σ die zwei Endpunkte des Schlitzes, so erhält man nach bekanntem elementarem Verfahren eine Abbildungsfunktion durch die Gleichung

$$(13) \quad \varphi(x') = \frac{(\varrho - \sigma)x'^2 + 2(\varrho + \sigma)x' + \varrho - \sigma}{4x'}$$

Zur Verifizierung kann man $\varphi(e^{i\theta})$ berechnen; man findet sofort

$$\varphi(e^{i\theta}) = \varrho \cos^2 \frac{\theta}{2} + \sigma \sin^2 \frac{\theta}{2},$$

also beschreibt $\varphi(x')$ einen geschlossenen Weg auf der reellen Strecke von ϱ bis σ , wenn x' die Kreisperipherie $|x'| = 1$ durchläuft.

Für $\varrho = \sigma$ wird $\varphi(x')$ gleich der Konstanten ϱ . Die Gleichung (13) wird also sämtliche oben angeführten möglichen Definitionen von $\varphi(x')$ in sich fassen, wenn wir die reellen Parameter ϱ und σ unabhängig voneinander das Intervall $(-1, 1)$ durchlaufen lassen.

20. Von der Funktion $\varphi(x')$ gehen wir jetzt zu

$$y' = f(x') = x' \varphi(x') = \frac{\varrho - \sigma}{4} x'^2 + \frac{1}{2} (\varrho + \sigma) x' + \frac{\varrho - \sigma}{4}$$

über; $f(x')$ ist für $|x'| < 1$ regulär und die Beziehung $y' = f(x')$ wird, für diese x' , im vierdimensionalen Raum durch ein charakteristisches Flächenstück F' dargestellt, dessen Rand

$$x' = e^{i\theta}, \quad y' = f(e^{i\theta}) = \left(\varrho \cos^2 \frac{\theta}{2} + \sigma \sin^2 \frac{\theta}{2} \right) e^{i\theta}$$

in (12) liegt. Daraus folgern wir, daß F' in K' enthalten ist.

In der Tat ist K' , als affines Bild von K , ein konvexes Gebiet. Wir fassen eine beliebige Stützhyperebene von K' ins Auge; sie kann durch eine Gleichung

$$\Re(a x' + b y' + c) = 0$$

dargestellt werden, worin a, b, c komplexe Konstanten sind und \Re den reellen Teil des in der Klammer stehenden Ausdrucks bedeutet. Dabei können wir das Vorzeichen der a, b, c so wählen, daß K' im Halbraum

$$\Re(a x' + b y' + c) > 0$$

zu liegen kommt. Nach dem Vorhergehenden ist nun für $|x'| = 1$

$$\Re(a x' + b f(x') + c) \geq 0,$$

denn für diese Werte von x' liegt ja der Punkt $(x', f(x'))$ auf der Begrenzung von K' . Mithin gilt, nach dem Satz über das Maximum einer harmonischen Funktion, für $|x'| < 1$

$$\Re(a x' + b f(x') + c) > 0.$$

Dies besagt aber, daß das Flächenstück F' im selben durch die betrachtete Stützhyperebene bestimmten Halbraum liegt wie K' selbst, und da die Stützhyperebene beliebig war, muß F' im Gebiet K' enthalten sein.

21. Wir können jetzt § 15 in Anwendung bringen und schließen, daß F' eine metrische Ebene von K' ist. Das Bild von F' durch die zu (11) inverse Transformation ist dann eine metrische Ebene von K , die durch die Gleichungen

$$(14) \quad \begin{aligned} x &= \frac{\varrho - \sigma}{8} u^2 + \frac{1}{2} \left(1 + \frac{\varrho + \sigma}{2} \right) u + \frac{\varrho - \sigma}{8}, \\ y &= \frac{\sigma - \varrho}{8} u^2 + \frac{1}{2} \left(1 - \frac{\varrho + \sigma}{2} \right) u + \frac{\sigma - \varrho}{8} \end{aligned}$$

dargestellt wird, wobei der an Stelle von x' stehende Flächenparameter u den Einheitskreis $|u| < 1$ zu beschreiben hat.

22. Aus den letzten Gleichungen bekommt man eine 2-parametrische Schar von metrischen Ebenen des Kreisgebietes K , wenn die reellen Para-

meter ϱ und σ das Intervall $(-1, 1)$ unabhängig voneinander durchlaufen. Diese Schar können wir zu einer 4-parametrischen erweitern, durch Benützung der analytischen Drehungen (3), die offenbar K in sich und somit, wegen der Invarianz der Carathéodoryschen Metrik gegenüber analytischen Abbildungen, metrische Ebenen von K in ebensolche transformieren. Wir erhalten auf diese Weise die 4-parametrische Schar von metrischen Ebenen:

$$(15) \quad \begin{aligned} x &= e^{i\alpha} \left(\frac{\varrho - \sigma}{8} u^2 + \frac{1}{2} \left(1 + \frac{\varrho + \sigma}{2} \right) u + \frac{\varrho - \sigma}{8} \right), \\ y &= e^{i\beta} \left(\frac{\sigma - \varrho}{8} u^2 + \frac{1}{2} \left(1 - \frac{\varrho + \sigma}{2} \right) u + \frac{\sigma - \varrho}{8} \right). \end{aligned}$$

Hingegen liefert die Ausübung einer Drehung (4), die ebenfalls K in sich überführt, nichts mehr Neues, weil dasselbe Resultat schon durch Änderung des Vorzeichens von ϱ und σ erzielt wird. Andererseits sind nicht alle durch (15) dargestellten metrischen Ebenen voneinander verschieden. Es genügt nämlich nicht, um lauter verschiedene zu haben, die reellen Parameter α und β modulo 2π einzuschränken, man muß noch die Möglichkeit ausschließen, daß zwei Darstellungen durch Transformation des Flächenparameters u auseinander hervorgehen. Eine leichte Diskussion zeigt nun, daß man zu diesem Zweck, für $\varrho \neq \sigma$, α und β etwa ≥ 0 und $< 2\pi$, aber nicht beide $\geq \pi$, und, für $\varrho = \sigma$, $\alpha = 0$ und $0 \leq \beta < 2\pi$ zu wählen hat.

Es sei noch bemerkt, daß die für $\varrho = \sigma$ herauskommenden metrischen Ebenen nichts anderes sind als die durch den Anfangspunkt, den Mittelpunkt von K , hindurchgehenden sämtlichen charakteristischen Ebenen des Raumes.

23. Die Zahl der in (15) vorkommenden, im ganzen sechs Parameter könnte den Gedanken nahelegen, daß vielleicht durch je zwei Punkte von K eine metrische Ebene der Schar (15) hindurchgelegt werden kann. Dann müßte es aber, aus Stetigkeitsgründen, auch zu zwei beliebigen Punkten der Begrenzung von K eine metrische Ebene der Schar geben, die dieselben zu Randpunkten hat, und das ist nicht wahr; denn nach Konstruktion behält der Quotient $\frac{y}{x}$ auf der Berandung einer jeden metrischen Ebene (15) ein konstantes Argument bei.

Man erkennt außerdem leicht folgendes. Abgesehen vom Mittelpunkt $(0, 0)$, enthalten die durch einen andern Punkt des Gebietes K hindurchgehenden metrischen Ebenen (15) keineswegs alle von diesem Punkt ausgehenden Richtungen des Raumes, wie es wiederum der Fall sein müßte, wenn der obige Gedanke richtig wäre. Z. B. treten die zu den Ebenen $x = 0$ oder $y = 0$ parallelen Richtungen nicht auf. Wir wollen nun im nächsten Paragraphen wenigstens die Bestimmung der metrischen Ebenen von K für diese letzten Richtungen vornehmen.

24. Es sei (a, b) ein beliebiger Punkt von K . Dann behaupten wir, daß die in K gelegenen charakteristischen Flächenstücke

$$(16) \quad x = a, \quad y = (1 - |a|)u$$

und

$$(17) \quad x = (1 - |b|)u, \quad y = b$$

für $|u| < 1$ metrische Ebenen von K sind.

Wegen der Drehungen (3) und (4), die K in sich transformieren, brauchen wir die Behauptung nur für das erste Flächenstück zu beweisen, und zwar unter der Annahme, daß a reell und nicht-negativ, also

$$0 \leq a < 1$$

ist. Wir betrachten dazu die analytische Abbildung

$$x' = x, \quad y' = \frac{y}{1-x},$$

welche in K regulär und umkehrbar eindeutig ist. Sie führt K in ein Gebiet K' über, das im Dizylinder

$$Z': \quad |x'| < 1, \quad |y'| < 1$$

enthalten ist, denn für die Bilder (x', y') der Punkte (x, y) von K gilt

$$|x'| = |x| < 1,$$

$$|y'| = \frac{|y|}{|1-x|} \leq \frac{|y|}{1-x} < 1.$$

Den Durchschnitt der Begrenzungen von K' und Z' finden wir, indem wir diejenigen Begrenzungspunkte von K' aufsuchen, für die $|x'| = 1$ oder $|y'| = 1$ ist. Es ergeben sich erstens die Punkte mit

$$|x'| = 1, \quad y' = 0,$$

zweitens die Punkte (x', y') mit reellem x' , für welche

$$(18) \quad 0 \leq x' \leq 1 \quad \text{und} \quad |y'| = 1$$

ist.

Fassen wir jetzt das charakteristische Flächenstück

$$(16') \quad x = a, \quad y = (1 - a)u \quad (|u| < 1)$$

ins Auge. Sein Bild

$$x' = a, \quad y' = u \quad (|u| < 1)$$

ist in K' enthalten und besitzt einen Rand

$$x' = a, \quad |y'| = 1,$$

der im Teil (18) des Durchschnitts der Begrenzungen von K' und Z' liegt. Also ist dasselbe nach § 15 eine metrische Ebene von K' , und mithin das Flächenstück (16') eine solche von K , wie zu beweisen war.

25. Wir gehen jetzt dazu über, mit Hilfe der gefundenen metrischen Ebenen von K gewisse Teile der Indikatrix von K in Punkten $P(a, b)$ zu bestimmen, die vom Mittelpunkt verschieden sind. Wir brauchen dabei nur diejenigen P in Betracht zu ziehen, die reelle, nicht-negative Koordinaten x, y besitzen, denn ein beliebiger anderer Punkt von K kann in einen solchen durch Ausführung einer Drehung (3) übergeführt werden. Von a und b werden wir also voraussetzen, daß sie reell und nicht-negativ sind, und daß

$$0 < a + b < 1$$

ist.

26. Wir fragen zunächst nach den metrischen Ebenen aus der 2-parametrischen Schar (14), die durch P hindurchgehen. Dafür ist notwendig und hinreichend, daß, für geeignete Werte von ϱ und σ und ein $u = u_0$ aus $|u| < 1$, die Gleichungen

$$a = \frac{\varrho - \sigma}{8} u_0^2 + \frac{1}{2} \left(1 + \frac{\varrho + \sigma}{2} \right) u_0 + \frac{\varrho - \sigma}{8},$$

$$b = \frac{\sigma - \varrho}{8} u_0^2 + \frac{1}{2} \left(1 - \frac{\varrho + \sigma}{2} \right) u_0 + \frac{\sigma - \varrho}{8}$$

oder die

$$a + b = u_0,$$

$$a - b = \frac{\varrho - \sigma}{4} u_0^2 + \frac{\varrho + \sigma}{2} u_0 + \frac{\varrho - \sigma}{4}$$

gelten. Der Wert u_0 ist also eindeutig festgelegt, und die Parameter ϱ, σ müssen einer linearen Bedingungsgleichung genügen, die wir in der Form

$$(19) \quad \varrho(1 + u_0)^2 - \sigma(1 - u_0)^2 = 4(a - b)$$

schreiben können; außerdem müssen ϱ und σ absolut ≤ 1 sein.

Wir lösen nun (19) nach ϱ auf:

$$(20) \quad \varrho = \frac{4(a - b) + \sigma(1 - u_0)^2}{(1 + u_0)^2}$$

und schreiben die Bedingung dafür hin, daß dieses ϱ absolut ≤ 1 sei; wir erhalten auf diese Weise

$$-(1 + u_0)^2 \leq 4(a - b) + \sigma(1 - u_0)^2 \leq (1 + u_0)^2$$

oder, mit Berücksichtigung von

$$(1 + u_0)^2 = (1 - u_0)^2 + 4(a + b),$$

$$-\frac{(1 - u_0)^2 + 8a}{(1 - u_0)^2} \leq \sigma \leq \frac{(1 - u_0)^2 + 8b}{(1 - u_0)^2}.$$

Diese Bedingung ist aber für alle zulässigen Werte von σ erfüllt. Wir bekommen also sämtliche metrischen Ebenen der Schar (14), die durch den Punkt P laufen, wenn wir in den Gleichungen (14) σ das Intervall $(-1, 1)$ durchlaufen lassen und ϱ als Funktion von σ gemäß (20) betrachten. Dabei entspricht der Punkt P allemal dem Wert $u = u_0 = a + b$ des Flächenparameters.

27. Außer dieser einparametrischen Gesamtheit von metrischen Ebenen durch P werden wir noch die beiden folgenden des § 24 brauchen:

$$x = a, \quad y = (1 - a)u$$

und

$$x = (1 - b)u, \quad y = b.$$

Bei diesen wird der Punkt P für $u = \frac{b}{1-a}$ bzw. $\frac{a}{1-b}$ erhalten.

28. Ist nun durch die Gleichungen

$$(21) \quad x = f(u), \quad y = g(u) \quad (|u| < 1)$$

eine metrische Ebene F von K gegeben, die für $u = u_0$ durch den Punkt $P(a, b)$ hindurchgeht, so bekommt man nach §§ 9, 12 die die Indikatrix $\Gamma_K(a, b)$ bildenden Vektoren für alle Richtungen, welche F im Punkte P berühren, indem man setzt:

$$(22) \quad \begin{aligned} \xi - a &= \lim_{\Delta u=0} \frac{f(u_0 + \Delta u) - f(u_0)}{E(u_0 + \Delta u, u_0)} = f'(u_0) \cdot \lim_{\Delta u=0} \frac{\Delta u}{E(u_0 + \Delta u, u_0)}, \\ \eta - b &= \lim_{\Delta u=0} \frac{g(u_0 + \Delta u) - g(u_0)}{E(u_0 + \Delta u, u_0)} = g'(u_0) \cdot \lim_{\Delta u=0} \frac{\Delta u}{E(u_0 + \Delta u, u_0)}; \end{aligned}$$

$E(u_0 + \Delta u, u_0)$ bedeutet dabei die nichteuklidische Entfernung der beiden Punkte u_0 und $u_0 + \Delta u$ des Einheitskreises $|u| < 1$, und der Grenzübergang des letzteren Punktes in den ersteren hat in einer bestimmten, aber beliebigen Richtung zu erfolgen.

Um den in (22) an letzter Stelle rechts stehenden Limes bequem auszurechnen, nehme man die Transformation

$$u = \frac{v - u_0}{\bar{u}_0 v - 1} \quad (\bar{u}_0 \text{ konjugiert komplex zu } u_0)$$

des Einheitskreises $|u| < 1$ in sich vor, die bekanntlich die nichteuklidischen Entfernungen invariant läßt. Man erhält dann

$$\lim_{\Delta u=0} \frac{\Delta u}{E(u_0 + \Delta u, u_0)} = \left(\frac{dv}{dv} \right)_{v=0} \cdot \lim_{\Delta v=0} \frac{\Delta v}{E(\Delta v, 0)}$$

und, mit Berücksichtigung von

$$E(\Delta v, 0) = \frac{1}{2} \lg \frac{1 + |\Delta v|}{1 - |\Delta v|},$$

schließlich

$$\lim_{\Delta u=0} \frac{\Delta u}{E(u_0 + \Delta u, u_0)} = (|u_0|^2 - 1) e^{i\theta},$$

wo θ einen festen, aber beliebigen reellen Wert hat.

Die so bestimmten Vektoren

$$(23) \quad \begin{aligned} \xi - a &= f'(u_0) (|u_0|^2 - 1) e^{i\theta} \\ \eta - b &= g'(u_0) (|u_0|^2 - 1) e^{i\theta} \end{aligned}$$

mit dem Anfangspunkt (a, b) füllen, für $0 \leq \theta < 2\pi$, eine Kreisfläche aus, die der Schnitt der Indikatrix $\Gamma_K(a, b)$ mit derjenigen charakteristischen Ebene ist, welche die metrische Ebene (21) im Punkte P berührt.

Statt von der Kreisfläche wollen wir im folgenden von dem dieselbe begrenzenden charakteristischen Kreise, dem Ort des Punktes (ξ, η) , sprechen.

29. Die Anwendung der Gleichungen (23) auf die in den §§ 26, 27 besprochenen metrischen Ebenen durch den Punkt P liefert jetzt folgende charakteristische Kreise der Begrenzung von $\Gamma_K(a, b)$: zunächst die einparametrische Gesamtheit

$$(24) \quad \begin{aligned} \xi - a &= - \left[\frac{1}{2} (1 - a - b) (1 + 3a - b) + \frac{\sigma}{2} (1 - a - b)^2 \right] e^{i\theta}, \\ \eta - b &= - \left[\frac{1}{2} (1 - a - b) (1 - a + 3b) - \frac{\sigma}{2} (1 - a - b)^2 \right] e^{i\theta} \end{aligned}$$

für

$$-1 \leq \sigma \leq 1,$$

sodann die zwei weiteren Kreise

$$(25) \quad \xi - a = 0, \quad \eta - b = (1 - a) \left(\frac{b^2}{(1 - a)^2} - 1 \right) e^{i\theta}$$

und

$$(26) \quad \xi - a = (1 - b) \left(\frac{a^2}{(1 - b)^2} - 1 \right) e^{i\theta}, \quad \eta - b = 0.$$

30. Unser letztes Ziel ist nun nach § 11, aus den soeben ermittelten Teilen der Indikatrix $\Gamma_K(a, b)$ den Nachweis für die Unmöglichkeit einer analytischen affinen Abbildung von $\Gamma_K(a, b)$ auf $\Gamma_K(0, 0)$ zu erbringen. Wir benützen dabei folgende von Herrn Carathéodory eingeführte Darstellung der Kreisgebiete durch Punktmengen eines dreidimensionalen reellen Raumes mit den Cartesischen Koordinaten α, β, γ .

Es sei

$$(27) \quad \begin{aligned} \xi - a &= x_0 e^{i\theta}, \\ \eta - b &= y_0 e^{i\theta} \end{aligned} \quad (0 \leq \theta < 2\pi)$$

ein (auf der Ebene $y_0(x-a) - x_0(y-b) = 0$ gelegener) charakteristischer Kreis mit dem Mittelpunkt (a, b) . Wir setzen, falls $x_0 \neq 0$,

$$\gamma = |\xi - a| = |x_0|$$

$$\alpha + i\beta = y_0 \frac{|x_0|}{x_0}$$

und, falls $x_0 = 0$,

$$\gamma = |x_0| = 0$$

$$\alpha + i\beta = y_0^* e^{i\theta}.$$

Für $x_0 \neq 0$ wird also der Kreis durch einen Punkt im Halbraum $\gamma > 0$ repräsentiert, für $x_0 = 0$ durch die Punkte einer ganzen Kreislinie in der Ebene $\gamma = 0$ mit dem Anfangspunkt O der α, β, γ als Zentrum. Umgekehrt entspricht jedem Punkte (α, β, γ) des Halbraums $\gamma > 0$ und jeder Kreislinie mit dem Mittelpunkt O in der Koordinatenebene $\gamma = 0$ eindeutig ein charakteristischer Kreis mit dem Mittelpunkt (a, b) ; dabei sind natürlich die in den Gleichungen (27) auftretenden Konstanten x_0 und y_0 bis auf einen gemeinsamen komplexen Faktor vom absoluten Betrag 1 bestimmt. Wir bemerken schließlich, daß diese gegenseitige Beziehung zwischen charakteristischen Kreisen des vierdimensionalen Raums mit dem Mittelpunkt (a, b) und Punkten des dreidimensionalen Raumes in gewissem unmittelbar verständlichem Sinne stetig ist.

31. Ein Kreisgebiet G mit dem Mittelpunkt (a, b) wird nun nach diesem Verfahren durch eine Punktmenge G^* des Halbraumes $\gamma \geq 0$ von folgender Beschaffenheit dargestellt: G^* ist beschränkt und relativ zum Halbraum $\gamma \geq 0$ offen, sie enthält den Anfangspunkt O , ist von ihm aus gesehen sternförmig, und ihre in der Ebene $\gamma = 0$ gelegenen Punkte füllen eine (unberandete) Kreisfläche aus. Umgekehrt ist jede solche in $\gamma \geq 0$ liegende Punktmenge G^* der Repräsentant eines Kreisgebietes mit dem Mittelpunkt (a, b) . Die relativ zum Halbraum $\gamma \geq 0$ definierten Begrenzungspunkte von G^* , das sind die Grenzpunkte von G^* mit positivem γ und die Berandung der in $\gamma = 0$ gelegenen Kreisfläche von G^* , entsprechen dabei den charakteristischen Kreisen, die die Begrenzung von G ausmachen.

32. Wir wenden jetzt diese Darstellung auf $\Gamma_K(0, 0)$ und auf die bekannten Teile von $\Gamma_K(a, b)$ an. Als Repräsentanten $\Gamma_K^*(0, 0)$ von $\Gamma_K(0, 0)$; d. h. \bar{K} , finden wir den zwischen den Ebenen $\gamma = 0$ und $\gamma = 1$ gelegenen Teil des Kegels

$$\alpha^2 + \beta^2 = (1 - \gamma)^2$$

mit seinem Innern.

Sodann erhalten wir aus den Begrenzungskreisen (24), (25) und (26) von $\Gamma_K(a, b)$ folgende Grenzpunkte von $\Gamma_K^*(a, b)$ in der Ebene $\beta = 0$:

erstens die Gesamtheit

$$(28) \quad \begin{aligned} \alpha &= \frac{1}{2}(1-a-b)(1-a+3b) - \frac{\sigma}{2}(1-a-b)^2, \\ \gamma &= \frac{1}{2}(1-a-b)(1+3a-b) + \frac{\sigma}{2}(1-a-b)^2 \end{aligned}$$

für

$$-1 \leq \sigma \leq 1;$$

zweitens die beiden Punkte

$$(29) \quad \alpha = \pm (1-a) \left(1 - \frac{b^2}{(1-a)^2}\right), \quad \gamma = 0;$$

und drittens den Punkt

$$(30) \quad \alpha = 0, \quad \gamma = (1-b) \left(1 - \frac{a^2}{(1-b)^2}\right).$$

33. Untersuchen wir nun die Lage dieser Punkte in der Ebene $\beta = 0$ mit dem Anfangspunkt O und den Koordinatenachsen α und γ !

Die Punkte (28) machen ein ganzes Stück der Geraden

$$(31) \quad \alpha + \gamma = 1 - (a+b)^2 = c$$

aus, nämlich die Strecke mit den beiden Endpunkten

$$(32) \quad \alpha = 2b(1-a-b), \quad \gamma = (1+a-b)(1-a-b)$$

und

$$(33) \quad \alpha = (1-a+b)(1-a-b), \quad \gamma = 2a(1-a-b).$$

Diese Strecke liegt ganz im ersten Quadranten und erreicht nur im Falle $a=0$ oder $b=0$ mit dem einen Endpunkt eine der Koordinatenachsen. Der Punkt (30) liegt auf der γ -Achse zwischen der Projektion des Endpunktes (32) und dem Durchstoßpunkt der Geraden (31), fällt also für $b=0$ mit diesem Endpunkt zusammen. Ähnlich liegt der Punkt (29) mit dem positiven α auf der α -Achse zwischen der Projektion des Endpunktes (33) und dem Durchstoßpunkt der Geraden (31), fällt also im Falle $a=0$ mit (33) zusammen. Symmetrisch zu ihm bezüglich O ist der zweite Punkt (29). Die Anordnung dieser Grenzpunkte von $\Gamma_K^*(a, b)$ ist offenbar derartig, daß ein System von Geraden, auf denen sie sämtlich liegen würden, mindestens drei Geraden aufweisen muß, von denen keine zwei symmetrisch in bezug auf O sind.

34. Unser Nachweis wird jetzt folgendermaßen zu Ende geführt. Wir leiten zunächst die Beziehung ab, die zwischen den dreidimensionalen Repräsentanten zweier Kreisgebiete besteht, wenn die Kreisgebiete selbst durch eine analytische Affinität miteinander zusammenhängen, und zeigen dann, daß eine solche Beziehung zwischen $\Gamma_K^*(0, 0)$ und $\Gamma_K^*(a, b)$ nicht möglich ist.

35. Da bei der Darstellung eines Kreisgebietes im dreidimensionalen Raum die Lage seines Mittelpunktes keine Rolle spielt, dürfen wir denselben in den Punkt $(0, 0)$ verlegen und die analytische Affinität in der Form

$$(34) \quad \begin{aligned} \xi' &= p\xi + q\eta = (p_1 + ip_2)\xi + (q_1 + iq_2)\eta, \\ \eta' &= r\xi + s\eta = (r_1 + ir_2)\xi + (s_1 + is_2)\eta \end{aligned}$$

ansetzen, wobei natürlich

$$\begin{vmatrix} p & q \\ r & s \end{vmatrix} \neq 0$$

ist. Ein charakteristischer Kreis mit dem Mittelpunkt $(0, 0)$ im Raume der (ξ, η) wird mit Hilfe der Koordinaten α, β, γ des ihn repräsentierenden Punktes und eines reellen Parameters θ durch die Gleichungen

$$\xi = \gamma e^{i\theta}, \quad \eta = (\alpha + i\beta) e^{i\theta}$$

dargestellt. Durch die Transformation (34) geht er in den Kreis

$$\begin{aligned} \xi' &= [p_1\gamma + q_1\alpha - q_2\beta + i(p_2\gamma + q_2\alpha + q_1\beta)] e^{i\theta}, \\ \eta' &= [r_1\gamma + s_1\alpha - s_2\beta + i(r_2\gamma + s_2\alpha + s_1\beta)] e^{i\theta} \end{aligned}$$

über. Wir setzen darin zur Abkürzung

$$\begin{aligned} A &= p_1\gamma + q_1\alpha - q_2\beta, & B &= p_2\gamma + q_2\alpha + q_1\beta, \\ C &= r_1\gamma + s_1\alpha - s_2\beta, & D &= r_2\gamma + s_2\alpha + s_1\beta \end{aligned}$$

und bekommen

$$\begin{aligned} \xi' &= (A + iB) e^{i\theta} = + \sqrt{A^2 + B^2} \left(\frac{A}{\sqrt{A^2 + B^2}} + i \frac{B}{\sqrt{A^2 + B^2}} \right) e^{i\theta}, \\ \eta' &= (C + iD) e^{i\theta} = \left(\frac{A}{\sqrt{A^2 + B^2}} + i \frac{B}{\sqrt{A^2 + B^2}} \right) \left(\frac{A}{\sqrt{A^2 + B^2}} - i \frac{B}{\sqrt{A^2 + B^2}} \right) (C + iD) e^{i\theta}. \end{aligned}$$

Dieser Kreis wird nach § 30 im dreidimensionalen Raum der $(\alpha', \beta', \gamma')$ den durch Punkt

$$\gamma' = |\xi'|, \quad \alpha' + i\beta' = \eta' \frac{|\xi'|}{\xi'}$$

oder die Kreislinie

$$\gamma' = |\xi'|, \quad \alpha' + i\beta' = \eta'$$

repräsentiert, je nachdem $\xi' \neq 0$ oder $= 0$ ist. Es gelten also für die α', β', γ' folgende Transformationsformeln:

$$(35) \quad \begin{aligned} \gamma' &= + \sqrt{A^2 + B^2}, \\ \alpha' &= \frac{AC + BD}{+ \sqrt{A^2 + B^2}}, \\ \beta' &= \frac{AD - BC}{+ \sqrt{A^2 + B^2}} \end{aligned}$$

mit dem Zusatz, daß, im Falle $A = B = 0$, die zwei letzten, unbestimmt werdenden Gleichungen durch

$$\alpha' + i\beta' = (C + iD)e^{i\theta}$$

zu ersetzen sind. Aus diesen Formeln entnimmt man die in allen Fällen gültigen Beziehungen

$$(36) \quad \begin{aligned} \alpha'^2 + \beta'^2 &= C^2 + D^2 \\ \alpha'^2 + \beta'^2 + \gamma'^2 &= A^2 + B^2 + C^2 + D^2. \end{aligned}$$

36. Gäbe es nun eine analytische affine Abbildung von $\Gamma_K(0, 0)$ auf $\Gamma_K(a, b)$, so würde man den Repräsentanten $\Gamma_K^*(a, b)$ durch eine Transformation (34) in $\Gamma_K^*(0, 0)$ überführen können. Die Begrenzung von $\Gamma_K^*(0, 0)$ wird aber, nach § 32, durch die Gleichung

$$\alpha'^2 + \beta'^2 = (1 - \gamma')^2 \quad (0 \leq \gamma' \leq 1)$$

gegeben. Also müßten nach (35) und (36) die Begrenzungspunkte von $\Gamma_K^*(a, b)$ die Gleichung

$$(37) \quad C^2 + D^2 = (1 - \sqrt{A^2 + B^2})^2$$

befriedigen. Die in der Ebene $\beta = 0$ gelegenen Begrenzungspunkte von $\Gamma_K^*(a, b)$ müßten mithin der Gleichung

$$(38) \quad \begin{aligned} &4[(p_1\gamma + q_1\alpha)^2 + (p_2\gamma + q_2\alpha)^2] \\ &= [1 + (p_1\gamma + q_1\alpha)^2 + (p_2\gamma + q_2\alpha)^2 - (r_1\gamma + s_1\alpha)^2 - (r_2\gamma + s_2\alpha)^2]^2 \end{aligned}$$

genügen, die man aus (37) erhält, wenn man $\beta = 0$ setzt und die Quadratwurzel entfernt. Nun gehören zu den letzten Begrenzungspunkten die Punkte einer ganzen Strecke der Geraden (31); folglich müßte die Gleichung (38) identisch in γ erfüllt sein, wenn, gemäß (31),

$$\alpha = c - \gamma$$

eingesetzt wird; d. h. der Ausdruck

$$((p_1 - q_1)\gamma + q_1c)^2 + ((p_2 - q_2)\gamma + q_2c)^2$$

müßte identisch mit dem Quadrat einer ganzen rationalen Funktion von γ sein. Daraus folgt

$$\begin{vmatrix} p_1 - q_1 & p_2 - q_2 \\ q_1 & q_2 \end{vmatrix} = \begin{vmatrix} p_1 & p_2 \\ q_1 & q_2 \end{vmatrix} = 0,$$

und man kann mithin

$$(p_1\gamma + q_1\alpha)^2 + (p_2\gamma + q_2\alpha)^2 \equiv (P\gamma + Q\alpha)^2$$

identisch in α und γ setzen. Die Gleichung (38) kann jetzt in der Form

$$(r_1\gamma + s_1\alpha)^2 + (r_2\gamma + s_2\alpha)^2 = [1 \pm (P\gamma + Q\alpha)]^2$$

geschrieben werden, und man schließt auf dieselbe Weise wie vorhin, daß eine Identität

$$(r_1\gamma + s_1\alpha)^2 + (r_2\gamma + s_2\alpha)^2 \equiv (R\gamma + S\alpha)^2$$

in α und γ gilt. Die Gleichung (38) nimmt also schließlich die Gestalt an:

$$(R\gamma + S\alpha)^2 = [1 \pm (P\gamma + Q\alpha)]^2$$

und stellt somit, in der Ebene $\beta = 0$, höchstens vier verschiedene Geraden dar, die paarweise symmetrisch in bezug auf den Anfangspunkt O sind. Diese Geraden müßten sämtliche in der Ebene $\beta = 0$ gelegenen Begrenzungspunkte von $\Gamma_K^*(a, b)$ enthalten, das ist aber schon für die in § 33 besprochenen Begrenzungspunkte unmöglich. Die Annahme also, daß es eine analytische affine Transformation der Indikatrix $\Gamma_K(0, 0)$ in die Indikatrix $\Gamma_K(a, b)$ gibt, ist unhaltbar, und damit ist der Beweis von Satz 2 vollständig erbracht.

(Eingegangen am 8. 5. 1927.)

Zum Entscheidungsproblem der mathematischen Logik.

Von

Paul Bernays in Göttingen und Moses Schönfinkel in Moskau.

§ 1.

Einleitende Übersicht: Fragestellungen und bisherige Ergebnisse.

Das zentrale Problem der mathematischen Logik, welches auch mit den Fragen der Axiomatik im engsten Zusammenhang steht, ist das *Entscheidungsproblem*. Es handelt sich dabei um folgendes.

Betrachtet werden logische Formeln, welche variable logische Funktionen, d. h. variable Prädikate und Relationen enthalten. Die Argumente dieser logischen Funktionen sind wiederum Variable, deren mögliche Werte von den Dingen eines gewissen „Individuenbereiches“ gebildet werden. Die Variablen für Individuen schreiben wir als kleine, die für logische Funktionen als große lateinische Buchstaben. Aus den Funktionsvariablen und ihren Argumenten setzen sich die zu betrachtenden Formeln mit Hilfe der logischen Zeichen zusammen; es sind dies folgende Symbole:

$\&$	„und“,	Zeichen der Konjunktion;
\vee	„oder“,	„ „ Disjunktion;
\rightarrow	„wenn — so“,	„ „ Implikation;
—	„nicht“,	„ „ Negation;
(x)	„für alle x “,	Allzeichen;
$(\exists x)$	„es gibt ein x “,	Seinszeichen.

Jede in einer Formel vorkommende Individuenvariable x, y, \dots ist auf ein gleichnamiges All- oder Seinszeichen bezogen.

Die logischen Zeichen werden nach der in der symbolischen Logik üblichen Weise inhaltlich gedeutet, und zwar soll auch, wie üblich, die Disjunktion nicht im Sinne des ausschließenden „Oder“ (aut—aut), sondern im Sinne von „vel—vel“ verstanden werden; ferner soll eine Implikation wie

$$P(x) \rightarrow Q(x)$$

gleichbedeutend sein mit

$$\overline{P(x)} \vee Q(x) \quad (,P(x) - \text{nicht, oder } Q(x)\text{“).$$

Bezüglich der Schreibweise der Negation wollen wir festsetzen, daß statt $\overline{F(x)}$, $\overline{G(x, y, z)}$ kürzer $\overline{F}(x)$, $\overline{G}(x, y, z)$ geschrieben werden kann.

Vermöge der inhaltlichen Deutung stellen die Formeln Schemata von *Aussagen* dar, die noch zweierlei variable Elemente enthalten, nämlich erstens den Individuenbereich, auf den sich die kleinen Variablen beziehen, und zweitens die Prädikate und Relationen, welche für die großen Variablen eingesetzt werden können.

Es soll nun ein Verfahren gefunden werden, durch welches man von einer vorgelegten Formel stets entscheiden kann, ob sie für jeden Individuenbereich und für jede Wahl der logischen Funktionen eine richtige Aussage liefert.

Die logischen Funktionen werden hierbei lediglich in Hinsicht auf ihren Wertverlauf betrachtet, d. h. es kommt bei einer logischen Funktion für unsere Frage nur darauf an, daß jedem Wertsystem ihrer Argumente eindeutig einer der Werte „wahr“ oder „falsch“ zugeordnet ist, gleichgültig, wie man zu dieser Zuordnung gelangt.

Wir wollen eine Formel, die bei jeder Einsetzung logischer Funktionen eine richtige Aussage ergibt, eine „allgemeingültige“ Formel nennen.

Die geforderte Entscheidung betrifft die Allgemeingültigkeit einer Formel für *jeden* Individuenbereich.

Ein Beispiel einer für jeden Individuenbereich allgemeingültigen Formel ist

$$(x) (F(x) \vee \overline{F}(x)).$$

Auf Grund der bekannten, die mathematische Logik beherrschenden *Dualität*, gemäß welcher *wahr* und *falsch*, *Konjunktion* und *Disjunktion*, *Allgemeinheit* und *Existenz* einander dual entsprechen, gibt es zu dem gestellten Problem der Entscheidung über *Allgemeingültigkeit* auch ein duales, nämlich die Frage der „*Erfüllbarkeit*“: es soll nach einem festen Verfahren von einer vorgelegten Formel entschieden werden, ob sie für gewisse Individuenbereiche bei passender Wahl der logischen Funktionen eine richtige Aussage liefert.

Die beiden Probleme sind sachlich gleichbedeutend. Nämlich die Frage der Allgemeingültigkeit ist für eine Formel \mathfrak{A} (große deutsche Buchstaben sollen hier, wie in Hilberts neueren Abhandlungen, zur Mitteilung von Formeln und von Formel-Bestandteilen dienen) dann und nur dann zu bejahen, wenn die Frage der Erfüllbarkeit für $\overline{\mathfrak{A}}$ zu verneinen ist. Es genügt daher, die Methoden der Behandlung jeweils für eines der beiden Probleme zu entwickeln. Die Ergebnisse lassen sich dann auf das andere

Problem dualistisch übertragen, ganz so wie es in der projektiven Geometrie auf Grund der dort bestehenden Dualität geschieht.

Eine naturgemäße Erweiterung unserer Fragestellung ergibt sich aus der Rolle des Individuenbereiches. Nämlich statt bei einer vorgelegten Formel zu untersuchen, ob sie für *jeden* Individuenbereich allgemeingültig ist, kann man fragen: Wie muß der Individuenbereich beschaffen sein, damit die Formel allgemeingültig ist?

Entsprechend läßt sich das Problem der Erfüllbarkeit in der Weise erweitern, daß man mit Bezug auf eine vorgelegte Formel fragt: Wie muß der Individuenbereich beschaffen sein, damit die Formel erfüllbar (d. h. bei einer passend gewählten Einsetzung von logischen Funktionen richtig) ist?

Die fragliche Beschaffenheit des Individuenbereiches kann nur in einer *Anzahlbestimmung* bestehen. Denn wenn zwei Individuenbereiche gleich viele Dinge enthalten, wenn also jedem Ding a des einen Bereiches umkehrbar eindeutig ein Ding a' des andern entspricht, so gehört auch zu jeder bestimmten logischen Funktion (von irgendeiner Anzahl k von Argumenten), welche sich auf den ersten Individuenbereich bezieht, eine solche auf den zweiten Bereich bezügliche Funktion (von ebenfalls k Argumenten), die für ein Wertsystem

$$a'_1, \dots, a'_k$$

von Dingen des zweiten Bereiches dann und nur dann wahr ist, wenn die erste Funktion für das entsprechende Wertsystem

$$a_1, \dots, a_k$$

von Dingen des ersten Bereiches wahr ist; und ebenso umgekehrt. Es kann demnach zwischen den beiden Individuenbereichen in Hinsicht auf Allgemeingültigkeit und Erfüllbarkeit logischer Formeln kein Unterschied stattfinden.

Überdies läßt sich von vornherein einsehen, daß die Bedingung für die Allgemeingültigkeit einer Formel nur eine *Höchstzahl*, diejenige für die Erfüllbarkeit nur eine *Mindestzahl* sein kann.

Es gilt nämlich der Satz: Ist eine Formel für irgendeinen Individuenbereich allgemeingültig, so ist sie es auch für jeden Bereich von *kleinerer* Individuenzahl, und ist eine Formel für einen Bereich erfüllbar, so ist es auch für jeden Bereich von *größerer* Individuenzahl.

Gemäß unserem Dualitätsprinzip genügt es, die eine der beiden Behauptungen zu beweisen; wir wählen die zweite. Es sei also eine Formel gegeben, die für einen gewissen Individuenbereich α erfüllbar ist. Wir wollen zeigen, daß die Formel für jeden Bereich β erfüllbar ist, welcher

α als Teilbereich enthält. Damit ist dann auch gezeigt, daß sie überhaupt für jeden Bereich erfüllbar ist, der mehr Individuen als α enthält, — da ja Bereiche von gleicher Individuenzahl sich hinsichtlich der Erfüllbarkeit von Formeln gleich verhalten.

Es seien

$$F, G, \dots, S$$

die in der Formel vorkommenden Funktionszeichen. Nach Voraussetzung gibt es ein System bestimmter logischer Funktionen

$$F_0, G_0, \dots, S_0,$$

welche sich auf den Bereich α beziehen und welche, für F, G, \dots, S eingesetzt, die betrachtete Formel zu einer richtigen machen. c sei ein Ding aus α , und γ sei die Gesamtheit, welche von c und ferner den nicht zu α gehörigen Dingen des Bereiches β gebildet wird. Wir ordnen nun jedem Ding von β ein Ding von α zu, indem wir jedem Ding der Gesamtheit γ das Ding c , und jedes nicht zu γ gehörige Ding von β sich selbst entsprechen lassen. Ferner ersetzen wir die logischen Funktionen F_0, \dots, S_0 durch neue Funktionen

$$F_1, \dots, S_1,$$

welche sich auf den Bereich β beziehen und deren Wahrheitswerte („wahr“ oder „falsch“) so bestimmt werden, daß sie für jedes (aus β entnommene) Wertsystem ihrer Argumente bezüglich mit den Wahrheitswerten der alten Funktionen für das System der zugeordneten Werte aus α übereinstimmen. Auf diese Weise erhalten wir eine auf den Bereich β bezügliche Aussage, welche auch durch Einsetzung aus der betrachteten Formel entsteht, und aus der Bildungsweise der Funktionen F_1, \dots, S_1 geht hervor, daß diese Aussage wiederum richtig ist¹⁾.

Somit ist in der Tat die betrachtete Formel auch in dem Bereich β erfüllbar.

Bei dieser Überlegung haben wir von einer Möglichkeit abgesehen, nämlich der, daß der Individuenbereich *überhaupt kein Ding* enthält. Diesen trivialen Ausnahmefall, der eine Sonderstellung einnimmt, wollen wir bei unseren Betrachtungen *durchweg ausschließen*. Seine Behandlung erledigt sich durch die Bemerkung, daß für einen leeren Individuenbereich eine Formel mit voranstehendem Allzeichen stets allgemeingültig, eine solche mit voranstehendem Seinszeichen stets unerfüllbar ist.

¹⁾ In betreff dieses letzten Punktes der Argumentation sei auf § 2, S. 353—354 hingewiesen, wo für einen ganz entsprechenden Fall die Überlegung ausführlicher angegeben wird.

Um ein paar Beispiele für die Bedingungen der Allgemeingültigkeit bzw. Erfüllbarkeit anzuführen, so ist die Formel

$$(Ex) F(x) \rightarrow (x) F(x)$$

für diejenigen Bereiche allgemeingültig, welche nur ein Ding enthalten; die Formel

$$\{(Ex)(F(x) \& G(x)) \& (Ex)(F(x) \& \bar{G}(x))\} \rightarrow (x) F(x)$$

ist allgemeingültig, sofern der Individuenbereich aus höchstens zwei Dingen besteht. Die Formel

$$(Ex) F(x) \& (Ex) \bar{F}(x)$$

ist erfüllbar für die Individuenbereiche, welche aus mindestens zwei Dingen bestehen, die Formel

$$(Ex) F(x) \& (Ex)(\bar{F}(x) \& G(x)) \& (Ex)(\bar{F}(x) \& \bar{G}(x))$$

ist erfüllbar für die Individuenbereiche, die aus mindestens drei Dingen bestehen.

Wie man leicht sieht, läßt sich *jede endliche Höchstzahl* durch die *Allgemeingültigkeit*, *jede endliche Mindestzahl* durch die *Erfüllbarkeit* einer Formel charakterisieren, und zwar braucht man dabei nur Funktionszeichen mit *einem* Argument.

Auch der Unterschied zwischen *endlichen* und *unendlichen* Individuenbereichen läßt sich durch die Allgemeingültigkeit bzw. Erfüllbarkeit gewisser Formeln zum Ausdruck bringen, wobei nun aber (nach einem noch zu erwähnenden Satz) das Auftreten von Funktionszeichen mit mindestens zwei Argumenten wesentlich ist:

So besteht z. B. die notwendige und hinreichende Bedingung für die Allgemeingültigkeit der Formel

$$\{(x) \bar{F}(x, x) \& (x)(y)(z) ((F(x, y) \& F(y, z)) \rightarrow F(x, z))\} \rightarrow (Ex)(y) \bar{F}(x, y)$$

in der Endlichkeit des Individuenbereiches, und demgemäß die Bedingung für die Erfüllbarkeit der Formel

$$(x) \bar{F}(x, x) \& (x)(y)(z) ((F(x, y) \& F(y, z)) \rightarrow F(x, z)) \& (x)(Ey) \bar{F}(x, y)$$

in der Unendlichkeit des Individuenbereiches.

Der Unterschied zwischen abzählbaren und überabzählbaren Individuenbereichen kann dagegen, nach einem zuerst von Löwenheim²⁾, später auf

²⁾ Leopold Löwenheim, „Über Möglichkeiten im Relativkalkül“, Math. Annalen 76. Leipzig 1915.

andere Weise von Skolem³⁾ bewiesenen Satz⁴⁾, nicht durch die Allgemeingültigkeit oder Erfüllbarkeit einer Formel ausgedrückt werden.

Um diesen Unterschied zur Darstellung zu bringen, muß man zu dem erweiterten Formalismus der „zweiten Stufe“ aufsteigen. Dieser besteht darin, daß das Allzeichen und das Seinszeichen nicht nur in Verbindung mit den kleinen Variablen, sondern auch mit den großen Variablen angewandt werden, so daß man Allgemeinheit und Existenz nicht nur in bezug auf Individuen, sondern auch in bezug auf logische Funktionen formal ausdrücken kann.

Z. B. die Formel

$$(P)(EQ)(x)(\bar{P}(x) \vee \bar{Q}(x))$$

besagt: „Zu jedem Prädikat gibt es ein mit ihm unverträgliches Prädikat“; die Formel

$$(R)(ES)(x)(y)((R(x, y) \rightarrow S(y, x)) \& (S(x, y) \rightarrow R(y, x)))$$

besagt: „Zu jeder zweigliedrigen Relation gibt es eine solche, die aus ihr durch Vertauschung der beiden Glieder entsteht.“

In dieser Symbolik kann die Allgemeingültigkeit und ebenso die Erfüllbarkeit einer Formel der „ersten Stufe“, d. h. einer Formel unseres bisherigen Formalismus, selbst durch eine Formel dargestellt werden.

Z. B. die Allgemeingültigkeit der Formel

$$(x)(F(x) \vee \bar{F}(x))$$

stellt sich dar durch die Formel

$$(F)(x)(F(x) \vee \bar{F}(x)),$$

die Erfüllbarkeit der Formel

$$(x)(y)(R(x, y) \rightarrow R(y, x))$$

durch die Formel

$$(ER)(x)(y)(R(x, y) \rightarrow R(y, x)).$$

Bei dieser Darstellung durch eine Formel der zweiten Stufe ist die Abhängigkeit von den variablen logischen Funktionen beseitigt. Es bleibt in der Formel nur noch der Individuenbereich unbestimmt.

Die beiden Fragen der Allgemeingültigkeit und der Erfüllbarkeit erscheinen nunmehr als Spezialfälle des viel allgemeineren Problems, von einer beliebigen Formel der zweiten Stufe zu entscheiden, ob sie richtig

³⁾ Th. Skolem, „Logisch-kombinatorische Untersuchungen über die Erfüllbarkeit oder Beweisbarkeit mathematischer Sätze nebst einem Theorem über dichte Mengen“. Videnskapselskabet Skrifter. I. Mat. Naturv. Kl., 1920, Nr. 4. Kristiania.

⁴⁾ Dieser Satz wird im folgenden genauer formuliert.

ist oder nicht, bzw. unter welchen Bedingungen für den Individuenbereich sie richtig ist. (Diese Bedingungen können wieder nur Anzahl-Bedingungen sein.)

Eine andere Erweiterung der Problemstellung besteht darin, daß in den zu untersuchenden Formeln neben den variablen logischen Funktionen die *Identitätsrelation* als bestimmte logische Funktion zugelassen wird.

Die ersten Untersuchungen auf dem Gebiete des Entscheidungsproblems sind von Schröder angestellt worden.

Zu systematischen Ergebnissen ist zuerst Löwenheim in seiner (vorhin zitierten) Abhandlung „Über Möglichkeiten im Relativkalkül“ gelangt.

Hierin behandelt er erstens das Entscheidungsproblem für den Fall von Prädikatenformeln, d. h. von solchen Formeln, in denen ausschließlich Funktionen *eines* Arguments auftreten; und zwar löst er es vollständig, für Formeln der ersten Stufe, unter Einbeziehung der Identitätsrelation und andeutungsweise auch für Formeln der zweiten Stufe.

Ferner zeigt er, daß jedes Problem der Entscheidung über Allgemeingültigkeit oder Erfüllbarkeit sich auf ein solches zurückführen läßt, in welchem die zu betrachtende Formel nur Funktionen von *zwei* Argumenten enthält.

Drittens beweist er den Satz, daß eine Formel der ersten Stufe, die überhaupt für irgendeinen Individuenbereich erfüllbar ist, auch für einen abzählbaren Individuenbereich erfüllbar ist oder, in der hierzu dualen Ausdrucksweise: daß eine Formel der ersten Stufe allgemeingültig für jeden Individuenbereich ist, falls sie für jeden abzählbaren Individuenbereich allgemeingültig ist.

Für diesen Satz hat hernach Skolem (in der oben zitierten Abhandlung) einen etwas einfacheren Beweis gegeben.

Das Entscheidungsproblem für die Prädikatenformeln hat, unabhängig von Löwenheim, später Behmann behandelt und zur vollständigen Erledigung gebracht⁵⁾.

Diese Untersuchung liefert insbesondere das Ergebnis, daß die Bedingung für die Individuenzahl nur in (unteren oder oberen) Abgrenzungen durch *endliche* Zahlen bestehen kann⁶⁾, und zwar gilt dies auch für die Formeln der zweiten Stufe und unter Einbeziehung der Identitätsrelation, so daß also der Unterschied zwischen endlichen und unendlichen Gesamtheiten nicht durch die Richtigkeit einer Prädikatenformel der zweiten Stufe, und erst recht nicht durch die Allgemeingültigkeit oder Erfüllbarkeit einer

⁵⁾ Heinrich Behmann, „Beiträge zur Algebra der Logik, insbesondere zum Entscheidungsproblem“. Math. Annalen 86, Heft 3/4 (1922); S. 163—229.

⁶⁾ Der Fall eines unendlichen, durch eine endliche Zahl nach unten abgegrenzten Anzahlintervalles kommt hiernach auch in Betracht.

Prädikatenformel der ersten Stufe — selbst bei Hinzunahme der Identitätsrelation — zum Ausdruck gebracht werden kann.

Im folgenden soll nun für den allereinfachsten bisher noch nicht erledigten Fall das Entscheidungsproblem gelöst werden. Es handelt sich dabei lediglich um Formeln der ersten Stufe unter Ausschluß der Identitätsrelation.

Man gewinnt eine Klassifikation dieser Formeln auf Grund der Tatsache, daß — gemäß einem Satze des Logikkalküls — jede logische Formel in eine solche „Normalform“ umgeformt werden kann, bei der die Allzeichen und Seinszeichen alle *der ganzen Formel voranstehen* und die mit ihr in bezug auf Wahrheit und Falschheit vollkommen gleichwertig ist. Beispielsweise kann die vorhin erwähnte Formel

$$\{(x) \bar{F}(x, x) \& (x)(y)(z) ((F(x, y) \& F(y, z)) \rightarrow F(x, z))\} \\ \rightarrow (Ex)(y) \bar{F}(x, y)$$

auf folgende Normalform gebracht werden:

$$(Ex)(Ey)(Ez)(u) \{F(x, x) \vee (F(x, y) \& F(y, z) \& \bar{F}(x, z)) \vee \bar{F}(x, u)\}.$$

Denken wir uns nun die zu betrachtenden Formeln auf eine solche Normalform gebracht, so sind die einfachsten Fälle von Formeln, die nicht Prädikatenformeln sind, diejenigen, in denen nur *zwei* Zeichen voranstehen. Auf diese Formeln beziehen sich die folgenden Betrachtungen. Von den vier möglichen Typen

$$(x)(y) \mathfrak{A}(x, y), (Ex)(Ey) \mathfrak{A}(x, y), (x)(Ey) \mathfrak{A}(x, y), (Ex)(y) \mathfrak{A}(x, y)$$

erweisen sich die drei ersten für das Problem der Allgemeingültigkeit als trivial. Überhaupt läßt sich, wie gezeigt wird, die Entscheidung über die Allgemeingültigkeit für alle die Formeln in trivialer Weise erledigen, in deren Normalform kein Seinszeichen vor einem Allzeichen steht.

Die Behandlung des vierten Typus wird auf diejenige einer Prädikatenformel zurückgeführt und ergibt eine explizite Übersicht über die möglichen Fälle von Allgemeingültigkeit.

Zur Vorbereitung wird die Entscheidung über Allgemeingültigkeit (bzw. Erfüllbarkeit) bei Prädikatenformeln (d. h. solchen der ersten Stufe, in denen auch die Identitätsrelation nicht auftritt) nach der direkten Methode, wie sie in der Abhandlung von Löwenheim kurz angegeben ist, ausführlich dargelegt, insbesondere im Hinblick auf die dabei resultierende Form der Anzahlbedingung für den Individuenbereich.

Die im folgenden mitgeteilten Überlegungen sind durch Vorlesungen von Hilbert über mathematische Logik angeregt worden und liegen schon um mehrere Jahre zurück. Die Durchführung des Entscheidungsverfahrens

im Falle einer einzigen auftretenden Funktion $F(x, y)$ rührt von M. Schönfinkel her, der zuerst das Problem in Angriff nahm ⁷⁾, von P. Bernays die Ausdehnung der Methode auf mehrere logische Funktionen, sowie die Abfassung der vorliegenden Arbeit.

§ 2.

Das Entscheidungsproblem bei Beschränkung auf Funktionen eines Arguments.

Der methodische Leitgedanke der folgenden Betrachtungen besteht in der Zurückführung des Entscheidungsproblems auf das Entscheidungsverfahren im *Aussagenkalkul*. Im Aussagenkalkul haben wir es zu tun mit Formeln, welche aus den Aussagenvariablen X, Y, Z, \dots (als solche nehmen wir große lateinische Buchstaben, die, zum Unterschied von den Funktionsvariablen, kein Argument bei sich führen) mittels der Konjunktion, Disjunktion, Implikation, Negation gebildet werden.

Die Entscheidung, ob eine solche Formel allgemeingültig ist, d. h. bei jeder Einsetzung von bestimmten Aussagen für die Variablen X, Y, \dots immer eine richtige Aussage liefert, kann direkt durch ein endliches Ausprobieren herbeigeführt werden; denn bei den einzusetzenden bestimmten Aussagen kommt es nur auf ihren Wahrheitswert an, und man braucht daher für die Aussagenvariablen nur die zwei Werte „wahr“ und „falsch“ in Betracht zu ziehen.

Ein besonders einfaches Kriterium für die Allgemeingültigkeit einer Aussagenformel erhält man an Hand der „konjunktiven Normalform“. Unter einer konjunktiven Normalform versteht man eine Formel, welche in einer konjunktiven Zusammensetzung von solchen Disjunktionen besteht, die ihrerseits als Disjunktionsglieder nur Variable oder negierte Variable enthalten ⁸⁾.

Nach einem bekannten Satze des Aussagenkalkuls läßt sich jede Aussagenformel in eine ihr wahrheitsgleiche konjunktive Normalform überführen. Als „wahrheitsgleich“ sollen zwei Aussagenformeln bezeichnet werden, wenn bei jeder bestimmten Einsetzung beide Formeln richtige oder beide falsche Aussagen ergeben.

⁷⁾ Über sein Ergebnis hat Herr Schönfinkel in der Göttinger mathematischen Gesellschaft im Wintersemester 1922/23 referiert.

⁸⁾ Hierin sollen die Fälle mit inbegriffen sein, wo die Formel nur aus einer einzigen Disjunktion besteht, aber auch, wo ein Konjunktionsglied nur aus einer Variablen bzw. einer negierten Variablen besteht. D. h. es sind „eingliedrige“ Konjunktionen und Disjunktionen zugelassen.

Bei einer solchen Normalform besteht nun das notwendige und hinreichende Kriterium für die Allgemeingültigkeit darin, daß jede der Disjunktionen zwei entgegengesetzte Glieder enthält, d. h. zwei solche Glieder, deren eines die Negation des anderen ist. So ist z. B. die Formel

$$(X \vee Y \vee \bar{Y}) \& (\bar{X} \vee X \vee Z)$$

allgemeingültig, nicht aber

$$X \& (\bar{Z} \vee Z).$$

Das duale Gegenstück zu der konjunktiven Normalform ist die *disjunktive* Normalform, die sich von jener dadurch unterscheidet, daß Konjunktion und Disjunktion ihre Rollen vertauschen.

Wie die konjunktive Normalform das Kriterium der Allgemeingültigkeit, so liefert die disjunktive Normalform das Kriterium der Erfüllbarkeit. Dieses Kriterium besteht für eine disjunktive Normalform darin, daß in mindestens einer der (disjunktiv verbundenen) Konjunktionen keine zwei entgegengesetzten Glieder enthalten sind. —

Die Herstellung einer konjunktiven Normalform⁹⁾ für eine gegebene Aussagenformel ist auf verschiedene Weisen möglich. Man kann daher diese Normalform noch weiteren Bedingungen unterwerfen. Insbesondere besagt ein Satz des Aussagenkalküls, daß jede Aussagenformel, welche mit den Aussagenvariablen X, Y, \dots, U gebildet werden kann (es wird nicht verlangt, daß diese Variablen alle in der Formel vorkommen), wahrheitsgleich ist mit einer solchen speziellen konjunktiven Normalform, bei der jede der (konjunktiv verbundenen) Disjunktionen entweder mit der Disjunktion

$$X \vee Y \vee \dots \vee U$$

übereinstimmt oder aus ihr durch Überstreichung einer oder mehrerer Variablen hervorgeht. Indem wir überdies noch verlangen, daß jede solche Disjunktion höchstens einmal als Glied auftritt, wird die Normalform eindeutig (abgesehen von der Reihenfolge der Konjunktionsglieder) festgelegt.

Diese besondere Art von konjunktiver Normalform, von der wir an einer späteren Stelle Gebrauch machen werden, wollen wir „ausgezeichnete“ *konjunktive Normalform* nennen. Dabei ist zu bemerken, daß wir als ausgezeichnete konjunktive Normalform einer *allgemeingültigen* Formel die 0-gliedrige Konjunktion anzusehen haben, welche überhaupt kein Glied enthält.

Die dem Problem der Allgemeingültigkeit angemessene Bevorzugung der konjunktiven Normalform bringt es mit sich, daß wir bei den An-

⁹⁾ Ganz entsprechendes gilt für die disjunktive Normalform.

wendungen des Aussagenkalküls im folgenden stets *die Konjunktion formal als Summe, die Disjunktion als Produkt* zu behandeln haben, — umgekehrt wie es gewöhnlich im Logikkalkül geschieht. —

Betrachten wir nun den einfachsten Fall des anfangs formulierten Entscheidungsproblems, nämlich den, wo als logische Funktionen nur solche mit *einem Argument*, also Prädikate, vorkommen. Hier gelingt in der Tat ganz allgemein die Zurückführung auf das Entscheidungsverfahren des Aussagenkalküls.

Zunächst erkennen wir die Möglichkeit dieser Zurückführung für den Fall eines Individuenbereiches mit einer *gegebenen endlichen Anzahl n* von Individuen.

Seien nämlich

$$a_1, \dots, a_n$$

diese Individuen, so ist die Operation des vorgesetzten Allzeichens (x) gleichbedeutend mit einer Konjunktion, bei der nacheinander für x die Werte a_1, \dots, a_n zu setzen sind, und ebenso ist die Operation des vorgesetzten Seinszeichens (Ex) gleichbedeutend mit einer Disjunktion, wo wiederum für x die Werte a_1, \dots, a_n zu setzen sind.

Ist nun P eine der vorkommenden Funktionsvariablen mit einem Argument, so erhält man bei jeder Einsetzung eines bestimmten Prädikates an Stelle von P für

$$P(a_1), \dots, P(a_n)$$

bestimmte Aussagen mit bestimmten Wahrheitswerten. Umgekehrt wird auch durch jede Verteilung der Werte „wahr“ und „falsch“ auf die Symbole

$$P(a_1), \dots, P(a_n)$$

der Wertverlauf eines Prädikates definiert. Demgemäß erhalten wir an Stelle jeder Funktionsvariablen n voneinander unabhängig variable Wahrheitswerte. Die Allgemeingültigkeit der vorgelegten Formel für eine gegebene Anzahl n von Individuen ist somit gleichbedeutend mit der Allgemeingültigkeit einer Aussagenformel, und für diese können wir die Frage der Allgemeingültigkeit entscheiden.

Hiermit ist freilich unsere Aufgabe nicht gelöst. Denn erstens wollen wir uns nicht auf endliche Individuenbereiche beschränken, und zweitens können wir ja auch nicht für *jede* endliche Individuenzahl die entsprechende Aussagenformel auf ihre Richtigkeit prüfen.

Nun läßt sich aber folgender Satz beweisen: Ist k die Anzahl der in einer Formel von der betrachteten Art vorkommenden Prädikatenvariablen, so ist die Formel allgemeingültig für jeden Individuenbereich, falls sie für einen Bereich von 2^k Individuen allgemeingültig ist.

Für den Beweis ist es vorteilhafter, die hierzu *duale Form* des Satzes zu wählen: Enthält eine Formel von der betrachteten Art k Prädikatenvariablen und ist sie überhaupt für irgendeinen Individuenbereich erfüllbar, so ist sie auch für einen Bereich von 2^k Individuen erfüllbar.

Dies ist folgendermaßen einzusehen.

Seien

$$F, G, \dots, P$$

die in der Formel vorkommenden k Prädikatenvariablen, und seien

$$F_0, G_0, \dots, P_0$$

gewisse auf einen Individuenbereich bezügliche bestimmte Prädikate, für welche die Formel eine richtige Aussage \mathfrak{A} ergibt. Dann teilen wir die Dinge des Individuenbereiches in Klassen ein, indem wir zwei Dinge a, b zur selben Klasse rechnen, wenn die Aussagen $F_0(a)$ und $F_0(b)$ wahrheitsgleich (d. h. beide wahr oder beide falsch) sind, und ebenso

$$G_0(a) \text{ mit } G_0(b), \dots, P_0(a) \text{ mit } P_0(b)$$

wahrheitsgleich ist.

Auf diese Weise erhalten wir höchstens 2^k Klassen, da es ja für ein Ding a höchstens 2^k Möglichkeiten der Verteilung von Wahrheit und Falschheit auf die Aussagen

$$F_0(a), G_0(a), \dots, P_0(a)$$

gibt.

Seien $\alpha_1, \dots, \alpha_n$ die verschiedenen Klassen, die wir so erhalten. Diese bilden einen Individuenbereich von n Dingen, wobei

$$n \leq 2^k$$

ist. Wir definieren nun Prädikate

$$F_1, G_1, \dots, P_1,$$

welche sich auf diesen neuen Individuenbereich beziehen, indem wir festsetzen, daß F_1 auf die Klasse α_p ($p = 1, \dots, n$) dann und nur dann zutrifft, wenn F_0 auf die zu α_p gehörigen Dinge (des anfänglichen Individuenbereichs) zutrifft, daß ebenso G_1 auf α_p dann und nur dann zutrifft, wenn G_0 auf die zu α_p gehörigen Dinge zutrifft usw.

Gemäß dieser Definition geht jede mit den Prädikaten F_0, \dots, P_0 und den logischen Zeichen gebildete, auf den anfänglichen Individuenbereich bezügliche Aussage, wo als Argumente der Prädikate entweder Variable oder bestimmte Dinge des anfänglichen Individuenbereiches stehen, in eine ihr wahrheitsgleiche Aussage über, wenn man darin F_0 durch F_1 , G_0 durch G_1 , \dots , P_0 durch P_1 ersetzt, ferner statt des anfänglichen den

neuen Individuenbereich der n Klassen zugrunde legt und für jedes evtl. als Argument auftretende Ding (des anfänglichen Bereiches) diejenige Klasse setzt, zu welcher das Ding gehört.

Diese Behauptung ist trivial für den Fall, daß die betreffende Aussage kein Allzeichen und kein Seinszeichen enthält. Um sie allgemein zu beweisen, machen wir Gebrauch von der bereits in der Einleitung erwähnten Tatsache, daß jede Formel, in der Variable auftreten, auf eine solche Normalform gebracht werden kann, daß die All- und Seinszeichen sämtlich voranstehen. Wir beschränken uns also auf Aussagen von dieser Normalform.

Der Satz gelte schon, wenn die Anzahl der All- und Seinszeichen in der Aussage kleiner als m ist. Nun betrachten wir eine Aussage mit m voranstehenden Zeichen ($m > 0$). Diese hat dann (bei passender Benennung der Variablen) entweder die Form

$$(x)\mathfrak{B}(x)$$

oder die Form

$$(Ex)\mathfrak{B}(x),$$

wobei in $\mathfrak{B}(x)$ nur noch $(m - 1)$ All- und Seinszeichen voranstehen. Durch die Ausführung der verlangten Änderungen erhält man die Aussage

$$(x)\tilde{\mathfrak{B}}(x) \text{ bzw. } (Ex)\tilde{\mathfrak{B}}(x).$$

Während die vorherige Aussage sich auf den ursprünglichen Individuenbereich bezieht, bezieht sich die geänderte Aussage auf den neuen Individuenbereich. Es ist zu zeigen, daß diese mit jener wahrheitsgleich ist. Dies erkennt man so: Sei a ein Ding des ursprünglichen Individuenbereiches und α die Klasse, zu der a gehört. Dann ist $\mathfrak{B}(a)$ eine Aussage mit $(m - 1)$ voranstehenden All- und Seinszeichen, welche durch Ausführung der verlangten Änderungen in $\tilde{\mathfrak{B}}(\alpha)$ übergeht; nach unserer Annahme ist daher $\mathfrak{B}(a)$ mit $\tilde{\mathfrak{B}}(\alpha)$ wahrheitsgleich. Wenn daher $\mathfrak{B}(x)$ für alle Dinge des ursprünglichen Individuenbereiches zutrifft, so trifft $\tilde{\mathfrak{B}}(x)$ für alle Dinge des neuen Individuenbereiches zu, und umgekehrt; ferner, wenn es ein Ding x im ursprünglichen Individuenbereich gibt, für welches $\mathfrak{B}(x)$ zutrifft, so gibt es auch im neuen Individuenbereich ein Ding x , für welches $\tilde{\mathfrak{B}}(x)$ zutrifft, und umgekehrt.

Somit ist die Behauptung durch den Schluß von n auf $n + 1$ bewiesen. Insbesondere folgt nun hieraus, daß die richtige Aussage \mathfrak{A} , die wir durch eine spezielle Einsetzung aus unserer Ausgangsformel erhalten, bei der Ersetzung von F_0, G_0, \dots, P_0 durch F_1, G_1, \dots, P_1 wieder in eine richtige Aussage übergeht, bei der die Variablen sich auf den Individuenbereich der Klassen $\alpha_1, \dots, \alpha_n$ beziehen, deren Anzahl $\leq 2^k$ ist. Das heißt: die Ausgangsformel ist, wenn sie überhaupt erfüllbar ist, auch in einem

Bereich von höchstens 2^k Individuen, und folglich auch in einem Bereiche von genau 2^k Individuen, erfüllbar.

Somit ist im Gebiete der logischen Funktionen eines Arguments das Problem der Allgemeingültigkeit sowie das der Erfüllbarkeit durch ein bestimmtes Entscheidungsverfahren gelöst. Zugleich ergibt sich bezüglich der Anzahl-Bedingungen für den Individuenbereich ein einfaches Resultat: Eine beschränkende Anzahl-Bedingung für die Allgemeingültigkeit einer Formel mit k Prädikaten-Variablen kann nur lauten: „der Individuenbereich besteht aus höchstens r Dingen“, wobei $1 \leq r < 2^k$; eine beschränkende Anzahl-Bedingung für die Erfüllbarkeit kann nur lauten: „der Individuenbereich besteht aus mindestens r Dingen“, wobei $1 < r \leq 2^k$.

§ 3.

Das Problem der Allgemeingültigkeit bei Formeln mit zwei voranstehenden Zeichen. Vorfagen. Triviale Fälle.

Wenden wir uns nun zu dem allgemeinen Fall, daß auch Funktionen mit mehreren Argumenten auftreten. Wir können von vornherein annehmen, daß die zu betrachtenden Formeln die Normalform haben, bei der die All- und Seinszeichen voranstehen.

Wenn nur ein einziges solches Zeichen voransteht, so haben wir noch kein neues Problem; denn dann kommt ja nur eine einzige Variable vor, und ein Funktionszeichen mit mehreren übereinstimmenden Argumenten, wie

$$F(x, x), \quad G(x, x, x),$$

ist ja in Hinsicht auf die möglichen Werte bei bestimmten Einsetzungen gar nicht verschieden von einem Funktionszeichen mit einem einzigen Argument. Die einfachste Möglichkeit eines wesentlichen Auftretens von Relationen (Funktionen mehrerer Argumente) ist also bei zwei voranstehenden Zeichen gegeben.

Dieser Fall zweier voranstehender Zeichen soll nunmehr behandelt werden.

Es sind folgende Typen möglich:

$$(x)(y) \mathfrak{A}(x, y), \quad (Ex)(Ey) \mathfrak{A}(x, y), \quad (x)(Ey) \mathfrak{A}(x, y), \quad (Ex)(y) \mathfrak{A}(x, y).$$

Wir wollen uns zunächst klarmachen, daß es keine Beschränkung der Allgemeinheit bedeutet, wenn wir annehmen, daß die in \mathfrak{A} vorkommenden Funktionszeichen lauter solche mit zwei Argumenten sind.

Die Vermeidung von Funktionen mit nur einem Argument geschieht einfach dadurch, daß wir z. B. statt $F(x)$ schreiben $F(x, x)$, indem wir also den Wertverlauf eines Prädikates als Teil des Wertverlaufes einer Relation auffassen, was sachlich keinen Unterschied macht.

Etwas mühsamer ist es, das Auftreten von Funktionen mit mehr als zwei Argumenten auszuschalten¹⁰⁾. Die Methode, nach der man hierbei verfährt, läßt sich an dem Fall der Funktionen mit drei Argumenten schon hinreichend deutlich darlegen.

Sei G ein Funktionszeichen mit drei Argumenten, welches in $\mathfrak{U}(x, y)$ vorkommt. Da in den Argumentstellen nur die Variablen x und y stehen, so zerlegt sich der in Betracht kommende Wertverlauf von G in denjenigen

$$\text{von } G(x, x, y), \text{ von } G(x, y, x) \text{ und von } G(y, x, x).$$

Dies sind drei Funktionen zweier Argumente, welche lediglich der Bedingung unterliegen, daß sie beim Einsetzen von x für y alle drei dieselbe Funktion von x liefern. Somit können wir für $G(x, x, y)$ eine ganz beliebige Funktion zweier Argumente $G(x, y)$ nehmen. Die beiden andern Funktionen von x und y müssen bei der Ersetzung von y durch x in je eine mit $G(x, x)$ wahrheitsgleiche Funktion von x übergehen; sonst aber sind sie auch ganz beliebig.

Nun erhält man eine Funktion dieser Art auf folgende Weise: Ist $F(x, y)$ eine beliebige Funktion, so ist

$$(F(x, y) \& F(y, x)) \vee (\bar{F}(x, y) \& \bar{F}(y, x))$$

für ein Individuenpaar x, y dann und nur dann richtig, wenn $F(x, y)$ mit $F(y, x)$ wahrheitsgleich ist. Schreiben wir für diesen Ausdruck zur Abkürzung $\mathfrak{C}(x, y)$, so ist $\mathfrak{C}(x, x)$ stets richtig, wie auch die Funktion $F(x, y)$ gewählt wird.

Ferner können wir eine spezielle Funktion $F_0(x, y)$ so wählen, daß für zwei verschiedene Individuen x, y die Aussagen

$$F_0(x, y) \text{ und } F_0(y, x)$$

in der Beziehung des ausschließenden „oder“ („aut — aut“) stehen, so daß $\mathfrak{C}(x, y)$ beim Einsetzen von F_0 für F falsch wird, außer wenn x dasselbe Individuum ist wie y .

Bilden wir nun den Ausdruck

$$(\bar{\mathfrak{C}}(x, y) \vee G(x, y)) \& (\mathfrak{C}(x, y) \vee H(x, y))$$

¹⁰⁾ Man könnte denken, daß hierzu der in der Einleitung erwähnte Satz von Löwenheim anzuwenden wäre, wonach jedes Problem der Allgemeingültigkeit (bzw. der Erfüllbarkeit) sich auf ein solches zurückführen läßt, bei dem nur Relationen mit zwei Argumenten auftreten. Diese Reduktion können wir jedoch hier nicht verwerten, weil durch sie die Zahl der in der logischen Formel voranstehenden Zeichen vermehrt wird, — während für unsern Zweck ein Verfahren erfordert wird, das den Typus der Formel ungeändert läßt.

— worin F und H als willkürliche Funktionen auftreten, während $G(x, y)$ die eben betrachtete Funktion ist —, so geht dieser beim Einsetzen von x für y über in

$$(\overline{\mathfrak{C}}(x, x) \vee G(x, x)) \& (\mathfrak{C}(x, x) \vee H(x, x)),$$

was mit $G(x, x)$ wahrheitsgleich ist, da $\mathfrak{C}(x, x)$ immer richtig ist. Im übrigen aber ist sein Wertverlauf ganz beliebig; denn setzen wir für F die Funktion F_0 ein, so wird für jedes Paar von verschiedenen Individuen x, y die Aussage $\mathfrak{C}(x, y)$ falsch, mithin wird

$$(\overline{\mathfrak{C}}(x, y) \vee G(x, y)) \& (\mathfrak{C}(x, y) \vee H(x, y))$$

wahrheitsgleich mit $H(x, y)$; und $H(x, y)$ ist ja ganz willkürlich.

Es können also in der Tat bei der Untersuchung der Allgemeingültigkeit einer Formel mit zwei voranstehenden Zeichen die Funktionszeichen mit drei Argumenten ausgeschaltet werden. Und nach demselben Verfahren gelingt es auch für die Funktionszeichen mit noch mehr Argumenten.

Wir brauchen somit bei der Untersuchung der vier Formeltypen

$$(x)(y)\mathfrak{A}(x, y), (Ex)(Ey)\mathfrak{A}(x, y), (x)(Ey)\mathfrak{A}(x, y), (Ex)(y)\mathfrak{A}(x, y)$$

nur den Fall in Betracht zu ziehen, daß in $\mathfrak{A}(x, y)$ ausschließlich Funktionszeichen mit zwei Argumenten vorkommen.

Es seien F, G, \dots, S diese Funktionszeichen; ihre Anzahl sei n . Der Ausdruck $\mathfrak{A}(x, y)$ setzt sich nach Art einer Aussagen-Verknüpfung zusammen aus den $4n$ „Komponenten“

$$\begin{array}{cccc} F(x, x), & F(y, y), & F(x, y), & F(y, x), \\ G(x, x), & G(y, y), & G(x, y), & G(y, x), \\ \vdots & \vdots & \vdots & \vdots \\ S(x, x), & S(y, y), & S(x, y), & S(y, x), \end{array}$$

die allerdings nicht sämtlich vorzukommen brauchen.

Wir sind nun zunächst einmal in der Lage, die Frage der Allgemeingültigkeit für die drei ersten Formeltypen direkt durch Zurückführung auf das Entscheidungsverfahren des Aussagenkalküls zu erledigen.

Die Allgemeingültigkeit von

$$(x)(y)\mathfrak{A}(x, y)$$

besagt, daß bei beliebiger Wahl von Funktionen

$$F, G, \dots, S$$

die Aussage $\mathfrak{A}(x, y)$ für jedes Individuenpaar x, y richtig wird.

Sind aber x, y zwei verschiedene Individuen, so können wegen der Willkürlichkeit der Funktionen F, G, \dots, S die Werte „wahr“ und „falsch“ ganz beliebig auf die $4n$ Komponenten verteilt werden. Ersetzen wir

daher in $\mathfrak{A}(x, y)$ die Komponenten durch $4n$ unabhängige Aussagenvariablen, so muß, damit $(x)(y)\mathfrak{A}(x, y)$ allgemeingültig ist, die entstehende Aussagenformel ebenfalls allgemeingültig sein. Andererseits ist diese Bedingung auch hinreichend; denn jeder spezielle Wert von $\mathfrak{A}(x, y)$ geht ja aus jener Aussagenformel durch eine Einsetzung hervor.

Somit ist die Allgemeingültigkeit von $(x)(y)\mathfrak{A}(x, y)$ gleichbedeutend mit derjenigen der Aussagenformel, die man aus $\mathfrak{A}(x, y)$ erhält, indem man die Komponenten durch unabhängige Aussagenvariablen ersetzt.

Betrachten wir nun gemeinsam die beiden Formeltypen

$$(Ex)(Ey)\mathfrak{A}(x, y) \quad \text{und} \quad (x)(Ey)\mathfrak{A}(x, y).$$

Die erste der beiden Formeln stellt (bei jeder Einsetzung spezieller Funktionen für F, \dots, S) eine schwächere Aussage dar als die zweite und diese wiederum eine schwächere Aussage als $(x)\mathfrak{A}(x, x)$. Finden wir daher eine Bedingung, welche für die Allgemeingültigkeit der ersten Formel notwendig, für die der dritten hinreichend ist, so ist diese für alle drei Formeln notwendig und hinreichend.

Eine solche Bedingung können wir aber leicht angeben. Damit nämlich die Formel

$$(Ex)(Ey)\mathfrak{A}(x, y)$$

allgemeingültig ist, muß sie insbesondere für alle diejenigen Funktionen

$$F, G, \dots, S$$

richtige Aussagen liefern, deren Werte unabhängig sind von der Wahl der Individuen x, y . (Z. B. stellt

$$R(x, y) \vee \bar{R}(x, y)$$

eine immer richtige,

$$R(x, y) \& \bar{R}(x, y)$$

eine immer falsche Relation dar, wenn für $R(x, y)$ irgendeine Relation gesetzt wird.)

Beim Einsetzen solcher Funktionen für F, G, \dots, S kann für jede einzelne der (von den Argumentwerten unabhängige) Wahrheitswert beliebig gewählt werden. Somit müssen wir eine allgemeingültige Aussagenformel erhalten, wenn wir in $\mathfrak{A}(x, y)$ je vier Komponenten, die zu demselben Funktionszeichen gehören, durch eine und dieselbe Aussagenvariable ersetzen, so daß an Stelle der $4n$ Komponenten nun n unabhängige Aussagenvariablen treten.

Ist andererseits diese Bedingung erfüllt, so wird für jedes Individuum x und für jede Wahl von Funktionen

$$F, G, \dots, S$$

die Aussage $\mathfrak{A}(x, x)$ richtig sein, d. h. die Formel

$$(x) \mathfrak{A}(x, x)$$

ist dann allgemeingültig.

Das gemeinsame Kriterium für die Allgemeingültigkeit einer Formel

$$(Ex)(Ey) \mathfrak{A}(x, y)$$

und der entsprechenden Formel

$$(x)(Ey) \mathfrak{A}(x, y)$$

besteht also in der Allgemeingültigkeit derjenigen Aussagenformel, die man aus $\mathfrak{A}(x, y)$ erhält, indem man die zu demselben Funktionszeichen gehörigen Komponenten jeweils durch eine und dieselbe Aussagenvariable ersetzt.

Was die *Anzahlbedingungen* für den Individuenbereich betrifft, so ergibt sich aus unserer Betrachtung sofort, daß eine Formel

$$(x)(y) \mathfrak{A}(x, y)$$

für jeden Individuenbereich allgemeingültig ist, falls sie für einen solchen mit *mindestens zwei* Individuen allgemeingültig ist, und daß bei den Formeln

$$(Ex)(Ey) \mathfrak{A}(x, y) \quad \text{und} \quad (x)(Ey) \mathfrak{A}(x, y)$$

die Anzahl der Individuen überhaupt nichts für die Allgemeingültigkeit ausmacht.

Das Entscheidungsverfahren für die Formeln

$$(x)(y) \mathfrak{A}(x, y), \quad (Ex)(Ey) \mathfrak{A}(x, y), \quad (x)(Ey) \mathfrak{A}(x, y)$$

läßt sich in entsprechender Weise auf alle die Formeln mit *mehreren voranstehenden Zeichen* ausdehnen, bei welchen jedes vorkommende Allzeichen jedem vorkommenden Seinszeichen vorausgeht.

Betrachten wir z. B. Formeln vom Typus

$$(x)(y)(z)(Eu)(Ev) \mathfrak{A}(x, y, z, u, v).$$

Damit eine solche Formel allgemeingültig ist für einen Individuenbereich, der mindestens drei Dinge

$$a, b, c$$

enthält, muß sie insbesondere für alle solche Funktionen⁸ richtige Aussagen liefern, deren Werte ungeändert bleiben, wenn jeder (evtl.) von a , b , c verschiedene Argumentwert durch a ersetzt wird. Also muß die 9-gliedrige Disjunktion

$$\mathfrak{A}(a, b, c, a, a) \vee \mathfrak{A}(a, b, c, a, b) \vee \mathfrak{A}(a, b, c, a, c) \vee \dots \vee \mathfrak{A}(a, b, c, c, c)$$

für alle jene Funktionen eine richtige Aussage darstellen, d. h. sie muß als Aussagenverknüpfung, gebildet aus den verschiedenen Komponenten, welche sich durch die möglichen Verteilungen der Werte a, b, c auf die Argumentstellen der vorkommenden Funktionszeichen ergeben, allgemeingültig sein in dem Sinne, daß jene Komponenten durch unabhängige Aussagenvariablen zu ersetzen sind.

Diese Bedingung ist aber offenbar auch hinreichend für die Allgemeingültigkeit der Ausgangsformel, und zwar bei einem beliebigen Individuenbereich.

Mit diesem Kriterium erhalten wir zugleich folgende Resultate: Eine Formel

$$(x)(y)(z)(Eu)(Ev) \mathfrak{A}(x, y, z, u, v)$$

ist dann und nur dann allgemeingültig für jeden Individuenbereich, wenn sie für einen Bereich von mindestens drei Individuen allgemeingültig ist; ihre Allgemeingültigkeit fällt zusammen mit derjenigen der Formel

$$(x)(y)(z) \{ \mathfrak{A}(x, y, z, x, x) \vee \mathfrak{A}(x, y, z, x, y) \vee \dots \vee \mathfrak{A}(x, y, z, z, z) \}. \quad -$$

Von den vier möglichen Formeltypen mit zwei voranstehenden Zeichen haben sich die drei ersten in Hinsicht auf unser Problem als trivial erwiesen. Es bleibt jetzt der vierte Formeltypus zu behandeln.

Hier sei noch darauf hingewiesen, daß bei der Frage der *Erfüllbarkeit* der nicht triviale von den vier Typen der zu dem vierten *duale*, also

$$(x)(Ey) \mathfrak{A}(x, y)$$

ist.

§ 4.

Die Behandlung der Formeln $(Ex)(y) \mathfrak{A}(x, y)$.

Wir haben nunmehr den Hauptfall unseres Problems zu betrachten: Vorgelegt ist eine Formel vom Typus

$$(Ex)(y) \mathfrak{A}(x, y),$$

wobei $\mathfrak{A}(x, y)$ aus den $4n$ Komponenten

$$\begin{array}{cccc} F(x, x), & F(y, y), & F(x, y), & F(y, x) \\ \vdots & \vdots & \vdots & \vdots \\ S(x, x), & S(y, y), & S(x, y), & S(y, x) \end{array}$$

in der Form einer Aussagenverknüpfung zusammengesetzt ist. Es handelt sich darum, zu entscheiden, ob diese Formel allgemeingültig ist, bzw. für welche Individuenbereiche sie allgemeingültig ist.

Wir diskutieren zunächst die Bedingungen der Allgemeingültigkeit für einen Individuenbereich mit einer festen endlichen Anzahl m von Individuen. Seien

$$a_1, a_2, \dots, a_m$$

diese Individuen; dann muß die Disjunktion

$$\mathfrak{A}(a_1, a_2) \vee \mathfrak{A}(a_2, a_3) \vee \dots \vee \mathfrak{A}(a_{m-1}, a_m) \vee \mathfrak{A}(a_m, a_1)$$

für jede Wahl der Funktionen

$$F, G, \dots, S$$

richtig sein. Denn es soll ja

$$(Ex)(y)\mathfrak{A}(x, y)$$

immer richtig sein, d. h. wie auch die Funktionen F, \dots, S gewählt werden, so soll es ein Individuum geben, welches, in $\mathfrak{A}(x, y)$ für x eingesetzt, bei beliebigem y die Formel $\mathfrak{A}(x, y)$ zu einer richtigen Aussage macht. Daher muß in der obigen Disjunktion jedenfalls ein Glied und mithin auch die Disjunktion selbst eine richtige Aussage sein.

Nun gilt ferner, wie wir im § 1 zeigten, daß eine Formel, die für einen Bereich von m Individuen allgemeingültig ist, auch für jeden kleineren Individuenbereich, insbesondere also auch für jeden Teilbereich diese Eigenschaft besitzt. Demnach muß jede Disjunktion von der Form

$$\mathfrak{A}(b_1, b_2) \vee \mathfrak{A}(b_2, b_3) \vee \dots \vee \mathfrak{A}(b_r, b_1),$$

worin $r \leq m$ ist und b_1, \dots, b_r irgendwelche r verschiedenen unter den Individuen a_1, \dots, a_m bedeuten, bei beliebiger Wahl der Funktionen F, G, \dots, S eine richtige Aussage ergeben.

Wir wollen eine solche Disjunktion eine r -gliedrige „zyklische Disjunktion“ nennen, und die Bedingung, daß diese Disjunktion für beliebige Funktionen F, G, \dots, S eine richtige Aussage darstellt, werde kurz als die „Bedingung B_r “ bezeichnet¹¹⁾.

Die Erfüllung der Bedingungen

$$B_1, B_2, \dots, B_m$$

erweist sich somit als notwendig für die Allgemeingültigkeit der Formel $(Ex)(y)\mathfrak{A}(x, y)$ bei einem m -zahligen Individuenbereich. Sie ist aber zugleich auch *hinreichend* für diese Allgemeingültigkeit. Denn ersetzt man in der Formel $(Ex)(y)\mathfrak{A}(x, y)$ die Operation (Ex) durch eine Disjunktion, (y) durch eine Konjunktion, erstreckt über die Individuen a_1, \dots, a_m , und stellt durch distributives Ausmultiplizieren eine konjunktive Normalform her, so sind die Glieder der entstehenden Konjunktion von der Form

$$\mathfrak{A}(a_1, c_1) \vee \mathfrak{A}(a_2, c_2) \vee \dots \vee \mathfrak{A}(a_m, c_m),$$

wobei mit c_a je eines der Individuen

$$a_1, \dots, a_m$$

¹¹⁾ Die Bedingung B_r hängt offenbar nicht von der Wahl der Individuen b_1, \dots, b_r ab.

bezeichnet ist, ohne daß jedoch

$$c_1, \dots, c_m$$

alle voneinander verschieden zu sein brauchen.

Jede solche Disjunktion enthält aber, wie man leicht sieht¹²⁾, als Teil-Disjunktion eine zyklische Disjunktion von höchstens m Gliedern. Sind daher die Bedingungen B_1, \dots, B_m erfüllt, so muß jede der Disjunktionen

$$\mathfrak{A}(a_1, c_1) \vee \dots \vee \mathfrak{A}(a_m, c_m)$$

bei beliebiger Wahl der Funktionen F, \dots, S eine richtige Aussage darstellen, und dasselbe muß daher von der ganzen konjunktiven Normalform gelten, welche wir als die Entwicklung von

$$(E x)(y) \mathfrak{A}(x, y)$$

für den Bereich der Individuen a_1, \dots, a_m erhalten haben.

Es sind also in der Tat die Bedingungen B_1, \dots, B_m notwendig und hinreichend für die Allgemeingültigkeit unserer betrachteten Formel bei einem m -zähligen Individuenbereich.

Diese Bedingungen sind nun keineswegs voneinander unabhängig. Insbesondere folgt aus der Erfüllung von B_{2k} die Erfüllung von B_k . Denn B_{2k} verlangt, daß für beliebige Funktionen

$$F, G, \dots, S$$

¹²⁾ Dies läßt sich z. B. so zeigen: Wird die Funktion $\varphi(k)$ für $k=1, \dots, m$ durch die Gleichung

$$a_{\varphi(k)} = c_k$$

und $\psi(n)$ für positive ganze n durch die Rekursion

$$\psi(1) = 1, \quad \psi(n+1) = \varphi(\psi(n))$$

definiert, so ist

$$a_{\psi(n+1)} = c_{\psi(n)},$$

mithin ist für jedes positive ganze p

$$\mathfrak{A}(a_{\psi(p)}, a_{\psi(p+1)}) \vee \mathfrak{A}(a_{\psi(p+1)}, a_{\psi(p+2)}) \vee \dots \vee \mathfrak{A}(a_{\psi(p+r-1)}, a_{\psi(p+r)})$$

eine Teil-Disjunktion von

$$\mathfrak{A}(a_1, c_1) \vee \dots \vee \mathfrak{A}(a_m, c_m),$$

und zwar eine zyklische, sofern

$$\psi(p), \psi(p+1), \dots, \psi(p+r-1)$$

alle voneinander verschieden sind, dagegen

$$\psi(p+r) = \psi(p)$$

ist. Diese Voraussetzung läßt sich aber mit einem $r \leq m$ erfüllen, da jeder Wert von $\psi(n)$ gleich einer der Zahlen

$$1, 2, \dots, m$$

ist, und daher unter den Zahlwerten

$$\psi(1), \psi(2), \dots, \psi(m+1)$$

jedenfalls zwei gleiche vorkommen müssen.

die Formel

$$\mathfrak{A}(b_1, b_2) \vee \mathfrak{A}(b_2, b_3) \vee \dots \vee \mathfrak{A}(b_{2k}, b_1)$$

eine richtige Aussage darstellt. Wählen wir insbesondere solche Funktionen, deren Wertverlauf ungeändert bleibt, wenn der Argumentwert b_{k+1} stets durch b_1 , ebenso b_{k+2} stets durch b_2, \dots, b_{2k} stets durch b_k ersetzt wird, so ergibt sich, daß für diese Funktionen die Formel

$$\mathfrak{A}(b_1, b_2) \vee \dots \vee \mathfrak{A}(b_k, b_1)$$

immer eine richtige Aussage liefert. Da aber die zugelassenen Funktionen in ihrer Abhängigkeit von

$$b_1, \dots, b_k$$

noch ganz beliebig sind, so muß die Bedingung B_k erfüllt sein.

Setzen wir speziell $k=1$ und $k=2$, so finden wir, daß die Bedingungen B_1 und B_2 in B_4 enthalten sind, daß also im Falle $m \geq 4$ bereits

$$B_3, B_4, \dots, B_m$$

zusammen ein notwendiges und hinreichendes Kriterium für die Allgemeingültigkeit der Formel

$$(E x)(y) \mathfrak{A}(x, y)$$

bei m Individuen darstellen.

Wir wollen nun die Form der Bedingung B_r für $r \geq 3$ genauer in Betracht ziehen.

Gemäß dem im § 2 erwähnten Satz über die ausgezeichnete konjunktive Normalform können wir zunächst die Formel $\mathfrak{A}(x, y)$ durch diejenige aus den $4n$ Komponenten

$$\begin{array}{cccc} F(x, x), & F(y, y), & F(x, y), & F(y, x) \\ \vdots & \vdots & \vdots & \vdots \\ S(x, x), & S(y, y), & S(x, y), & S(y, x) \end{array}$$

gebildete ausgezeichnete konjunktive Normalform

$$\mathfrak{N}(x, y)$$

ersetzen, welche in Hinsicht auf die Verknüpfung der $4n$ Komponenten mit $\mathfrak{A}(x, y)$ wahrheitsgleich ist¹³). Die Bedingung B_r lautet hiernach: Es muß die Disjunktion

$$\mathfrak{N}(b_1, b_2) \vee \mathfrak{N}(b_2, b_3) \vee \dots \vee \mathfrak{N}(b_r, b_1)$$

für jede Wahl der Funktionen F, G, \dots, S eine richtige Aussage ergeben.

¹³ Im Falle, daß $\mathfrak{A}(x, y)$ bereits als Aussagenverknüpfung allgemeingültig ist, wird die Normalform $\mathfrak{N}(x, y)$ 0-gliedrig, und die Bedingung B_r ist dann in trivialer Weise erfüllt.

Die Glieder dieser Disjunktion

$$\mathfrak{N}(b_1, b_2), \mathfrak{N}(b_2, b_3), \dots, \mathfrak{N}(b_r, b_1)$$

setzen sich wiederum aus Komponenten zusammen; und zwar gehören hier zu jedem Funktionszeichen $3r$ Komponenten. Z. B. gehören zu F die Komponenten

$$\begin{array}{ccc} F(b_1, b_1), & F(b_1, b_2), & F(b_2, b_1), \\ F(b_2, b_2), & F(b_2, b_3), & F(b_3, b_2), \\ \vdots & \vdots & \vdots \\ F(b_r, b_r), & F(b_r, b_1), & F(b_1, b_r). \end{array}$$

Die möglichen Werte dieser $3r$ Komponenten bei Einsetzung einer bestimmten Funktion für F sind, da $r \geq 3$ ist, voneinander ganz unabhängig. Ferner sind die Werte von Komponenten, die zu verschiedenen Funktionszeichen gehören, gewiß auch voneinander unabhängig.

Die Bedingung B_r ist somit gleichbedeutend mit der Allgemeingültigkeit derjenigen Aussagenformel, welche man erhält, indem man in der Disjunktion

$$\mathfrak{N}(b_1, b_2) \vee \mathfrak{N}(b_2, b_3) \vee \dots \vee \mathfrak{N}(b_r, b_1)$$

die $3r \cdot n$ Komponenten durch ebenso viele unabhängige Aussagenvariablen ersetzt.

Führen wir dies nun aus, d. h. ersetzen wir

$$F(b_k, b_k), G(b_k, b_k), \dots, S(b_k, b_k)$$

bezüglich durch

$$X_k, Y_k, \dots, U_k \quad (\text{für } k = 1, \dots, r)$$

und

$$F(b_k, b_l), G(b_k, b_l), \dots, S(b_k, b_l)$$

durch

$$X_{kl}, Y_{kl}, \dots, U_{kl} \quad (\text{für } k, l = 1, \dots, r; k \neq l),$$

so erhalten wir an Stelle von

$$\mathfrak{N}(b_1, b_2), \mathfrak{N}(b_2, b_3), \dots, \mathfrak{N}(b_r, b_1)$$

bezüglich die Aussagenformeln

$$\mathfrak{N}_1, \mathfrak{N}_2, \dots, \mathfrak{N}_r,$$

welche sich voneinander nur durch die Indizes der vorkommenden Aussagenvariablen unterscheiden.

\mathfrak{N}_1 ist eine aus den Variablen

$$X_1, Y_1, \dots, U_1, \quad X_2, Y_2, \dots, U_2, \quad X_{12}, Y_{12}, \dots, U_{12}, \quad X_{21}, Y_{21}, \dots, U_{21}$$

gebildete ausgezeichnete konjunktive Normalform, und ganz entsprechend sind

$$\mathfrak{N}_2, \mathfrak{N}_3, \dots, \mathfrak{N}_r$$

gebildet.

Die Bedingung B_r stellt sich jetzt dar durch die Allgemeingültigkeit der Formel

$$\mathfrak{N}_1 \vee \mathfrak{N}_2 \vee \dots \vee \mathfrak{N}_r.$$

Entwickeln wir diese durch distributives Ausmultiplizieren in eine konjunktive Normalform¹⁴⁾, so hat jedes Konjunktionsglied (jeder „Summand“) dieser Normalform die Gestalt

$$\mathfrak{G}_1 \vee \mathfrak{G}_2 \vee \dots \vee \mathfrak{G}_r,$$

wobei \mathfrak{G}_1 ein Summand von \mathfrak{N}_1 , \mathfrak{G}_2 ein Summand von $\mathfrak{N}_2, \dots, \mathfrak{G}_r$ ein Summand von \mathfrak{N}_r ist. Und das Kriterium für die Allgemeingültigkeit einer Aussagenformel, wie es im § 2 angegeben wurde, besagt demnach in der Anwendung auf die Formel

$$\mathfrak{N}_1 \vee \mathfrak{N}_2 \vee \dots \vee \mathfrak{N}_r,$$

daß in jeder solchen Disjunktion

$$\mathfrak{G}_1 \vee \mathfrak{G}_2 \vee \dots \vee \mathfrak{G}_r$$

mindestens eine Aussagenvariable einmal ohne Negation und einmal mit Negation vorkommen muß.

Nun erkennen wir aber sofort, daß für eine der Aussagenvariablen mit zwei Indizes

$$X_{kl}, Y_{kl}, \dots, U_{kl} \quad (k \neq l)$$

der Fall ausgeschlossen ist, daß sie in einer Disjunktion

$$\mathfrak{G}_1 \vee \mathfrak{G}_2 \vee \dots \vee \mathfrak{G}_r$$

sowohl ohne Negation wie mit Negation auftritt. Denn eine solche Variable kommt überhaupt nur in einer von den Normalformen $\mathfrak{N}_1, \dots, \mathfrak{N}_r$ vor, und in dieser, welche ja eine ausgezeichnete Normalform ist, tritt sie in jedem Summanden entweder nur ohne Negation oder nur mit Negation auf.

Demnach sind diese Aussagenvariablen

$$X_{kl}, Y_{kl}, \dots, U_{kl}$$

ohne Einfluß auf die Allgemeingültigkeit der Formel

$$\mathfrak{N}_1 \vee \mathfrak{N}_2 \vee \dots \vee \mathfrak{N}_r,$$

d. h. es ändert sich in Hinsicht auf die Erfüllung der Bedingung B_r nichts, wenn wir in den Summanden der Normalformen $\mathfrak{N}_1, \dots, \mathfrak{N}_r$ überall diese Variablen (bzw. ihre Negationen) wegstreichen.

Vergegenwärtigen wir uns, daß bei der Bildung der Formeln $\mathfrak{N}_1, \mathfrak{N}_2, \dots, \mathfrak{N}_r$ aus $\mathfrak{N}(x, y)$ die Variablen

$$X_{kl}, Y_{kl}, \dots, U_{kl}$$

¹⁴⁾ Diese ist keine ausgezeichnete Normalform.

den Komponenten

$$F(x, y), F(y, x), G(x, y), G(y, x), \dots, S(x, y), S(y, x)$$

zugeordnet sind, so gelangen wir zu folgendem Ergebnis: Für die Erfüllung der Bedingung B_r macht es keinen Unterschied, wenn wir in den Konjunktionsgliedern von $\mathfrak{N}(x, y)$ überall die Komponenten

$$F(x, y), F(y, x), \dots, S(x, y), S(y, x)$$

(bzw. ihre Negationen) streichen, also nur die Komponenten

$$F(x, x), F(y, y), G(x, x), G(y, y), \dots, S(x, x), S(y, y)$$

beibehalten.

Dies gilt nun für jede der Bedingungen B_3, B_4, \dots, B_m . Diese bilden aber, falls $m \geq 4$ ist, in ihrer Gesamtheit die notwendige und hinreichende Bedingung für die Allgemeingültigkeit der Formel

$$(Ex)(y) \mathfrak{A}(x, y)$$

bei m Individuen. Es folgt demnach, daß für einen m -zahligen Individuenbereich, falls $m \geq 4$ ist, unsere betrachtete Formel dann und nur dann allgemeingültig ist, wenn auch diejenige Formel

$$(Ex)(y) \mathfrak{N}'(x, y)$$

es ist, bei der $\mathfrak{N}'(x, y)$ aus der zu $\mathfrak{A}(x, y)$ gehörigen ausgezeichneten konjunktiven Normalform $\mathfrak{N}(x, y)$ durch Wegstreichen der Komponenten

$$F(x, y), F(y, x), G(x, y), G(y, x), \dots, S(x, y), S(y, x)$$

und Weglassen aller dadurch eventuell auftretenden Wiederholungen von Summanden gebildet ist.

In dieser Formel

$$(Ex)(y) \mathfrak{N}'(x, y)$$

kommt der Relationscharakter der Funktionen F, G, \dots, S gar nicht zur Geltung, vielmehr ist sie eine reine *Prädikatenformel*, gebildet aus den n Prädikaten

$$F(x, x), G(x, x), \dots, S(x, x).$$

Auf diese Weise wird unser Problem der Allgemeingültigkeit auf eines von denjenigen zurückgeführt, die wir im § 2 behandelt haben. Dadurch werden wir zugleich frei von der Beschränkung auf eine bestimmte endliche Individuenzahl. Denn indem wir das Endergebnis des § 2 verwerten, gewinnen wir folgendes Resultat:

Die Formel

$$(Ex)(y) \mathfrak{A}(x, y)$$

ist allgemeingültig für jeden Individuenbereich, falls sie allgemeingültig ist für einen Bereich, dessen Individuenzahl ≥ 4 und $\geq 2^n$ ist.

Mit diesen Feststellungen ist die Frage der Allgemeingültigkeit für die Formeln

$$(Ex)(y) \mathfrak{A}(x, y)$$

grundsätzlich erledigt. Wir wollen aber noch die erhaltene Bedingung für die Allgemeingültigkeit (bei einem Bereich von mindestens 4 Individuen) auf eine explizite Form bringen. Als notwendig und hinreichend dafür, daß $(Ex)(y) \mathfrak{A}(x, y)$ für einen Bereich von m Individuen ($m \geq 4$) allgemeingültig ist, haben wir erkannt, daß die Formel

$$(Ex)(y) \mathfrak{N}'(x, y)$$

für diesen Bereich allgemeingültig ist, und dies wiederum ist gleichbedeutend damit, daß $\mathfrak{N}'(x, y)$ die Bedingungen

$$B_1, B_2, \dots, B_m$$

erfüllt. Die Bedingung B_r besagt in Anwendung auf $\mathfrak{N}'(x, y)$, daß diejenige Aussagenformel

$$\mathfrak{N}'_1 \vee \mathfrak{N}'_2 \vee \dots \vee \mathfrak{N}'_r$$

allgemeingültig ist, welche aus

$$\mathfrak{N}'(b_1, b_2) \vee \mathfrak{N}'(b_2, b_3) \vee \dots \vee \mathfrak{N}'(b_r, b_1)$$

entsteht, indem an Stelle von

$$F(b_k, b_k), G(b_k, b_k), \dots, S(b_k, b_k)$$

bezüglich die Aussagen-Variablen

$$X_k, Y_k, \dots, U_k$$

gesetzt werden ($k = 1, \dots, r$). Und zwar brauchen wir hier B_1 und B_2 nicht mehr auszunehmen.

Um nun eine übersichtliche Schreibweise zu gewinnen, führen wir zunächst für die 2^n verschiedenen n -gliedrigen Disjunktionen, welche als erstes Glied $F(x, x)$ bzw. $\bar{F}(x, x)$, als zweites Glied $G(x, x)$ bzw. $\bar{G}(x, x)$, ..., als n -tes Glied $S(x, x)$ bzw. $\bar{S}(x, x)$ haben, eine Numerierung ein und bezeichnen sie in dieser Reihenfolge mit

$$\mathfrak{P}_1(x), \mathfrak{P}_2(x), \dots, \mathfrak{P}_{2^n}(x).$$

Die ausgezeichnete Normalform $\mathfrak{N}'(x, y)$ setzt sich konjunktiv zusammen aus Gliedern von der Form

$$\mathfrak{P}_\alpha(x) \vee \mathfrak{P}_\beta(y) \quad (\alpha, \beta = 1, 2, \dots, 2^n),$$

und wir stellen sie symbolisch dar in der Form

$$\sum_{\alpha, \beta=1}^{2^n} C_{\alpha\beta} (\mathfrak{P}_\alpha(x) \vee \mathfrak{P}_\beta(y)),$$

wobei das Summenzeichen für die Konjunktion steht und der Koeffizient $C_{\alpha\beta}$ gleich 1 oder 0 zu setzen ist, je nachdem in $\mathfrak{N}'(x, y)$ das Glied $\mathfrak{P}_\alpha(x) \vee \mathfrak{P}_\beta(y)$ vorkommt oder fehlt.

Diese Darstellungsweise ist so beschaffen, daß beim distributiven Entwickeln von Disjunktionen die gewöhnlichen Rechenregeln der Algebra anwendbar sind.

Unsere Aufgabe besteht nun darin, diejenigen Koeffizientensysteme $C_{\alpha\beta}$ zu bestimmen, für welche die Formel

$$(Ex)(y) \mathfrak{N}'(x, y)$$

bei einem m -zähligen Individuenbereich allgemeingültig ist. Dazu müssen wir die Bedingungen

$$B_1, \dots, B_m$$

als Gleichungen für die $C_{\alpha\beta}$ ausdrücken.

Die Bedingung B_r verlangt, daß die Aussagenformel

$$\mathfrak{N}'_1 \vee \dots \vee \mathfrak{N}'_r$$

allgemeingültig ist. \mathfrak{N}'_k wird aus $\mathfrak{N}'(x, y)$ erhalten, indem

$$F(x, x), G(x, x), \dots, S(x, x)$$

bezüglich durch

$$X_k, Y_k, \dots, U_k$$

und

$$F(y, y), G(y, y), \dots, S(y, y)$$

durch

$$X_{k+1}, Y_{k+1}, \dots, U_{k+1} \quad \text{für } k \neq r$$

bzw. durch

$$X_1, Y_1, \dots, U_1 \quad \text{für } k = r$$

ersetzt werden.

Bei diesen Ersetzungen geht — so wählen wir die Bezeichnung —

$$\mathfrak{P}_\alpha(x) \text{ in } \mathfrak{P}_\alpha^{(k)}$$

und

$$\mathfrak{P}_\alpha(y) \text{ in } \mathfrak{P}_\alpha^{(k+1)} \quad \text{für } k \neq r$$

$$\text{bzw. in } \mathfrak{P}_\alpha^{(1)} \quad \text{für } k = r$$

über. Und die distributive Entwicklung von

$$\mathfrak{N}'_1 \vee \dots \vee \mathfrak{N}'_r$$

liefert auf Grund unserer symbolischen Darstellung den Ausdruck

$$\sum_{\alpha, \beta, \gamma, \delta, \dots, \mu, \nu=1}^{2^m} C_{\alpha\beta} \cdot C_{\gamma\delta} \dots C_{\mu\nu} \mathfrak{P}_\alpha^{(1)} \vee \mathfrak{P}_\beta^{(2)} \vee \mathfrak{P}_\gamma^{(2)} \vee \mathfrak{P}_\delta^{(3)} \vee \dots \vee \mathfrak{P}_\mu^{(r)} \vee \mathfrak{P}_\nu^{(1)}.$$

Damit diese Aussagenformel allgemeingültig ist, dürfen nur diejenigen Koeffizienten

$$C_{\alpha\beta} \cdot C_{\gamma\delta} \dots C_{\mu\nu}$$

von 0 verschieden sein, für welche die zugehörige Disjunktion

$$\mathfrak{P}_{\alpha}^{(1)} \vee \mathfrak{P}_{\beta}^{(2)} \vee \mathfrak{P}_{\gamma}^{(2)} \vee \mathfrak{P}_{\delta}^{(3)} \vee \dots \vee \mathfrak{P}_{\mu}^{(r)} \vee \mathfrak{P}_{\nu}^{(1)}$$

die Eigenschaft hat, daß in ihr mindestens eine Aussagen-Variable sowohl ohne Negation wie mit Negation als Glied auftritt. Dies ist aber, wie man sich leicht überlegt, dann und nur dann der Fall, wenn unter den \mathfrak{P} -Faktoren zwei mit demselben oberen, aber verschiedenem unteren Index vorkommen, d. h. wenn von den unteren Indizes

$$\alpha, \beta, \gamma, \delta, \dots, \mu, \nu$$

entweder der zweite verschieden von dem dritten oder der vierte verschieden von dem fünften ..., oder der $2r$ -te verschieden von dem ersten ist. Für alle anderen Indexsysteme muß also der zugehörige Koeffizient gleich 0 sein.

Diese Bedingung gilt nun für

$$r = 1, 2, \dots, m$$

und liefert folgende für die Allgemeingültigkeit von $(Ex)(y) \mathfrak{N}'(x, y)$ notwendigen und hinreichenden Bedingungsgleichungen:

$$\begin{aligned} (1) \quad C_{\alpha\alpha} &= 0 & (\alpha = 1, \dots, 2^n), \\ (2) \quad C_{\alpha\beta} \cdot C_{\beta\alpha} &= 0 & (\alpha, \beta = 1, \dots, 2^n), \\ (3) \quad C_{\alpha\beta} \cdot C_{\beta\gamma} \cdot C_{\gamma\alpha} &= 0 & (\alpha, \beta, \gamma = 1, \dots, 2^n), \\ \vdots & & \\ (m) \quad C_{\alpha\beta} \cdot C_{\beta\gamma} \cdot \dots \cdot C_{\mu\nu} \cdot C_{\nu\alpha} &= 0 & (\alpha, \beta, \gamma, \dots, \mu, \nu = 1, \dots, 2^n). \end{aligned}$$

Hierbei genügt es, die Werte der Indizes

$$\alpha, \beta, \gamma, \dots$$

alle voneinander verschieden zu nehmen, da sonst die betreffende Gleichung sich bereits aus den vorhergehenden Gleichungen ergibt.

Somit erhalten wir nun folgendes Entscheidungsverfahren: Ist eine Formel

$$(Ex)(y) \mathfrak{A}(x, y)$$

vorgelegt, bei der $\mathfrak{A}(x, y)$ in der Form einer Aussagen-Verknüpfung aus den $4n$ Komponenten

$$\begin{array}{cccc} F(x, x), & F(y, y), & F(x, y), & F(y, x) \\ \vdots & \vdots & \vdots & \vdots \\ S(x, x), & S(y, y), & S(x, y), & S(y, x) \end{array}$$

zusammengesetzt ist, so ersetzen wir zunächst $\mathfrak{A}(x, y)$ durch die zugehörige (aus den $4n$ Komponenten gebildete) ausgezeichnete konjunktive Normalform $\mathfrak{N}(x, y)$ — (welche im Falle, wo $\mathfrak{A}(x, y)$ schon als Aussagenformel allgemeingültig ist, überhaupt kein Glied enthält). Von $\mathfrak{N}(x, y)$ gehen wir durch Streichen der Komponenten

$$F(x, y), F(y, x), \dots, S(x, y), S(y, x)$$

und Weglassen der dadurch evtl. entstehenden Wiederholungen von Konjunktionsgliedern zu der Normalform $\mathfrak{N}'(x, y)$ über und schreiben diese symbolisch in der Form

$$\sum_{\alpha, \beta=1}^{2^n} C_{\alpha\beta} (\mathfrak{P}_\alpha(x) \vee \mathfrak{P}_\beta(y)).$$

Dann stellen, falls $4 \leq m \leq 2^n$ ist, die obigen Gleichungen

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

die notwendige und hinreichende Bedingung dafür dar, daß die Formel

$$(Ex)(y) \mathfrak{A}(x, y)$$

für jeden m -zähligen Individuenbereich allgemeingültig ist; und die Gleichungen

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

stellen die notwendige und hinreichende Bedingung dafür dar, daß die Formel für jeden beliebigen Individuenbereich allgemeingültig ist.

Besonders einfach gestaltet sich das Ergebnis im Falle $n=1$, wo die Formel $\mathfrak{A}(x, y)$ nur ein einziges Funktionszeichen F enthält. Wir haben dann nur die vier Koeffizienten

$$C_{11}, C_{22}, C_{12}, C_{21},$$

und die Gleichungen (1), (2) lauten:

$$C_{11} = 0, \quad C_{22} = 0, \quad C_{12} \cdot C_{21} = 0.$$

Hiernach kann die Normalform $\mathfrak{N}'(x, y)$ — wenn sie überhaupt ein Glied enthält — nur eine der beiden Formen

$$F(x, x) \vee \bar{F}(y, y), \quad \bar{F}(x, x) \vee F(y, y)$$

haben, und $\mathfrak{A}(x, y)$ muß sich daher (nach den Regeln des Aussagenkalküls) umformen lassen in eine Formel entweder von der Gestalt

$$F(x, x) \vee \bar{F}(y, y) \vee \mathfrak{B}$$

oder von der Gestalt

$$\bar{F}(x, x) \vee F(y, y) \vee \mathfrak{B},$$

wobei \mathfrak{B} aus den Komponenten

$$F(x, y) \text{ und } F(y, x)$$

zusammengesetzt ist.

Diese Beschaffenheit von $\mathfrak{A}(x, y)$ ist also im Falle eines einzigen auftretenden Funktionszeichens notwendig und hinreichend dafür, daß die Formel

$$(Ex)(y)\mathfrak{A}(x, y)$$

für jeden Individuenbereich allgemeingültig ist. (Daß die Bedingung hinreichend ist, kann man leicht direkt verifizieren.)

Wenn die Individuenzahl $m < 4$ ist, so besteht in Hinsicht auf die Allgemeingültigkeit keine Gleichwertigkeit zwischen den Formeln

$$\mathfrak{N}(x, y) \text{ und } \mathfrak{N}'(x, y);$$

infolgedessen sind dann die Bedingungen

$$(1) \dots (m)$$

zwar hinreichend, aber nicht notwendig für die Allgemeingültigkeit.

So ist z. B. die Formel

$$(Ex)(y)(F(x, y) \vee \bar{F}(y, x))$$

allgemeingültig für einen Bereich von zwei Individuen, während die zugehörige Normalform $\mathfrak{N}'(x, y)$, deren symbolische Darstellung

$$\sum_{\alpha, \beta=1}^2 (\mathfrak{B}_\alpha(x) \vee \mathfrak{B}_\beta(y))$$

lautet, keiner der Gleichungen (1), (2) genügt. Und die Formel

$$(Ex)(y) \left((F(x, x) \vee \bar{F}(y, y) \vee F(x, y) \vee \bar{F}(y, x)) \right. \\ \left. \& (G(x, x) \vee \bar{G}(y, y) \vee F(x, y) \vee \bar{F}(y, x)) \right)$$

ist allgemeingültig für einen Bereich von drei Individuen, während die Normalform $\mathfrak{N}'(x, y)$, in deren symbolischer Darstellung (bei geeigneter Numerierung) die Koeffizienten

$$C_{12}, C_{13}, C_{14}, C_{23}, C_{24}, C_{32}, C_{34}$$

gleich 1, die übrigen gleich 0 zu setzen sind, nur die Gleichungen (1), nicht aber (2) erfüllt.

Solche besonderen Fälle von Allgemeingültigkeit für weniger als vier Individuen, bei denen die Bedingungen (1) ... (m) nicht erfüllt werden, sind gegenüber dem Hauptfall dadurch ausgezeichnet, daß bei ihnen der Charakter der Relationen als Funktionen *zweier* Argumente wesentlich ist.

Die Entscheidung über die Allgemeingültigkeit einer Formel für weniger als vier Individuen ist natürlich durch endliches Ausprobieren zu gewinnen.

In den hier behandelten Fällen des Entscheidungsproblems beruht der Erfolg des angewandten Verfahrens auf dem Umstand, daß von einer gewissen Individuenzahl an, welche durch den Typus der Formel und die Anzahl der in ihr vorkommenden Funktionszeichen bestimmt ist, keine neue Bedingung für die Allgemeingültigkeit hinzutritt.

Dieser Umstand ist nun durchaus auf spezielle Formeltypen beschränkt. Wie schon im § 1 erwähnt wurde, gibt es ja Formeln — und zwar bereits solche mit nur vier voranstehenden Zeichen —, die für jeden endlichen Individuenbereich allgemeingültig sind, dagegen nicht mehr für einen unendlichen.

Bei der Betrachtung solcher Formeln, für deren Allgemeingültigkeit (bzw. Erfüllbarkeit) die unendlichen Individuenbereiche eine wesentliche Rolle spielen, macht sich die Problematik des Unendlichen geltend, und wir kommen damit an die Stelle, wo das Entscheidungsproblem mit den prinzipiellen Fragen der Grundlegung der Mathematik verflochten ist.

(Eingegangen am 24. 3. 1927.)

Über die Definition durch transfiniten Induktion und verwandte Fragen der allgemeinen Mengenlehre.

Von

J. v. Neumann in Göttingen.

Einleitung.

1. Der Begriff der Ordnungszahl und die Möglichkeit und Eindeutigkeit der Definition durch transfiniten Induktion gehören gewiß zu den charakteristischsten Merkmalen der allgemeinen Mengenlehre. Sie sind es in erster Linie, die über die mit der Analysis oder Topologie verwandten Probleme und Methoden der Punktmengentheorie hinausführen ins oft (und nicht ganz grundlos) als uferlos und paradox bezeichnete, aber nichtsdestoweniger recht interessante Gebiet der allgemeinen, abstrakten Mengenlehre. Trotz ihrer fundamentalen Bedeutung haben aber diese Begriffsbildungen weder in der naiven, noch in der seit 1908 entstandenen formalistischen Mengenlehre eine erschöpfende und strenge Begründung erfahren. Dies muß um so sonderbarer erscheinen, als sie einer solchen, selbst in der naiven Mengenlehre, fähig sind.

(Für die Definition durch finite — vollständige — Induktion hat hingegen bereits Dedekind die Notwendigkeit eines strengen Beweises ihrer Möglichkeit erkannt. Er hat sie auch — unabhängig von der damals entstehenden Cantorschen Mengenlehre — bewiesen und durch geeignete Gegenbeispiele dargetan, daß sie mit dem Beweisen durch finite Induktion nicht verwechselt werden darf¹.)

Die seit Cantor übliche (und bis zum scharfen Auftreten der Kritik der Mengenlehre rücksichtslos geübte) Definition der Ordnungszahlen durch Abstraktion einer Eigenschaft gewisser Mengen (nämlich aller, die einer

¹) R. Dedekind, Was sind und was sollen die Zahlen? Braunschweig, mehrere Auflagen seit 1887. Vgl. insbesondere Satz 126 und die daran anschließende Bemerkung 130.

gegebenen, wohlgeordneten, Menge ähnlich sind) ist in einer formalistischen Mengenlehre jedenfalls gänzlich unbrauchbar, und auch in der naiven Mengenlehre wäre ihr wohl eine direkte, eigentliche Definition (die die Ordnungszahlen als wohldefinierte Mengen eindeutig festlegt) vorzuziehen. Der Nachweis, daß dies möglich ist, ist ein Ziel dieser Arbeit²⁾.

Die *Möglichkeit* der Definition durch transfiniten Induktion ist auch alles andere als selbstverständlich. (Nicht so die Eindeutigkeit; diese ist, ebenso wie das Beweisen durch transfiniten Induktion, eine unmittelbare Folgerung aus der Grundeigenschaft der Wohlordnung.) Denn die direkte Konstruktion einer, der transfinit-induktiven Definition genügenden Funktion würde das sukzessive Ausführen unendlich (in der Regel überabzählbar) vieler Schritte erfordern; es ist also ebensowenig unmittelbar-evident wie etwa das Auswahlprinzip. Während aber das Auswahlprinzip seit seiner Entdeckung stets als unabhängiges logisches Prinzip angesehen wurde, war dies bei der Definition durch transfiniten Induktion nie der Fall, und wie hier gezeigt werden soll, ist dies auch nicht notwendig: denn die Erfüllbarkeit solcher Definitionen kann bewiesen werden³⁾.

2. Die vorliegende Arbeit soll möglichst allgemein gehalten bleiben, unabhängig von dem Umstande, daß es heute eigentlich keine einheitliche allgemeine Mengenlehre gibt: sondern eine naive, eine intuitionistische, sowie mehrere formalistisch-axiomatische Systeme.

Dementsprechend werden wir in den Teilen II, III die Sprache der naiven Mengenlehre benützen, aber stets parallel darauf hinweisen, wie unsere Überlegungen in die Zermelo-Fraenkelsche Mengenlehre⁴⁾ zu übertragen sind. Es sei noch darauf hingewiesen, daß der Verf. ein vom Zermeloschen wesentlich verschiedenes System der axiomatischen Mengen-

²⁾ Die Aufstellung der Ordnungszahlenreihe (in der naiven Mengenlehre) habe ich bereits in meiner Arbeit „Zur Einführung der transfiniten Ordnungszahlen“ (Acta litt. ac. sc. reg. univ. Hung. Franc. Jos. Szeged 1 (1923), 4., Seite 199–208) durchgeführt. (Ein mit dem meinen nahe verwandter Ordnungszahlenbegriff war, wie ich nachträglich aus mündlichen Mitteilungen erfuhr, Zermelo 1916 bekannt gewesen. Indessen konnte damals der Fundamentalsatz, wonach es zu jeder wohlgeordneten Menge eine ähnliche Ordnungszahl gibt – vgl. A. 5. – nicht streng bewiesen werden, da das Ersetzungsaxiom unbekannt war.) Auch hatte ich dort die Notwendigkeit des Ersetzungsaxioms festgestellt.

³⁾ Eine Skizze des Beweises habe ich schon in meiner Arbeit ²⁾ gegeben, allerdings für die Ordnungszahlenreihe; indessen kann er in jeder wohlgeordneten Menge fast wörtlich gleich geführt werden.

⁴⁾ Das Axiomensystem von Zermelo (Untersuchungen über die Grundlagen der Mengenlehre I, Math. Ann. 65 (1908)) wurde von Fraenkel präzisiert und abgeschlossen (Unters. über die Grundl. d. Mengenl., Math. Zeitschr. 22 (1925)); Axiomatische Theorie der geordneten Mengen, Journal f. Math. 155 (1926)). Beim Zitieren und bei der Bezeichnungsweise soll für uns die letztere Arbeit von Fraenkel maßgebend sein.

lehre angegeben hat⁵⁾, in dem diese Überlegungen auch angestellt werden können. Da jedoch dies an einem anderen Orte behandelt wird⁶⁾, soll darauf hier nicht Bezug genommen werden.

Andere Systeme der formalen Mengenlehre, so von Russell, Hilbert und ihren Schülern, sind vorderhand noch nicht über die ersten Mächtigkeiten ausgebaut, kommen also hier nicht in Betracht.

3. Die Einteilung der Arbeit ist diese: Zuerst werden die zu Zermelo-Fraenkel notwendigen Zusätze diskutiert⁷⁾ (I); dann wird der Ordnungszahlenbegriff aufgestellt und seine Grundeigenschaften bewiesen (II), und schließlich die Möglichkeit (und Eindeutigkeit) der Definition durch transfinite Induktion bewiesen (III).

I.

Die formalen Hilfsmittel zur Aufstellung des Ordnungszahlenbegriffes⁸⁾.

1. Wir werden in II. die Ordnungszahlen als eine Klasse wohlgeordneter Mengen definieren und beweisen, daß zu jeder wohlgeordneten Menge eine und nur eine ähnliche in dieser Klasse vorkommt (diese ist dann ihre Ordnungszahl). Es wird bei diesem Beweise wiederholt und wesentlich darauf ankommen, gewisse Systeme von Dingen, die auf gegebene Mengen ein-eindeutig bezogen sind, als Mengen zu erweisen (im Sinne der Fraenkelschen Axiomatik). Bei solchen Beweisen ist aber jedenfalls zu erwarten, daß das sog. Fraenkelsche Ersetzungsaxiom herangezogen werden muß⁹⁾: denn dieses ermöglicht in der Regel den Nachweis, daß ein ein-eindeutiges Bild einer Menge wieder eine Menge ist.

In der Tat wird es notwendig sein, das Ersetzungsaxiom heranzuziehen, um die Schlüsse von II. ins genannte formale System übersetzen zu können. Indessen ist es außerdem noch notwendig, ein weiteres Prinzip zum

⁵⁾ J. v. Neumann, Eine Axiomatisierung der Mengenlehre; *Journal f. Math.* 154 (1925).

⁶⁾ J. v. Neumann, Die Axiomatisierung der Mengenlehre, erscheint demnächst in der *Math. Zeitschrift* und enthält die Herleitung der Hauptergebnisse der allgemeinen Mengenlehre aus dem genannten Axiomensystem.

⁷⁾ Der Teil I kann von denen, die sich nur für die Resultate von II, III in der naiven Mengenlehre interessieren, überschlagen werden; ebenso die Fußnoten der Teile II, III.

⁸⁾ Vgl. Fußnote 7).

⁹⁾ Das Ersetzungsaxiom wurde zuerst von Fraenkel formuliert (Zu den Grundlagen der Cantor-Zermeloschen Mengenlehre, *Math. Ann.* 86 (1922)), in sein Axiomensystem nahm er es aber zunächst nicht auf.

Fraenkelschen Systeme hinzuzufügen. Es erweist sich nämlich die Fraenkelsche Definition der Funktionen als zu eng: um vom Ersetzungsaxiom wirklich Gebrauch machen zu können, müssen wir in der Lage sein, wesentlich mehr Funktionen bilden zu können, als es im unveränderten Systeme der Fall ist. (Die Notwendigkeit eines weiteren Zusatzes ist mir im Briefwechsel mit Herrn Fraenkel klar geworden.)

2. Wir formulieren zuerst die beiden genannten Zusätze:

Ersetzungsaxiom. Wenn M eine Menge ist, und f eine Funktion, so gibt es eine Menge M' , deren Elemente sämtliche $f(x)$, $x \in M$ sind, und nur diese¹⁰⁾.

Ferner wäre die Fraenkelsche Definition der Funktion¹¹⁾ folgendermaßen abzuändern:

In der Definition 1 b verlangt Fraenkel: M sei eine Menge, $\varphi(y)$, $\psi(y)$ zwei Funktionen, \circ eine der Relationen $=$, \neq , ε , ε , in M , φ , ψ gehe noch die Hilfsvariable x ein. Dann ist die Menge $M_{\varphi(-) \circ \psi(-)}$ („Menge aller $y \in M$, für die $\varphi(y) \circ \psi(y)$ gilt“) eine Funktion von x . An Stelle hiervon verlangen wir:

F. $\varphi(y)$, $\psi(y)$ seien zwei Funktionen, \circ eine der Relationen $=$, \neq , ε , ε ; in φ , ψ gehe noch die Hilfsvariable x ein. Wenn für jedes x eine Menge existiert, deren Elemente alle y mit $\varphi(y) \circ \psi(y)$ sind, und nur diese, so ist diese Menge eine Funktion von x ¹²⁾. —

Man sieht sofort, in welcher Richtung diese Verallgemeinerung liegt: in der ursprünglichen Fraenkelschen Definition wird der Funktionscharakter unserer von x abhängigen Menge davon abhängig gemacht, ob eine bereits als Funktion erkannte Majorante (d. h. eine Obermenge) vorliegt; während jetzt nur der Mengencharakter für alle x verlangt wird. **F.** erzeugt also mehr Funktionen.

Ehe wir zur Diskussion unseres eigentlichen Gegenstandes übergehen, soll noch gezeigt werden, daß das Ersetzungsaxiom allein, ohne **F.**, gar keine wesentliche Erweiterung des Fraenkelschen Systems ist. Im Gegenteil: es folgt aus Fraenkels übrigen Axiomen. (Nach Hinzufügen von **F.** zu denselben ist das natürlich nicht mehr der Fall: **F.** erweitert ja den Funktionsbegriff weitgehend.)

¹⁰⁾ Es würde genügen, die Existenz einer Menge zu postulieren, die alle diese Elemente (und vielleicht noch andere) enthält; die Existenz einer Menge mit genau diesen Elementen folgt dann aus den übrigen Axiomen. (Vgl. 12 § 2, Fraenkel, Untersuchungen...).

¹¹⁾ Definition s. auf S. 132—133 der zitierten Fraenkelschen Arbeit.

¹²⁾ Es würde genügen, dies allein für die Relation $=$ zu verlangen. Fraenkel verlangt neuerdings nur ε und ε .

3. Wir wollen zeigen: wenn die Fraenkelschen Axiome gelten, und die Funktionen durch die unveränderte Definition beschrieben sind (vgl. Fußnote ¹¹⁾), so gilt das Ersetzungsaxiom.

Hierzu genügt es, zu beweisen: wenn M eine Menge ist, und $f(x)$ eine Funktion, so gibt es eine Menge $N = N(M, f)$, so daß alle $f(x)$, $x \in M$, zu N gehören (vgl. Fußnote ¹⁰⁾).

Zu diesem Zwecke schließen wir induktiv, parallel mit der (induktiven) Fraenkelschen Funktionendefinition.

Erstens ist nach Fraenkel x und jede Konstante c eine Funktion (von x). Hier gilt natürlich unsere Behauptung: wir setzen $N = M$ bzw. $N = (c)$.

Zweitens sind $\mathbb{U}(x)$ und $\mathfrak{S}(x)$ (Potenzmenge und Vereinigungsmenge) Funktionen. Hier können wir

$$N = \mathbb{U}\mathbb{U}\mathfrak{S}M \quad \text{bzw.} \quad N = \mathbb{U}\mathfrak{S}\mathfrak{S}M$$

setzen: denn aus $x \in M$ folgt offenbar

$$\mathbb{U}x \in \mathbb{U}\mathbb{U}\mathfrak{S}M, \quad \mathfrak{S}x \in \mathbb{U}\mathfrak{S}\mathfrak{S}M.$$

Drittens: seien $f(x)$, $g(x)$ zwei Funktionen, so sind auch $(f(x), g(x))$ und $f(g(x))$ Funktionen. Wenn unsere Behauptung für f, g bereits gilt, so gilt sie auch für die neuen Funktionen; und zwar mit

$$N = \mathbb{U}(N(f, M) + N(g, M)) \quad \text{bzw.} \quad N = N(f, N(g, M)).$$

Das letztere ist trivial, das erstere gilt, da aus

$$\begin{aligned} f(x) \in N(f, M), \quad g(x) \in N(g, M) \\ (f(x), g(x)) \in \mathbb{U}(N(f, M) + N(g, M)) \end{aligned}$$

folgt.

Viertens sei $M(x)$ eine Funktion, und seien $\varphi(x, y)$, $\psi(x, y)$ zwei Funktionen von y , in die auch x eingeht, und \circ sei eine der Relationen $=$, \neq , ε , ξ . Dann ist auch $M(x)_{\varphi(x, \bar{y}) \circ \psi(x, \bar{y})}$ eine Funktion von x . Wenn unsere Behauptung für $M(x)$ bereits gilt, so gilt sie auch für $M(x)_{\varphi(x, \bar{y}) \circ \psi(x, \bar{y})}$. In der Tat: aus $x \in M$ folgt

$$M(x) \in N(M, M), \quad M(x)_{\varphi(x, \bar{y}) \circ \psi(x, \bar{y})} \in M(x),$$

also

$$M(x)_{\varphi(x, \bar{y}) \circ \psi(x, \bar{y})} \in \mathbb{U}\mathfrak{S}N(M, M).$$

Also können wir

$$N = \mathbb{U}\mathfrak{S}N(M, M)$$

setzen.

Da es im Fraenkelschen Systeme keine weiteren Funktionen gibt, ist unsere Behauptung damit restlos bewiesen.

II.

Die Ordnungszahlen.

A. Definition, Existenzbeweis.

1. Es sei noch einmal darauf hingewiesen, daß wir die Terminologie der naiven Mengenlehre benützen werden; die zur Übertragung in den Zermelo-Fraenkelschen Formalismus nötigen Zusätze sollen in Fußnoten angedeutet werden.

Von allgemein mengentheoretischen Begriffen setzen wir nur die der Wohlordnung und der Ähnlichkeit voraus; jedoch keinen einzigen der auf sie bezüglichen Sätze (Vergleichbarkeit, Wohlordnungssatz). Unsere Überlegungen sind unabhängig vom Wohlordnungssatz und von der Existenz einer unendlichen Menge.

Unsere Bezeichnungen sind die folgenden:

$x \in M$: x ist Element von M . $M \subseteq N$, $N \supseteq M$: M ist Teilmenge von N . $M < N$, $N > M$: M ist echte Teilmenge von N . Wenn $f(x)$ eine Funktion ist, und $E(x)$ eine Eigenschaft, $M(f(x); E(x))$: die Menge der $f(x)$, wenn x alle Dinge mit $E(x)$ durchläuft¹³).

Wenn M eine geordnete Menge ist, und $x \in M$, $y \in M$, so bedeutet $x < y$: x ist (in der betreffenden Ordnung) vor y , $A(x; M)$: $M(y; y < x)$ (der Abschnitt vor x)¹⁴).

Die Ordnungszahlen (kurz: *OZ*) werden wir als gewisse wohlgeordnete Mengen definieren; und zwar werden wir ihnen eine derartige weitere Beschränkung auferlegen, daß zu jeder wohlgeordneten Menge eine und nur eine ähnliche *OZ* existiert. Eine andere fundamentale Eigenschaft dieser *OZ* ist, daß jede unter ihnen die Menge aller vorangehenden *OZ* ist. (Vgl. B. 1 sowie das Folgende.)

Unsere Definition lautet so:

Eine wohlgeordnete Menge M ist dann und nur dann *OZ*, wenn für alle $x \in M$

$$x = A(x; M)$$

¹³ Eine solche Menge braucht natürlich nicht immer zu existieren. Wenn aber $M(x; E(x))$ (die Menge aller x , für die $E(x)$ gilt) existiert, so folgt ihr Vorhandensein aus dem Ersetzungsaxiom; und nur in diesem Falle werden wir sie im folgenden benutzen

¹⁴ Existiert nach Fraenkel (Axiomatische Theorie...),⁵ da es aus der dort (16) definierten Menge μ_x (in Fraenkels Bezeichnung μ_m) leicht zu erhalten ist:

$$A(x; M) = M - \mu_x - \{x\}.$$

gilt (d. h. jedes Element sein eigener Abschnitt in M ist)¹⁵⁾. Dieser *OZ*-Begriff soll nun näher untersucht werden.

2. Wir weisen zuerst einige Eigenschaften der *OZ* nach, die zur einleitenden Orientierung nützlich sind.

Erstens: die Ordnung einer *OZ* ist stets die Subsumtionsordnung (d. h. diejenige, bei der $x < y$ mit $x < y$ gleichbedeutend ist)¹⁶⁾.

In der Tat: $x < y$ ist offenbar mit $A(x; P) < A(y; P)$ (wir bezeichnen die *OZ* mit P, Q, R, S, \dots) gleichbedeutend, und dies, nach Definition der *OZ*, mit $x < y$.

Zweitens: Zwei ähnliche *OZ* müssen identisch sein.

Es sei nämlich P auf Q durch die Abbildung $u = f(x)$ ähnlich bezogen ($x \in P, u \in Q$). Wir werden zeigen: es ist stets $f(x) = x$, dies hat dann $P = Q$ zur Folge.

Wäre das nicht der Fall, so wäre die Menge P' aller $x \in P$ mit $f(x) \neq x$ ¹⁷⁾ nicht leer, sie hätte also ein erstes Element x_0 . Wir setzen $u_0 = f(x_0)$. Dann ist für alle $y < x_0$ jedenfalls $f(y) = y$, also bildet $v = f(y)$ den $A(x_0; P)$ auf sich selber ab. Wegen der Ähnlichkeit der Abbildung muß aber $A(x_0; P)$ auf $A(u_0; P)$ abgebildet werden, also ist

$$A(x_0; P) = A(u_0; P),$$

und weil P, Q *OZ* sind, hat dies $x_0 = u_0$, d. h. $x_0 = f(x_0)$ zur Folge. Und das widerspricht der Annahme über x_0 .

Wir haben damit gezeigt, daß es zu jeder wohlgeordneten Menge höchstens eine ähnliche *OZ* gibt. Im folgenden soll bewiesen werden, daß es auch mindestens eine solche gibt. Diesen Beweis werden wir in drei Schritten führen.

3. Nehmen wir an, die wohlgeordnete Menge M sei einer *OZ* P ähnlich. x_0 sei ein Element von M und $f(x)$ die ähnliche Abbildung von M auf P .

¹⁵⁾ Eine „geordnete Menge“ besteht formal eigentlich aus zwei Dingen, einer Menge und einer Ordnung von ihr. Es ist aber leicht, die zwei durch eines zu ersetzen: man kann z. B. die Ordnung allein angeben, die Menge ist, wie man leicht zeigt, deren Vereinigungsmenge. (Wenn nämlich die Ordnung nach Fraenkel definiert wird.)

Man zeigt leicht, daß $A(x; M)$ eine Funktion von x und M ist: ist es doch eine recht einfach in Abhängigkeit von x und M gebildete Aussonderungsmenge von M . Ferner gibt es eine Funktion $f(M)$, derart, daß $f(M) \neq 0$ dem *OZ*-Charakter von M gleichbedeutend ist. Auch dies ergibt sich aus dem Aussonderungsaxiom mit der bei Fraenkel sonst üblichen Methode.

¹⁶⁾ Der Begriff stammt von Hessenberg.

¹⁷⁾ Diese Menge existiert: sie entsteht durch Aussonderung $P_{f(x) \neq x}$.

Das durch $u = f(x)$ vermittelte Bild von x_0 sei u_0 , dann wird $A(x_0; M)$ auf $A(u_0; P)$ abgebildet und ist ihm ähnlich. $A(u_0; P)$ ist (weil P OZ ist) $= u_0 = f(x_0)$, wir behaupten: *es ist eine OZ.*

Wegen $A(u_0; P) \subset P$ ist es jedenfalls wohlgeordnet, und aus $v \in A(u_0; P)$ folgt

$$A(v; A(u_0; P)) = A(v; P) = v.$$

Also ist es wirklich eine OZ.

Wir haben also bewiesen: $f(x_0)$ ist eine OZ und dem $A(x_0; M)$ ähnlich.

4. Zweitens nehmen wir an, daß M eine wohlgeordnete Menge ist, und zu jedem $A(x; M)$, wo $x \in M$, eine ähnliche OZ existiert. Es soll gezeigt werden: auch zu M gibt es eine ähnliche OZ.

Nach 2. gibt es zu jedem $x \in M$ eine einzige OZ, die dem $A(x; M)$ ähnlich ist, wir nennen sie $f(x)$.¹⁸⁾ Wir halten für einen Augenblick ein $x \in M$ fest. Die ähnliche Abbildung von $A(x; M)$ auf $f(x)$ sei etwa $v = \varphi_x(y)$. Dann ist für ein $y < x$ (d. h. $y \in A(x; M)$) $\varphi_x(y)$ eine OZ, die dem

$$A(y; A(x; M)) = A(y; M)$$

ähnlich ist, also ist

$$\varphi_x(y) = f(y).$$

Hieraus folgern wir erstens: $\varphi_x(y)$ ist ein Element der OZ $f(x)$, also sein eigener Abschnitt in ihr, also eine echte Teilmenge von $f(x)$; also gilt dasselbe von $f(y)$. D. h.: aus $y < x$ folgt $f(y) \subset f(x)$. Auch die Umkehrung gilt: denn wenn nicht $y < x$ ist, so ist $y \geq x$, $x \leq y$, also $f(x) \subset f(y)$, und nicht $f(y) \subset f(x)$.

Zweitens schließen wir so: $f(x)$ ist das durch $v = \varphi_x(y)$, d. h. das durch $v = f(y)$ vermittelte Bild von $A(x; M)$. Es ist also die Menge aller $f(y)$ mit $y < x$. Nach dem Obigen ist aber $y < x$ mit $f(y) \subset f(x)$ gleichbedeutend, wir können also auch sagen: $f(x)$ ist die Menge aller $f(y)$ mit $f(y) \subset f(x)$.

Nun werde die Menge

$$P = M(f(x); x \in M)$$

¹⁸⁾ Es fehlt natürlich der Beweis, daß $f(x)$ wirklich eine Funktion ist. Diesen führt man so:

Daß P eine zu $A(x; M)$ ähnliche OZ ist, kann man mit der bei Fraenkel üblichen Methode auf die Form $g(x, P) \neq 0$ bringen. (Die Abhängigkeit von M mag in g selbst enthalten sein.)

Da für jedes $x \in M$ ein einziges P der Bedingung $g(x, P) \neq 0$ genügt, bilden diese P eine Menge (ein Element bildet stets eine Menge!), wir können also das Zusatz-Prinzip F. anwenden. Dann ist diese Menge gleich $h(x)$, und ihr einziges Element $\in (h(x))$. Der letztere Ausdruck ist aber leicht in eine Funktion $f(x)$ zu überführen.

gebildet¹⁹⁾. Da $y < x$ mit $f(y) < f(x)$ gleichbedeutend ist, ist P in der Subsumtionsordnung dem M ähnlich.

Wir behaupten aber: P OZ . In der Tat: erstens ist P dem M ähnlich, also wohlgeordnet, und zweitens gilt für jedes $u \in P$

$$u = f(x) = M(f(y); f(y) < f(x)) = M(v; v \in P, v < u) = A(u; P).$$

Damit ist unsere Behauptung restlos bewiesen.

5. Wir vollziehen nun den dritten und letzten Schritt, indem wir zeigen: zu jeder wohlgeordneten Menge M gibt es eine ähnliche OZ .

Gäbe es keine zu M ähnliche OZ , so müßte es unter den Elementen x von M nach 4. auch solche geben, daß keine zu $A(x; M)$ ähnliche OZ existiert. Die Menge aller dieser x sei N ,²⁰⁾ nach Annahme ist N nicht leer.

Es ist aber $N \notin M$ und M wohlgeordnet, also besitzt N ein erstes Element x_0 . Für jedes $y \in A(x_0; M)$, d. h. $y < x_0$, ist $y \in N$, also gibt es eine zu

$$A(y; M) = A(y; A(x_0; M))$$

ähnliche OZ . Nach 4. muß es aber dann auch eine zu $A(x_0; M)$ ähnliche OZ geben, entgegen $x_0 \in N$.

Damit ist alles bewiesen.

B. Haupteigenschaften.

1. Wir werden beweisen: Eine OZ P ist die Menge aller OZ Q , für welche $Q \subset P$ gilt. D. h.: $Q \in P$ ist mit Q OZ , $Q \subset P$ gleichbedeutend. Erstens sei $Q \in P$. Dann ist

$$Q = A(Q; P) \subset P.$$

Ferner ist Q (als Teilmenge von P) wohlgeordnet, und es gilt für jedes $x \in Q$:

$$A(x; Q) = A(x; A(Q; P)) = A(x; P) = x.$$

Also besitzt Q die für die OZ charakteristischen Eigenschaften, es ist Q eine OZ .

Zweitens sei Q OZ , $Q \subset P$. Wenn $x \in Q$, $y \in P$ und $y < x$ ist, so gehört y zu $A(x; P)$; da aber P, Q beide OZ sind, so ist

$$A(x; P) = x = A(x; Q) \subset Q.$$

Also gehört auch y zu Q . Nun ist P eine wohlgeordnete Menge, und

¹⁹⁾ Ihre Existenz folgt aus dem Ersetzungs-Axiom.

²⁰⁾ Daß die Menge N existiert, zeigt man so: Wir gehen wieder von der Funktion $g(x, P)$ in Fußnote ¹⁸⁾ aus. Für jedes $x \in M$ gibt es keine zu $A(x; M)$ ähnliche OZ oder genau eine, also kein P mit $g(x, P) \neq 0$ oder genau eines. Jedenfalls bilden diese P eine Menge, und diese ist nach F. gleich $h(x)$. Diejenigen $x \in M$, für die keine zu $A(x; M)$ ähnliche OZ existiert, sind durch $h(x) = 0$ gekennzeichnet, also ist $N = M_{h(x)=0}$.

$Q < P$, und aus $x \varepsilon Q$, $y \varepsilon P$, $y < x$ folgt $y \varepsilon Q$. Hieraus kann man bekanntlich folgern, daß Q der Abschnitt von einem Elemente x_0 von P ist. Dann ist aber:

$$Q = A(x_0; P) = x_0,$$

also $Q \varepsilon P$.

Damit ist alles bewiesen. Insbesondere wissen wir jetzt, daß für zwei OZ P, Q $P \varepsilon Q$ mit $P < Q$ gleichbedeutend ist. $P \varepsilon P$ ist also ausgeschlossen.

2. Wir zeigen weiter: Von zwei verschiedenen OZ P, Q ist eine gewiß echte Teilmenge der anderen.

Zu diesem Zwecke betrachten wir den Durchschnitt R von P und Q . Sowohl P wie Q sind als OZ in der Subsumtionsordnung gegeben, also stimmen ihre Ordnungen in R überein. Als Teilmenge von ihnen ist dann auch R wohlgeordnet.

Ferner sei x ein Element von R . Da P, Q OZ sind, ist

$$x = A(x; P), \quad x = A(x; Q),$$

und $A(x; R)$ ist der Durchschnitt beider, also auch gleich x . Damit haben wir aber gezeigt, daß R eine OZ ist.

Nun ist $R \notin P$, $R \notin Q$. Wenn $R = P$ oder $R = Q$ ist, so ist $P \notin Q$ bzw. $Q \notin P$, wegen $P \neq Q$ also $P < Q$ oder $Q < P$; und dann ist die Behauptung bewiesen.

$R \neq P$, $R \neq Q$ aber ist unmöglich. Dann wäre nämlich $R < P$, $R < Q$, wegen R, P, Q OZ also $R \varepsilon P$, $R \varepsilon Q$, d. h. $R \varepsilon R$. Und dies ist wegen R OZ ausgeschlossen.

3. M sei irgendeine Menge von OZ ; nach dem soeben Bewiesenen kann M durch Subsumtion geordnet werden. Wir wollen zeigen, daß es in der Subsumtionsordnung wohlgeordnet ist.

Es gilt also zu zeigen: wenn $N \notin M$, $N \neq 0$ ist, so hat N ein erstes Element.

Wegen $N \neq 0$ können wir ein $P \varepsilon N$ auswählen. Sind alle $Q \varepsilon N$ auch $Q \geq P$, so ist P das gesuchte erste Element. Ist das nicht der Fall, so betrachten wir diejenigen Elemente Q von N , für die nicht $Q \geq P$ ist. Nach Annahme gibt es solche.

Da P, Q OZ sind, ist nicht $Q \geq P$ gleichbedeutend mit $Q < P$, d. h. mit $Q \varepsilon P$. Die genannten Q machen also gerade den Durchschnitt von N und P aus. Wir nennen diesen Durchschnitt N^* .

Da $N^* \notin P$, $N^* \neq 0$ ist, und P (in der Subsumtionsordnung) wohlgeordnet ist, hat N^* ein erstes Element, P^* . Wegen $P^* \varepsilon N^*$ ist insbesondere $P^* \varepsilon P$, $P^* \varepsilon N$, also P^* OZ , $P^* < P$, $P^* \varepsilon N$. Wir behaupten nun: P^* ist das erste Element von N .

Denn für ein $Q \in N$ ist entweder $P \in Q$, also $P^* \subset P \in Q$, oder aber $Q \in N^*$, also $P^* \in Q$.

4. Wir behaupten: P ist dann und nur dann eine OZ , wenn aus $P \in P$ folgt: $P \subset OZ$, $P \in P$.

Die Bedingung ist notwendig, denn aus $P \subset OZ$ folgt für $P \in P$ sogar $P \subset OZ$, $P \subset P$.

Daß sie hinreichend ist, zeigen wir so: Als Menge von OZ ist P in der Subsumtionsordnung wohlgeordnet. Es bleibt übrig, für alle $P \in P$

$$P = A(P; P)$$

zu beweisen. Nun ist für $P \in P$ wegen $P \in P$

$$P = M(Q; Q \in P) = M(Q; Q \in P, Q \in P)$$

und dies ist wegen $P \subset OZ$

$$= M(Q; Q \subset OZ, Q \subset P, Q \in P);$$

dies ist aber, da aus $Q \in P$ $Q \subset OZ$ folgt,

$$= M(Q; Q \subset P, Q \in P) = A(P; P).$$

Damit ist der Beweis vollendet.

C. Grundoperationen mit OZ .

1. Mit Hilfe des Kriteriums in B. 4. können wir einfach die wesentlichsten Erzeugungsweisen für die OZ gewinnen.

Erstens stellen wir fest: 0 ist eine OZ . In der Tat: 0 hat überhaupt keine Elemente, es ist also nach B. 4. eine OZ .

Zweitens behaupten wir: Wenn M eine Menge von OZ ist, so ist $\mathfrak{S}(M)$ eine OZ . Denn aus $P \in \mathfrak{S}(M)$ folgt $P \in Q$, $Q \in M$, also ist $Q \subset OZ$, und folglich $P \subset OZ$, $P \subset Q \in \mathfrak{S}(M)$.

Drittens behaupten wir: wenn P eine OZ ist, so ist auch $P + \{P\}$ eine OZ . Denn aus $Q \in (P + \{P\})$ folgt $Q \in P$ oder $Q \in \{P\}$. Im ersten Falle ist $Q \subset OZ$, $Q \subset P \in P + \{P\}$, im zweiten $Q = P$, also wieder $Q \subset OZ$, $Q \in P + \{P\}$.

Die OZ $P + \{P\}$ wollen wir (nach Möglichkeit im Anschluß an die übliche Bezeichnung) mit $P \oplus 1$ bezeichnen.

2. Wir behaupten: In der Subsumtionsordnung ist 0 die erste OZ , und $P \oplus 1$ die auf P unmittelbar folgende.

Das erste ist trivial. Das zweite beweisen wir so: Zunächst ist $P \in (P + \{P\})$, $P \in P \oplus 1$, also $P \subset P \oplus 1$. Ferner folgt aus $P \subset Q$ ($Q \subset OZ$) auch $P \oplus 1 \in Q$. Sonst wäre nämlich $P \oplus 1 \supset Q$, also

$$P \subset Q \subset P + \{P\},$$

was unmöglich ist, da $P + \{P\}$ nur ein Element mehr hat als P . Damit ist alles bewiesen.

Weiter zeigen wir: $P \supseteq Q$ ist gleichbedeutend bzw. $P \oplus 1 \supseteq Q \oplus 1$ (P, Q OZ).

Es genügt zu zeigen, daß aus $P \supseteq Q$ folgt $P \oplus 1 \supseteq Q \oplus 1$; die Umkehrung muß gelten, da diese drei Fälle (nach B. 2.) eine vollständige Disjunktion bilden.

Es genügt ferner aus $P < Q$ $P \oplus 1 < Q \oplus 1$ zu folgern: denn aus $P = Q$ folgt jedenfalls $P \oplus 1 = Q \oplus 1$, und $>$ entsteht aus $<$ durch Vertauschen von P, Q . Aus $P < Q$ folgt aber in der Tat

$$P \oplus 1 \in Q < Q \oplus 1.$$

3. Schließlich zeigen wir: Wenn P eine OZ ist und es auf die Form $P = R \oplus 1$ (R OZ) gebracht werden kann, so ist dies für $R = \mathfrak{S}(P)$ der Fall, und nur für dieses. Ist dies aber nicht möglich, so ist $P = \mathfrak{S}(P)$. Im letzteren Falle nennen wir P eine Limeszahl.

Die erste Behauptung kommt offenbar darauf heraus, $R = \mathfrak{S}(R \oplus 1)$ zu zeigen. Nun ist

$$\mathfrak{S}(R \oplus 1) = \mathfrak{S}(R + \{R\}) = \mathfrak{S}(R) + R.$$

Da aus $S \varepsilon R$ $S \in R$ folgt, ist $\mathfrak{S}(R) \in R$, folglich ist

$$\mathfrak{S}(R) + R = R.$$

Die zweite Behauptung besagt: aus $\mathfrak{S}(P) \neq P$ folgt $P = \mathfrak{S}(P) \oplus 1$. Ebenso wie vorhin zeigt man $\mathfrak{S}(P) \in P$, wegen $\mathfrak{S}(P) \neq P$ muß also $\mathfrak{S}(P) < P$, $\mathfrak{S}(P) \oplus 1 \in P$ sein. Wäre nun $\mathfrak{S}(P) \oplus 1 \neq P$, so müßte $\mathfrak{S}(P) \oplus 1 < P$ sein, wir hätten also:

$$\mathfrak{S}(P) < \mathfrak{S}(P) \oplus 1 < P,$$

$$\mathfrak{S}(P) \varepsilon (\mathfrak{S}(P) \oplus 1), \quad (\mathfrak{S}(P) \oplus 1) \varepsilon P,$$

$$\mathfrak{S}(P) \varepsilon \mathfrak{S}(P).$$

Dies ist aber mit $\mathfrak{S}(P)$ OZ unvereinbar.

D. Anwendungen.

1. Der vorliegende OZ-Begriff erlaubt, wenn man sich auf endliche OZ beschränkt, einen dem Dedekindschen²¹⁾ analogen Aufbau der Theorie der positiven ganzen (endlichen) Zahlen, ohne Voraussetzung der Existenz einer unendlichen Menge. (In meiner Arbeit⁶⁾, Kap. VIII. 4. ist das etwas ausführlicher dargestellt. Vgl. auch die Arbeit⁵⁾, S. 237.) Ohne hierauf

²¹⁾ Vgl. Fußnote 1).

näher einzugehen, wollen wir an Hand der Resultate aus C. 2. die ersten OZ angeben. (Man sieht direkt das Zutreffen des Satzes B. 1., wonach jede die Menge aller vorangehenden ist.)

Gewöhnliche							
Bezeichnung	0	1	2	3	4		
Systematische							
Bezeichnung	0	$0 \oplus 1$	$0 \oplus 1 \oplus 1$	$0 \oplus 1 \oplus 1 \oplus 1$		$0 \oplus 1 \oplus 1 \oplus 1 \oplus 1$	
Die OZ . . .	0	$\{0\}$	$\{0, \{0\}\}$	$\{0, \{0\}, \{0, \{0\}\}\}$	$\{0, \{0\}, \{0, \{0\}\}, \{0, \{0\}, \{0, \{0\}\}\}$		

Wie man sieht, sind sie etwas komplizierter, als die Zermeloschen „Zahlen“ ²²⁾

$$0, \{0\}, \{\{0\}\}, \{\{\{0\}\}\}, \{\{\{\{0\}\}\}\}, \dots,$$

aber sie können dafür ungehindert über ω hinaus fortgesetzt werden.

2. Wenn wir die zur wohlgeordneten Menge M ähnliche OZ kurz als die OZ von M bezeichnen, so ist es klar, daß zwei wohlgeordnete Mengen M, N dann und nur dann ähnlich sind, wenn sie dieselbe OZ haben.

Weiter ist jede Menge ihrer OZ ähnlich, also ist M dann und nur dann einem Abschnitte von N ähnlich, wenn die OZ von M, P , einem Abschnitte

$$A(R; Q) = R$$

der OZ von N, Q , ähnlich ist. Da auch R OZ ist, besagt das $P = R$, d. h. $P \varepsilon Q$, oder was dasselbe ist, $P < Q$.

Auf Grund dieser zwei Bemerkungen sieht man ohne weiteres ein, daß der Satz in B. 2. die allgemeine Vergleichbarkeit der wohlgeordneten Mengen enthält.

3. Durch unser OZ -System ist auch das Problem der Angabe der Alephs gelöst, wir müssen nur den Begriff der Anfangszahl oder Kardinalzahl (kurz: KZ) einführen.

Wir nennen eine OZ P eine KZ , wenn für keine OZ $Q < P$ Q und P äquivalent sind; d. h. wenn P mit keinem seiner Elemente äquivalent ist ²³⁾.

²²⁾ Vgl. Fußnote 4).

²³⁾ Es gibt eine Funktion f , so daß $f(P) = 0$ mit P KZ gleichbedeutend ist. Um sie zu konstruieren, brauchen wir nur eine Funktion $g(M, N)$, so daß $g(M, N) \neq 0$ die Äquivalenz von M und N andeutet: dann leistet $f(P) = P_g(P, x) \neq 0$ das Gewünschte. Für elementfremde Mengen M, N leistet dies zunächst die Funktion $g'(M, N) = \Xi$ (vgl. Fraenkel, Untersuchungen . . ., Seite 261, 13), wo Ξ die Menge aller ähnlichen Abbildungen von M, N ist; man sieht leicht den Funktionscharakter von Ξ ein. Für nicht-elementfremde M, N muß es ein zu beiden fremdes L geben, so daß $g'(M, L) \neq 0$, $g'(N, L) \neq 0$ ist (a. a. O. Seite 262, 16); und dieses L kann, wenn überhaupt vorhanden, als Teilmenge von $\mathfrak{U}(M + \mathfrak{U}(M + \mathfrak{S}(M + N)))$ gewählt werden

(Fortsetzung der Fußnote 23 auf nächster Seite.)

Man sieht unschwer ein, daß es zu jeder wohlordenbaren Menge M eine und nur eine äquivalente KZ gibt.

Erstens gibt es höchstens eine: denn sind die KZ P, Q dem M äquivalent, so sind sie es auch untereinander, also ist $P \supset Q$ unmöglich, also ist $P = Q$.

Zweitens gibt es mindestens eine; dabei sind aber, da die „Menge aller OZ “ ein bekanntlich paradoxes Gebilde ist²⁴⁾, gewisse Vorsichtsmaßregeln notwendig.

M sei nämlich eine wohlordenbare Menge, dann gibt es eine zum wohlgeordneten M ähnliche OZ , also gibt es mit M äquivalente OZ überhaupt. P sei eine solche. Wenn kein $Q \varepsilon P$ mit M äquivalent ist, so ist auch keines mit P äquivalent, also ist $P \varepsilon KZ$, und wir sind am Ziele. Im entgegengesetzten Falle bilden wir die Menge P aller mit M äquivalenter $Q \varepsilon P$ ²⁵⁾. Dann ist $P \in P, P \neq 0$.

Da P wohlgeordnet ist, hat P ein erstes Element, P^* . P^* ist mit M äquivalent, wir behaupten aber, daß es eine KZ ist. Denn wäre es mit einem $Q \varepsilon P^*$ äquivalent, so wäre Q mit M äquivalent. Ferner ist $Q \varepsilon P^*, P^* < P$, also $Q \varepsilon P$; folglich ist $Q \varepsilon P$. Und dann muß $Q \supseteq P^*$ sein, entgegen $Q \varepsilon P^*, Q < P^*$.

Damit ist unsere Behauptung vollständig bewiesen. Nehmen wir nun den Wohlordnungssatz hinzu (der ja bisher nicht benutzt wurde), so folgt: zu jeder Menge M gibt es eine und nur eine äquivalente KZ . Also sind die KZ geeignet, die Mächtigkeiten zu vertreten.

III. Die Definition durch transfiniten Induktion.

1. Das Definieren durch transfiniten Induktion ist ein Prozeß, der in der Ordnungszahlenreihe oder auch in einer beliebigen wohlgeordneten Menge Anwendung findet. In diesem letzteren Falle kann seine *Möglichkeit* (falls auch der Wertevorrat der zu definierenden Funktion von vornherein in einer gegebenen Menge liegt) im ungeänderten Fraenkelschen

(folgt leicht aus den Überlegungen a. a. O. Seite 262, 16 und Seite 257, 5). Die Existenz eines solchen L wird also durch

$$\{\{ \cup \cup (M + \cup (M + \mathcal{C}(M + N))) \}_{g'(M \neq 0)} \}_{g'(N \neq 0)} \neq 0$$

ausgesagt, und die linke Seite ist offenbar ein $g(M, N)$.

²⁴⁾ Sie führt zur bekannten Burali-Fortischen Antinomie, die wir hier ziemlich kurz formulieren können. Gäbe es eine Menge Ω , die alle OZ , und nur diese enthält, so folgte aus $P \varepsilon \Omega$ erstens $P \varepsilon OZ$ und zweitens $P \in \Omega$ (aus $Q \varepsilon P$ folgt ja $Q \varepsilon OZ$). Nach **B. 1.** wäre also $\Omega \varepsilon OZ$, d. h. $\Omega \varepsilon \Omega$, entgegen **B. 4.** (Man sieht übrigens, daß das vielumstrittene „Hinzufügen eines Elementes zu Ω “ gar nicht nötig ist, um sich in den Widerspruch zu verwickeln.)

²⁵⁾ Nach ²³⁾ existiert diese Menge, sie ist gleich $P_{g(M, x) \neq 0}$.

Systeme bewiesen werden, und zwar im wesentlichen mit der Methode unseres folgenden Beweises. Wir wollen jedoch die induktive Definition von *OZ*-Funktionen untersuchen, wobei dann die Zusätze aus I notwendig sind. (Dabei brauchen wir gar keine Einschränkungen über den Wertevorrat zu machen.)

Zunächst soll der Prozeß des „Definierens durch transfiniten Induktion“ bzw. die Behauptung seiner allgemeinen und stets eindeutigen Möglichkeit genau beschrieben werden. Und zwar soll dies in möglichst allgemeiner Form erfolgen; wir wählen die folgende *Formulierung*:

J. Wenn f eine Funktion und P eine *OZ* ist, so bezeichnen wir mit $F(f, P)$ den Verlauf der Funktion f in P . Darunter verstehen wir diejenige nur in P (d. h. nur für die $OZ < P$) definierte Funktion, die dort durchweg dieselben Werte hat, wie f .²⁶⁾

Wenn φ eine Funktion (von zwei Variablen) ist, so gibt es eine und nur eine Funktion f (einer Variablen), die nur für die *OZ* definiert ist, und bei der für jede *OZ* P

$$f(P) = \varphi(F(f, P), P)$$

gilt.

(Die obige Relation drückt den Grundgedanken der induktiven Definition aus: f ist an der Stelle P zu berechnen aus P und allen seinen Werten vor P .)

2. Um die *Möglichkeit* der *Definition* durch transfiniten Induktion zu beweisen, werden wir die Behauptung dahin umformen, daß sie eines *Beweises* durch transfiniten Induktion fähig wird. Zu diesem Zwecke führen wir die folgenden Begriffe ein:

Eine *OZ* P ist „normal“, wenn es eine (in P definierte, d. h. für die $OZ < P$ definierte) Funktion f gibt, so daß für alle $Q \in P$ (d. h. $Q \subset OZ$, $Q < P$)

$$f(Q) = \varphi(F(f, Q), Q)$$

ist.

Ein solches f nennen wir ein „Funktionselement bis P “. Und

$$\varphi(F(f, P), P)$$

nennen wir einen „Funktionswert von P “. ($F(f, P)$ ist vom naiven Standpunkte aus dasselbe wie f , da dieses nur in P definiert ist, im axiomatischen Systeme ist es anders (vgl. Fußnote²⁶⁾).

²⁶⁾ Es ist angemessen, diesen „Verlauf“ $F(f, P)$ nicht als Funktion, sondern als Menge einzuführen: denn wir werden ihn (vgl. w. u.) als Argument in eine Funktion einsetzen müssen, und dies ist im Fraenkelschen Systeme nur für Mengen (nicht aber für Funktionen) möglich. Es gilt also, den in P definierten Funktionen eineindeutig Mengen zuzuordnen.

(Fortsetzung der Fußnote 26 auf nächster Seite.)

Wir werden nacheinander zeigen: Wenn P normal ist, so gibt es — bei vorgegebenem φ — genau ein Funktionselement bis P , und folglich auch genau einen Funktionswert von P (3.); alle OZ P sind normal (4.—5.), und hieraus schließen: es existiert eine und nur eine Funktion f mit den in 1. genannten Eigenschaften (5.).

3. Die erste der soeben genannten Behauptungen läuft offenbar darauf hinaus, zu zeigen: Wenn f, g Funktionselemente bis P sind — bei vorgegebenem φ —, so sind sie identisch, d. h. es ist für alle $Q \varepsilon P$

$$f(Q) = g(Q).$$

Wäre das nicht der Fall, so wäre die Menge P aller $Q \varepsilon P$ mit $f(Q) \neq g(Q)$ ²⁷⁾ nicht leer, d. h. $P \in P$, $P \neq 0$. Also hätte P ein erstes Element Q_0 .

Für $Q \varepsilon Q_0$, d. h. Q OZ , $Q < Q_0$ ist also $f(Q) = g(Q)$, so daß $F(f, Q_0) = F(g, Q_0)$ ist, und dies hat

$$\varphi(F(f, Q_0), Q_0) = \varphi(F(g, Q_0), Q_0), f(Q_0) = g(Q_0)$$

zur Folge, entgegen $Q_0 \varepsilon P$.

Für normale OZ sind also Funktionselement und Funktionswert eindeutig festgelegte Begriffe, die wir mit $FE(P)$ und $FW(P)$ bezeichnen wollen.

4. Wir behaupten: Wenn $Q < P$ ist, so folgt aus der Normalität von P auch die von Q , und es ist

$$FW(Q) = FE(P)(Q).$$

($FE(P)$ entspricht einer Funktion, wir können also ihren Wert an der Stelle Q bilden!)

Die gewöhnliche Wiedergabe der Funktion f durch die Paarmenge der $(x, f(x))$ versagt hier, denn die x und die $f(x)$ können nicht von vornherein auf elementfremde Mengen beschränkt werden. Doch wird diese Schwierigkeit leicht umgangen durch Einführung des „geordneten Paares“

$$\langle u, v \rangle = ((u), (u, v)).$$

Man zeigt erstens leicht, daß aus $\langle u, v \rangle = \langle u', v' \rangle$ stets $u = u', v = v'$ folgt; zweitens gibt es offenbar eine Funktion ξ mit

$$\xi(u, v) = \langle u, v \rangle.$$

Wenn nun eine Funktion f und eine OZ (es könnte eine beliebige Menge sein) P gegeben ist, so bilden wir $\langle x, f(x) \rangle$ für alle x von P . Diese bilden eine Menge, denn $\langle x, f(x) \rangle$ kann auf die Form $g(x)$ gebracht werden (g eine von f abhängige Funktion), und dann genügt die Anwendung des Ersetzungsaxioms.

Die so gewonnene Menge nennen wir $F(f, P)$. (Natürlich ist F keine Funktion von f, P im axiomatischen Sinne, f kann ja gar nicht Argument sein.)

²⁷⁾ Entsteht durch Aussonderung $P_{f(x) \neq g(x)}$.

Es sei nämlich $f = FE(P)$ und f^* diejenige nur in Q definierte Funktion, die dort mit f übereinstimmt. Dann ist für $R \varepsilon Q$

$$f^*(R) = f(R) = \varphi(F(f, R), R) = \varphi(F(f^*, R), R).$$

Also ist Q normal und f^* sein Funktionselement; und es ist

$$FW(Q) = \varphi(F(f^*, Q), Q) = \varphi(F(f, Q), Q) = f(Q),$$

womit alles bewiesen ist.

Wir behaupten weiter: Wenn alle $Q \varepsilon P$, d. h. alle $Q < P$, normal sind, so ist es auch P .

Um dies zu zeigen, setzen wir

$$f(Q) = FW(Q),$$

was für alle $Q \varepsilon P$ eindeutig sinnvoll ist²⁸). Dann ist für $Q \varepsilon P$ und $R \varepsilon Q$ nach dem vorigen Resultate

$$f(R) = FW(R) = FE(Q)(R),$$

²⁸) Die Existenz einer Funktion f , so daß für alle normalen P $f(P) = FW(P)$ ist, muß bewiesen werden.

Wenn P normal ist, gibt es ein g , so daß für alle $Q \varepsilon P$

$$g(Q) = \varphi(F(g, Q), Q)$$

ist. Wir setzen $F = F(g, P)$, dann ist die Existenz von F die notwendige und hinreichende Bedingung für die Normalität. Wir werden eine Funktion χ aufsuchen, derart, daß $\chi(P, F) \neq 0$ die notwendige und hinreichende Bedingung dafür ist, daß F in der gewünschten Beziehung zu P steht. (Wir mußten von der Funktion g zur Menge F übergehen, um ein χ konstruieren zu können.)

Von F wird offenbar dies verlangt:

Jedes Element von F hat die Form $\langle u, v \rangle$, $u \varepsilon P$. Zu jedem $u \varepsilon P$ gibt es genau ein v mit $\langle u, v \rangle \varepsilon F$. Wenn $Q \varepsilon P$ ist und F' eine Menge ist, deren Elemente alle $\langle u, v \rangle \varepsilon F$ mit $u \varepsilon Q$ sind, so ist

$$\langle Q, \varphi(F', Q) \rangle \varepsilon F.$$

Um dies in $\chi(P, F) \neq 0$ umzuformen, bemerken wir: Es ist

$$u = \mathfrak{D}(\langle (u), (u, v) \rangle) = \mathfrak{D}(\langle u, v \rangle),$$

und da \mathfrak{D} durch eine Funktion ω ersetzt werden kann (man setzt dies in Evidenz, indem man etwa

$$M_y^* = M_{x \varepsilon y}, \quad \mathfrak{D}(M) = \mathfrak{S}(M)_{M_y^* = M}$$

setzt), so gilt

$$u = \omega(\langle u, v \rangle).$$

Ferner konstruiert man leicht eine Funktion ξ , für die $\xi(x) \neq 0$ damit gleichbedeutend ist, daß x die Form $\langle u, v \rangle = (\langle u \rangle, \langle u, v \rangle)$ hat.

Von F wird also verlangt:

Aus $x \varepsilon F$ folgt $\xi(x) \neq 0$ und $\omega(x) \varepsilon P$. Aus $x \varepsilon F$, $y \varepsilon F$ und $\omega(x) = \omega(y)$ folgt $x = y$. Zu jedem $u \varepsilon P$ gibt es ein $x \varepsilon F$ mit $\omega(x) = u$. Für jedes $Q \varepsilon P$ gilt, wenn

$$F' = F_{\omega(x) \varepsilon Q}$$

gesetzt wird,

$$\langle Q, \varphi(F', Q) \rangle \varepsilon F.$$

(Fortsetzung der Fußnote 28 auf nächster Seite.)

also ist für $Q \varepsilon P$

$$F(f, Q) = F(FE(Q), Q).$$

Und hieraus folgt weiter

$$f(Q) = FW(Q) = \varphi(F(FE(Q), Q), Q) = \varphi(F(f, Q), Q).$$

Damit ist aber gezeigt, daß P normal ist.

5. Nun können wir zeigen, daß alle OZ P normal sind. Im entgegengesetzten Falle wäre etwa P eine nicht normale OZ . Nach 4. können dann nicht alle $Q \varepsilon P$ normal sein, ihre Menge P ²⁹⁾ ist also nicht leer, d. h. $P \in P$, $P \neq 0$. Also hat P ein erstes Element Q_0 . Aus $Q \varepsilon Q_0$ folgt wegen $P < P$ $Q_0 \varepsilon P$, und andererseits $Q \varepsilon OZ$, $Q < Q_0$, so daß Q nicht zu P gehören kann; also ist Q normal, nach 4. muß folglich auch Q_0 normal sein, entgegen $Q_0 \varepsilon P$.

Damit ist unsere Behauptung bewiesen.

Es bleibt noch übrig, hieraus unsere eigentliche Behauptung, wonach eine und nur eine Funktion f den Bedingungen **J.** in **1.** genügt, herzuleiten.

Wir zeigen zuerst: es gibt eine solche. Es sei

$$f(P) = FW(P)$$
³⁰⁾.

Dann gilt für jede OZ P und jedes $Q \varepsilon P$

$$f(Q) = FW(Q) = FE(P)(Q),$$

also ist für jede OZ P

$$F(f, P) = F(FE(P), P).$$

Dies ist aber mit den bei Fraenkel üblichen Methoden unschwer auf die Form $\chi(P, F) \neq 0$ zu bringen.

Nach diesen Vorbereitungen ist unser Ziel wie folgt zu erreichen:

Da P normal ist, existiert genau ein F mit

$$\chi(P, F) \neq 0,$$

es bildet also eine Menge, die nach dem Zusatzprinzip **F.** eine Funktion $\sigma(P)$ von P ist; und F ist dann, als einziges Element, gleich $\ominus(\sigma(P))$, $FW(Q)$ aber gleich $\varphi(F, P) = \varphi(\ominus(\sigma(P)), P)$, was natürlich auf die Form $f(P)$ gebracht werden kann.

²⁹⁾ Die Existenz von P steht fest, sobald eine Funktion g vorliegt, derart, daß $g(Q) \neq 0$ die Normalität von Q bedeutet: dann ist $P = P_g(Q) \neq 0$.

Um sie zu gewinnen, greifen wir auf das χ der Fußnote ²⁸⁾ zurück. Je nachdem Q normal ist oder nicht, ist für kein F oder genau eines

$$\chi(Q, F) \neq 0,$$

sie bilden also jedesmal eine Menge. Nach **F.** ist diese aber eine Funktion $\sigma(Q)$ von Q ; Q ist normal, wenn die genannte Menge, d. h. $\sigma(Q)$, nicht Null ist. Damit haben wir die gewünschte Funktion hergestellt.

³⁰⁾ Da alle P normal sind, liefern die Überlegungen der Fußnote ²⁸⁾ die Existenz dieser Funktion f .

Hieraus folgt aber

$$f(P) = FW(P) = \varphi(F(FE(P), P), P) = \varphi(F(f, P), P).$$

Also besitzt f in der Tat die gewünschten Eigenschaften.

Und nun zeigen wir: keine andere Funktion leistet dies. Es genüge nämlich g den Bedingungen von **J**. Es sei P eine OZ , und g^* diejenige nur in P definierte Funktion, die dort mit g übereinstimmt. Dann ist für jedes $Q \varepsilon P$

$$F(g^*, Q) = F(g, Q),$$

also

$$g^*(Q) = g(Q) = \varphi(F(g, Q), Q) = \varphi(F(g^*, Q), Q),$$

d. h. g^* ist gleich $FE(P)$, und hieraus folgt

$$\begin{aligned} g(P) &= \varphi(F(g, P), P) = \varphi(F(g^*, P), P) \\ &= \varphi(F(FE(P), P), P) = FW(P). \end{aligned}$$

Damit ist alles bewiesen.

Ich möchte zum Schlusse nicht versäumen, meinen Dank an **Frl. E. Noether** auszusprechen, die mich zur Abfassung dieser Arbeit veranlaßt hat.

(Eingegangen am 20. 3. 1927.)

Zusatz zu vorstehendem Aufsatz Herrn v. Neumanns.

Von

Adolf Fraenkel in Marburg.

Dem Abschnitt I der vorstehenden bedeutungsvollen Arbeit Herrn v. Neumanns seien einige Bemerkungen angefügt über den Unterschied, der zwischen *allgemeiner* und *spezieller* Mengenlehre hinsichtlich der axiomatischen Erfordernisse besteht. Damit wird gleichzeitig Mißverständnissen vorgebeugt, die daraus entstehen könnten, daß der vorstehende Aufsatz es naturgemäß mit den Bedürfnissen der Ordnungszahlentheorie — also eines Gebietes der speziellen Mengenlehre — zu tun hat.

1. Für die Erfordernisse der „*allgemeinen Mengenlehre*“¹⁾, die auch die Theorie der geordneten und wohlgeordneten Mengen umfaßt, reicht der von mir geprägte Funktionsbegriff²⁾ und das darauf gestützte Aussonderungsaxiom völlig aus. Dasselbe gilt auch noch für die „*Zermelosche Mengenlehre*“, die durch Hinzunahme des „*Axioms des Unendlichen*“ zur allgemeinen Mengenlehre entsteht; darin sind nicht nur die abzählbaren Mengen und das Kontinuum gesichert, sondern alle durch Potenzmengenbildung daraus entstehenden Mengen, die also ihrer Mächtigkeit nach zum mindesten bis zu jedem Aleph mit endlichem Index aufsteigen. Daher genügt der bisherige Funktionsbegriff auch zur Lösung allgemeiner Fragen wie derjenigen der Unabhängigkeit des Auswahlaxioms³⁾.

2. Wie ich 1921 zeigte³⁾, gibt es einfache (sogar abzählbare) Mengen, zu deren Sicherung die Zermelosche Mengenlehre nicht mehr ausreicht; mit Hilfe derartiger Mengen gelangt man dann leicht zu Mächtigkeiten $\geq \aleph_\omega$. Statt des damals — vor Aufstellung des Funktionsbegriffs —

¹⁾ Zermelo, *Math. Annalen* 65 (1908), S. 266 unten.

²⁾ *Sitzungsber. der Preußischen Akademie d. Wiss., Phys.-Math. Kl.*, 1922, S. 253; vgl. auch Fußnote ¹¹⁾ der vorstehenden Arbeit.

³⁾ *Jahresber. d. D. Math.-Ver.* 30 (1921), S. 97; vgl. auch Fußnote ⁹⁾ der vorstehenden Arbeit.

zur Sicherung solcher Mengen angeregten loseren Ersetzungsaxioms ist es einfacher und zur Sicherung derartiger Mengen dennoch hinreichend, dem Zermeloschen Axiom des Unendlichen die folgende erweiternde Fassung zu geben (unter Benutzung meines Funktionsbegriffs): Wenn eine Menge m und eine Funktion $\varphi(x)$ gegeben sind, so gibt es mindestens eine Menge M , die erstens m und zweitens mit jedem $y \in M$ auch noch $\varphi(y)$ als Element enthält⁴⁾.

3. Da auch dieses Axiom noch nicht alle etwa wünschenswerten speziellen Mengen sichert, so wurde die Hinzufügung weiterer Axiome für einschlägige Untersuchungen vorbehalten. Namentlich zeigt sich, daß zur Ermöglichung der v. Neumannschen Theorie der Ordnungszahlen⁵⁾ weder das vorstehende Axiom noch das Ersetzungsaxiom mit dem bisherigen Funktionsbegriff ausreicht; für letztere Tatsache liefert die vorstehend von v. Neumann gegebene Ableitung des so gefaßten Ersetzungsaxioms aus den übrigen Axiomen den inneren Grund. Die von mir deshalb als notwendig bezeichnete Ausdehnung des Funktionsbegriffs im Ersetzungsaxiom⁶⁾ hat Herr v. Neumann vorstehend (I, Nr. 2) scharfsinnig und in einem Ausmaß durchgeführt, das für sehr weite Zwecke genügt; dabei wird das Ersetzungsaxiom nicht nur unabhängig, sondern auch kräftig genug, um zusammen mit Zermelos Axiom des Unendlichen die Ordnungszahlen v. Neumanns zu sichern. Für die Teilmengenbildung auf Grund des Aussonderungsaxioms kann es beim engeren Funktionsbegriff verbleiben, der nur die elementaren Prozesse Zermelos benutzt.

4. Ob mit dieser dritten Etappe in der Aufstellung zusätzlicher Axiome zwecks Sicherung *spezieller* Mengen das Ziel endgültig erreicht ist, erscheint zweifelhaft. (Man denke etwa an noch offene Fragen wie die nach der Existenz regulärer Anfangszahlen mit Limeszahlindex⁷⁾.) Dagegen bleibt nach wie vor das Axiomensystem der *allgemeinen* Mengenlehre (einschließlich der Theorien der Äquivalenz, der Ordnung und Wohlordnung) durch diese Ausdehnungen unberührt.

⁴⁾ Fraenkel, Einleitung in die Mengenlehre (2. Aufl. 1923), S. 217.

⁵⁾ Vgl. Fußnote ²⁾ des vorstehenden Aufsatzes.

⁶⁾ Math. Zeitschrift 22 (1925), S. 271.

⁷⁾ Vgl. Kuratowski in den Annales de la Société Polonaise de Mathématique 3 (1925), S. 147.

Eine charakteristische Eigenschaft der abgeschlossenen konvexen Punktengen.

Von

Heinrich Tietze in München.

1. Dafür, daß eine in der Ebene gelegene abgeschlossene Punktmenge \mathfrak{M} , welche innere Punkte besitzt, konvex sei, ist es bekanntlich notwendig und hinreichend, daß durch jeden Randpunkt P von \mathfrak{M} wenigstens eine *Stützgerade* an \mathfrak{M} geht. Als *Stützstrecke von der Länge 2ρ* (> 0) durch einen Randpunkt P einer Menge \mathfrak{M} bezeichnen wir¹⁾ eine Strecke CD der Länge 2ρ mit P als Mittelpunkt, wenn im Inneren des Kreises mit dem Mittelpunkt P und dem Radius ρ auf (wenigstens) einer der beiden Seiten seines Durchmessers CD kein Punkt von \mathfrak{M} liegt^{1a)}. Geht durch P eine Stützstrecke an \mathfrak{M} von der Länge 2ρ , so geht durch P an \mathfrak{M} auch eine Stützstrecke der Länge $2\rho'$ für jedes positive $\rho' < \rho$. Das Vorhandensein einer Stützgeraden an \mathfrak{M} durch P ist gleichbedeutend mit dem Vorhandensein von Stützstrecken beliebiger Länge an \mathfrak{M} durch P . Wir wollen nun folgenden Satz beweisen: *Dafür, daß eine zusammenhängende abgeschlossene ebene Punktmenge \mathfrak{M} , welche innere Punkte besitzt, konvex sei* (d. h. mit je zweien ihrer Punkte A, B auch alle Punkte der Strecke AB enthalte) *ist es bereits hinreichend (natürlich auch notwendig), daß es eine positive Zahl ρ gibt, derart, daß durch jeden Randpunkt von \mathfrak{M} eine Stützstrecke an \mathfrak{M} von der Länge 2ρ geht.*

Zunächst mögen die folgenden Beispiele nicht-konvexer Mengen²⁾ zeigen, daß von den gemachten Voraussetzungen keine entbehrt werden

¹⁾ Vgl. die Note „Über konvexe Figuren“, Journal f. Math. 158, p. 168.

^{1a)} (Zusatz bei der Korrektur.) Über die Beziehung des Begriffs „Stützstrecke“ zu dem von Leja und Wilkosz, Ann. de la Soc. polon. de math., 2 (1923), Krakau 1924, p. 222, eingeführten Begriff „Konkavitätspunkt“ vgl. eine demnächst im Journ. f. Math. erscheinende Note.

²⁾ Vgl. auch die in Nr. 4 meiner Note loc. cit.¹⁾ angegebenen Beispiele.

kann: 1. die Vereinigungsmenge \mathfrak{M} zweier getrennt liegender abgeschlossener Kreisscheiben (\mathfrak{M} ist nicht zusammenhängend); 2. die Menge \mathfrak{M} der Punkte im Innern eines Dreiecks einschließlich der Eckpunkte (\mathfrak{M} ist nicht abgeschlossen); 3. die Menge \mathfrak{M} der Punkte auf einem einfachen geschlossenen oder ungeschlossenen Polygonzug (\mathfrak{M} enthält keine inneren Punkte); 4. die Menge \mathfrak{M} der Punkte, die gleichzeitig im Inneren eines Kreises und im Äußern eines dem Kreise eingeschriebenen Dreiecks liegen, einschließlich der Punkte auf dem Kreise und auf den Dreieckseiten (zwar gibt es durch jeden Randpunkt von \mathfrak{M} eine Stützstrecke an \mathfrak{M} , aber ihre Längen sinken auf den Dreieckseiten bei Annäherung an die Ecken unter jedes $\varrho > 0$).

2. Der ausgesprochene Satz möge sogleich in seiner Verallgemeinerung von zwei auf n Dimensionen bewiesen werden. Dabei werde im n -dimensionalen Raum \mathfrak{R}^n ($n \geq 2$) der Punkte (x_1, x_2, \dots, x_n) als *Stützscheibe vom Radius ϱ* (> 0) durch einen Randpunkt $P = (p_1, p_2, \dots, p_n)$ einer Menge \mathfrak{M} eine Punktmenge

$$K(x) \equiv \sum (x_i - p_i)^2 - \varrho^2 < 0$$

$$L(x) \equiv \sum u_i (x_i - p_i) = 0$$

$$(\sum u_i^2 \neq 0)$$

dann bezeichnet, wenn wenigstens eine der beiden Punktmenge $K < 0$, $L > 0$ und $K < 0$, $L < 0$ keinen Punkt von \mathfrak{M} enthält. Der zu beweisende Satz lautet dann:

Dafür, daß eine zusammenhängende, abgeschlossene, innere Punkte enthaltende Menge \mathfrak{M} des \mathfrak{R}^n konvex sei, ist es hinreichend, daß für ein festes $\varrho > 0$ durch jeden Randpunkt von \mathfrak{M} eine Stützscheibe an \mathfrak{M} vom Radius ϱ geht.

3. Der Beweis soll vorerst für den Fall geführt werden, daß \mathfrak{M} beschränkt ist^{2a)} und im übrigen alle gemachten Voraussetzungen erfüllt. Sei dann A ein innerer Punkt von \mathfrak{M} , \mathfrak{M}_0 die Menge aller inneren Punkte von \mathfrak{M} , die sich mit A durch einen ganz im Inneren von \mathfrak{M} verlaufenden Streckenzug verbinden lassen, \mathfrak{M}^* die abgeschlossene Hülle von \mathfrak{M}_0 . Dann ist $\mathfrak{M}^* \subseteq \mathfrak{M}$, jeder Randpunkt von \mathfrak{M}^* ist auch Randpunkt von \mathfrak{M} , daher geht durch ihn eine Stützscheibe an \mathfrak{M} , die natürlich auch Stützscheibe an \mathfrak{M}^* ist. An anderer Stelle³⁾ habe ich bewiesen, daß die abgeschlossene

^{2a)} (Zusatz bei der Korrektur.) Diese Einschränkung ergab sich, weil zur Zeit der Abfassung der vorliegenden Arbeit der im folgenden herangezogene Satz (vgl. ³⁾) nur für beschränkte Mengen ausgesprochen und bewiesen war.

³⁾ loc. cit. ¹⁾, Nr. 3.

Hülle einer zusammenhängenden beschränkten^{3a)} offenen Menge, an die es durch jeden ihrer Randpunkte eine Stützscheibe gibt, konvex ist. Demnach ist \mathfrak{M}^* konvex, und unsere Behauptung ist nachgewiesen, sobald gezeigt ist, daß $\mathfrak{M} = \mathfrak{M}^*$ sein muß.

Zu dem Zwecke setzen wir $\mathfrak{M} = \mathfrak{M}^* + \mathfrak{N}$ und zeigen, daß die Annahme, \mathfrak{N} sei nicht leer, zu einem Widerspruch führt.

In der Tat kann, da \mathfrak{M}^* abgeschlossen ist, die zu \mathfrak{M}^* punktfremde Menge \mathfrak{N} keinen Häufungspunkt von \mathfrak{M}^* enthalten und es muß also \mathfrak{M}^* wenigstens einen Häufungspunkt P von \mathfrak{N} enthalten, weil andernfalls $\mathfrak{M} = \mathfrak{M}^* + \mathfrak{N}$ gegen die Voraussetzung nicht zusammenhängend wäre⁴⁾. Dabei ist P notwendig ein Randpunkt von \mathfrak{M}^* und von \mathfrak{M} . Nun werden wir in Nr. 4 den folgenden Hilfssatz beweisen:

Hilfssatz. Seien \mathfrak{M}^* und \mathfrak{N} zwei punktfremde (nicht leere) Mengen des n -dimensionalen Raumes, welche folgenden Voraussetzungen genügen: \mathfrak{M}^* besitzt innere Punkte, ist abgeschlossen und konvex; jeder Randpunkt von \mathfrak{M}^* ist Randpunkt der Vereinigung $\mathfrak{M} = \mathfrak{M}^* + \mathfrak{N}$; wenigstens ein Randpunkt P von \mathfrak{M}^* ist zugleich Häufungspunkt von \mathfrak{N} ; durch jeden Randpunkt von \mathfrak{M} gibt es eine Stützscheibe an \mathfrak{M} . Unter diesen Voraussetzungen gibt es zu jedem $\varrho > 0$ einen Randpunkt Q von \mathfrak{M} , durch den es keine Stützscheibe an \mathfrak{M} mit einem Radius $\geq \varrho$ gibt (und zwar gibt es solche Punkte Q in beliebiger Nähe von P).

Unsere vorher eingeführten Mengen \mathfrak{M}^* , \mathfrak{N} , \mathfrak{M} erfüllen nun die Voraussetzungen dieses Hilfssatzes. Es müßte also Randpunkte Q von \mathfrak{M} geben mit Stützscheiben an \mathfrak{M} von einem Radius, der nicht $\geq \varrho$ sein kann, im Widerspruch zu der Voraussetzung, daß es durch jeden Randpunkt von \mathfrak{M} eine Stützscheibe vom Radius ϱ gibt.

4. Wir gehen an den Beweis des Hilfssatzes von Nr. 3. Wir können annehmen, daß an \mathfrak{M} durch P eine Stützscheibe \mathfrak{S} vom Radius ϱ existiert, da andernfalls bereits P selbst ein Randpunkt Q der verlangten Art wäre. Sei $X_1, X_2, \dots, X_n, \dots$ eine Folge von Punkten aus \mathfrak{N} mit P als Limes. Die Stützscheibe \mathfrak{S} sei wie oben analytisch durch $K < 0, L = 0$ gegeben. Dabei sei der Bereich $K < 0, L > 0$ frei von Punkten aus \mathfrak{M} . Sei A ein innerer Punkt von \mathfrak{M}^* . Alle Punkte, die auf der Strecke AP zwischen A und P liegen, sind ebenso wie A innere Punkte der konvexen Menge \mathfrak{M}^* . Daher muß A dem Bereich $L < 0$ angehören, da andernfalls die Punkte der Strecke AP

^{3a)} (Zusatz bei der Korrektur.) Die Voraussetzung der Beschränktheit kann übrigens entbehrt werden; vgl. den nachträglichen Zusatz loc. cit.¹⁾, Nr. 5.

⁴⁾ Eine Menge \mathfrak{M} heißt zusammenhängend (Lennes-Hausdorff), wenn bei jeder Zerlegung $\mathfrak{M} = \mathfrak{M}_1 + \mathfrak{M}_2$ in zwei punktfremde nicht-leere Teilmengen $\mathfrak{M}_1, \mathfrak{M}_2$ wenigstens eine dieser Teilmengen einen Häufungspunkt der anderen enthält.

mit einem Abstand $< \varrho$ von P zu $K < 0$, $L \geq 0$ gehören würden und der Bereich $K < 0$, $L > 0$ nicht frei von Punkten aus \mathfrak{M} wäre. Wegen $\lim X_\nu = P$ gibt es ein ν_0 , so daß für $\nu \geq \nu_0$ die Abstände $\overline{PX_\nu} < \varrho$ sind. Für $\nu \geq \nu_0$ liegt also X_ν in $K < 0$ und kann daher nur in $L \leq 0$ liegen, da andernfalls der Bereich $K < 0$, $L > 0$ nicht frei von Punkten aus \mathfrak{M} wäre. Sei P_ν der auf der Geraden AX_ν zwischen A und X_ν gelegene Randpunkt der abgeschlossenen konvexen Menge \mathfrak{M}^* . Wegen $\lim X_\nu = P$ und der leicht beweisbaren Gleichung⁵⁾ $\lim P_\nu = P$ gibt es einen Index $\mu \geq \nu_0$, für den $\overline{P_\mu X_\mu} < \varrho$ ist. Sei (analog wie oben \textcircled{C}) durch $K_\mu < 0$, $L_\mu = 0$ eine Stützscheibe \mathfrak{S}_μ vom Radius ϱ_μ durch P_μ an \mathfrak{M} gegeben, wobei $K_\mu < 0$, $L_\mu > 0$ frei von \mathfrak{M} sei. Da alle Punkte, die zwischen A und P_μ auf der Strecke AP_μ liegen, ebenso wie A innere Punkte von \mathfrak{M}^* sind, muß gemäß einer schon oben angewendeten Überlegung A dem Bereich $L_\mu < 0$ angehören, hingegen X_μ , weil auf der Verlängerung von AP_μ über P_μ hinaus gelegen, dem Bereiche $L_\mu > 0$. Daher kann der Radius ϱ_μ von \mathfrak{S}_μ nicht größer als $\overline{P_\mu X_\mu}$ sein, da sonst der Bereich $K_\mu < 0$, $L_\mu > 0$ den Punkt X_μ aus \mathfrak{M} enthielte. Es muß also $\varrho_\mu \leq \overline{P_\mu X_\mu}$, daher $\varrho_\mu < \varrho$ sein. Der Punkt P_μ ist somit ein Randpunkt Q von \mathfrak{M} von der in unserem Hilfssatz verlangten Art. (Dabei kann man für ein beliebiges $\sigma > 0$ noch μ so wählen, daß $\overline{PQ} = \overline{PP_\mu} < \sigma$.)

Es werde (wegen der nachfolgenden Anwendung) hervorgehoben, daß für die Gültigkeit des Hilfssatzes und für den vorliegenden Beweis keine Voraussetzung über Beschränktheit der Mengen \mathfrak{M} , \mathfrak{M}^* , \mathfrak{N} nötig ist.

5. Sei nunmehr \mathfrak{M} eine *nicht-beschränkte* Menge, die den Voraussetzungen des Satzes von Nr. 2 genügt. Sei \mathfrak{R}_m die Punktmenge $\sum x_i^2 \leq m^2$ ($m = 1, 2, 3, \dots$) und $\mathfrak{M}\mathfrak{R}_m$ der Durchschnitt von \mathfrak{M} mit \mathfrak{R}_m , der für $m \geq m_0$ nicht leer sei und innere Punkte besitze. Sei A ein innerer Punkt von $\mathfrak{M}\mathfrak{R}_{m_0}$. Sei \mathfrak{M}_m (für $m \geq m_0$) die Menge derjenigen inneren Punkte von $\mathfrak{M}\mathfrak{R}_m$, die sich mit A durch einen ganz im Inneren von $\mathfrak{M}\mathfrak{R}_m$ verlaufenden Streckenzug verbinden lassen, ferner \mathfrak{M}_m^* die abgeschlossene Hülle von \mathfrak{M}_m und \mathfrak{M}^* die Vereinigung aller Mengen \mathfrak{M}_m^* . (Es ist $\mathfrak{M}_m^* \subseteq \mathfrak{M}_{m+1}^* \subseteq \mathfrak{M}^*$.) Jeder Randpunkt von \mathfrak{M}_m^* ist, sofern er nicht

⁵⁾ Sei $\varepsilon > 0$ (wobei $\varepsilon < \varrho$ und $\varepsilon < \overline{AP}$ sei) und O der Punkt im Abstand $\frac{\varepsilon}{2}$ von P auf der Strecke AP . Da O innerer Punkt von \mathfrak{M}^* ist, gibt es ein positives $\delta < \frac{\varepsilon}{2}$, so daß die Menge \mathfrak{U} aller Punkte mit einem Abstand $\leq \delta$ von O ganz dem Inneren von \mathfrak{M}^* angehört. Wegen $\lim X_\nu = P$ müssen für $\nu \geq \nu_1$ die Abstände $\overline{OX_\nu} < \varepsilon$ sein und die Geraden AX_ν den Bereich \mathfrak{U} in einer Strecke UV treffen. Da V , dem Inneren von \mathfrak{M}^* angehört, muß P_ν zwischen V und X_ν liegen und daher ebenso wie V , und X_ν dem Inneren der Kugel vom Radius ε um P angehören: $\overline{PP_\nu} < \varepsilon$.

Randpunkt von \mathfrak{M} ist, ein Randpunkt von \mathfrak{R}_m und es geht also durch ihn entweder eine Stützscheibe vom Radius ϱ an \mathfrak{M} oder andernfalls eine Stützebene an \mathfrak{R}_m , auf alle Fälle also eine Stützscheibe vom Radius ϱ an \mathfrak{M} , \mathfrak{R}_m und an \mathfrak{M}_m^* . Außerdem ist \mathfrak{M}_m^* zusammenhängend und beschränkt, erfüllt somit alle Voraussetzungen, unter denen in Nr. 3 der Satz von Nr. 2 bereits bewiesen ist. Somit ist jede Menge \mathfrak{M}_m^* ($m \geq m_0$) konvex. Sind nun B, C zwei Punkte von \mathfrak{M}^* , so muß es ein m geben, so daß beide Punkte B, C und dann auch ihre Verbindungsstrecke BC in \mathfrak{M}_m^* und damit in \mathfrak{M}^* liegen; es ist also auch \mathfrak{M}^* konvex. Ferner ist \mathfrak{M}^* abgeschlossen; ist nämlich D ein Randpunkt von \mathfrak{M}^* (wobei D in \mathfrak{R}_{m_1} liege und $m_1 \geq m_0$ sei), so gehören, da A ein innerer Punkt von \mathfrak{M}^* ist, alle zwischen A und D gelegenen Punkte der Strecke AD zu \mathfrak{M}^* und somit zu \mathfrak{M} ; es gehört also D selbst zu \mathfrak{M} und daher zu \mathfrak{M}_m^* , und zu \mathfrak{M}^* . Da \mathfrak{M}^* konvex und $\mathfrak{M} \supseteq \mathfrak{M}^*$ ist, bleibt nur noch zu zeigen, daß $\mathfrak{M} = \mathfrak{M}^*$ sein muß.

Zu dem Zwecke setzen wir, analog wie in Nr. 3, $\mathfrak{M} = \mathfrak{M}^* + \mathfrak{N}$. Wäre nun \mathfrak{N} nicht leer, so müßte es, weil \mathfrak{M} zusammenhängend ist, wie in Nr. 3 in \mathfrak{M}^* einen Häufungspunkt P von \mathfrak{N} geben, woraus sich wie dort auf Grund desselben Hilfssatzes ein Widerspruch mit der Voraussetzung ergäbe, daß durch jeden Randpunkt von \mathfrak{M} eine Stützscheibe an \mathfrak{M} vom Radius ϱ geht.

Damit ist der Satz von Nr. 2 allgemein bewiesen.

München, den 27. Juni 1927.

Topological equivalence of complexes.

Von

M. H. A. Newman in Cambridge (England).

The "internal transformation" (simple subdivision of a cell), which is the basis of combinatory analysis situs as developed in the works of Dehn, Heegaard, Kneser¹), and others, does not lend itself very readily to the direct investigation of topological invariants: owing to the great variety of ways in which a single n -cell may be subdivided it is necessary to interpolate discussions of the properties of spheres and their intersections not always germane to the real subject matter. It is the object of this paper to shew that a certain general transformation, of which "internal transformation" is a special case, can always be carried out by a series of "moves" of a much more specialised kind, whose effects are easily followed without the help of general theorems about spheres.

The moves employed are those that I defined in an earlier paper²), but the standpoint there adopted has been modified in one respect. The " n -arrays", discussed in §§ 1 and 2 below, are no longer regarded as being themselves the subject matter of combinatory topology, but are thought of rather as instruments for proving theorems about the n -complexes introduced in § 3. This does not necessitate any change in the definitions or results of F I and F II, save that the symbol " \rightarrow " is no longer read as "is topologically equivalent to", which is given a different sense.

§ 1 is a summary of the definitions and leading theorems of F I and F II, § 2 contains a new proof and generalisation of F II Theorem 10, § 3 gives its application to n -complexes.

¹) Dehn and Heegaard, *Encykl. der Math. Wiss.* III, AB 3 *Analysis Situs*; H. Kneser, *Proc. Amsterdam* 27 (1924), p. 601; E. Bilz, *Math. Zeitschr.* 18 (1923), S. 1.

²) The Foundations of Combinatory Analysis Situs, I and II, *Proc. Amsterdam* 29 (1926), p. 611 and 627, referred to as F I and F II. See also "Additions and Corrections" to these papers, *Proc. Amsterdam* 30 (1927), p. 670, cited as F . Add.

§ 1.

Arrays³⁾.

If n is a positive integer or zero, an n -array is formed from a finite or enumerable set of objects by choosing as the *units* of the array certain of the groups of $n + 1$ objects contained in the set⁴⁾. The choice is unrestricted save that every object must belong to at least one unit. The objects are called the *vertices* of the n -array, and if $0 \leq k \leq n$ any $k + 1$ vertices all belonging to the same unit form a k -component of the array. An $(n - 1)$ -component is called a *face*; units with a face in common are *adjacent*. An n -simplex is an n -array with only one unit. The *logical sum*, $\Gamma \dot{+} \Delta \dot{+} \dots$, of a number of n -arrays, Γ, Δ, \dots , is the n -array whose units are all the units of all the arrays⁵⁾. — The *sum mod 2* of Γ, Δ, \dots , denoted by $\Gamma + \Delta + \dots$, is the n -array whose units are the n -simplexes contained in⁶⁾ an odd number of the arrays. When no two of Γ, Δ, \dots have a common unit the logical sum and the sum mod 2 are the same and may be called the “sum”, simply, and denoted by $\Gamma + \Delta + \dots$. If the n -array Γ contains the n -array Δ , $\Gamma - \Delta$ is the sum of the n -simplexes contained in Γ but not in Δ .

If S and T are simplexes with no common vertex ST is the simplex containing all the vertices of both. If Γ and Δ are arrays with no common vertex $\Gamma\Delta$ is the sum of all products ST , where S is a unit of Γ and T of Δ .

An n -array is *regular* if each face belongs to at most two units and each vertex to a finite number. It is *connected* if every two units are the extreme members of a sequence of units such that adjacent members of the sequence are adjacent units of the array.

The sum of the faces of an n -array, Γ , that belong each to only one unit is the *boundary* of Γ , denoted by $\bar{\Gamma}$, (or, when that is inconvenient, by $B(\Gamma)$). A component not contained in the boundary is *internal*. The sum mod 2 of the boundaries of the units of Γ is the *margin* of Γ , denoted by $\hat{\Gamma}$.

³⁾ The definitions in § 1 are taken, with certain slight modifications, from $F I$ and $F II$, where proofs of the theorems will be found. It should be noticed that Theorem 10 of $F II$ is not assumed. The present paper replaces § 9 of $F II$.

⁴⁾ P. Alexandroff pointed out almost simultaneously (*Math. Annalen* 96 (1926), p. 489) that an n -simplex may be regarded as being simply its $n + 1$ vertices. See also the same author's paper, *Math. Annalen* 94 (1925), p. 296.

⁵⁾ The corresponding notation for sets of points was introduced by Carathéodory, *Reelle Funktionen*, p. 28.

⁶⁾ The array Γ^1 contains the array Γ^2 if every unit of Γ^2 is a unit or component of Γ^1 .

It is easily proved that if Γ and Δ have no common vertex and $d(\Gamma\Delta) > 0$), $\overline{\Gamma\Delta}$ is $\Gamma \cdot \overline{\Delta} + \Delta \cdot \overline{\Gamma}$ (FI 1) and $\widehat{\Gamma\Delta}$ is $\Gamma \cdot \widehat{\Delta} + \Delta \cdot \widehat{\Gamma}$ (FI 1a). Here it is agreed that if Γ is a vertex $\overline{\Gamma}$ (or $\widehat{\Gamma}$) shall be omitted from terms containing it, and if Γ is unbounded (or has no margin) terms containing $\overline{\Gamma}$ (or $\widehat{\Gamma}$) are to be omitted. E. g., if α is a vertex $\overline{\alpha\Gamma}$ is $\Gamma + \alpha \cdot \overline{\Gamma}$).

(In the rest of this paper all arrays are supposed to contain only a finite number of vertices.)

S_n is said to have *regular contact* with Γ_n , not containing it, if it is the product of two components U and V , (S_n is UV), such that (I) U belongs to $\overline{\Gamma}_n$, (II) U is interior to $\Gamma_n + UV$, (III) V does not belong to Γ_n . (U and V must each contain at least one vertex.)

If Γ_n is regular the common faces of Γ_n and UV form the array $U \cdot \overline{V}$, lying in $\overline{\Gamma}_n$ (FI 4 and 5).

If Γ_n is bounded it is a *move of type 1* to add to (Γ_n) a simplex, S_n , having regular contact with it, and a *move of type 2* to remove T_n having regular contact with $\Gamma_n - T_n$.

If Γ_n contains $S_k \cdot \overline{T}_{n-k}$, but $\Gamma_n - S_k \cdot \overline{T}_{n-k}$ contains neither S_k nor T_{n-k} , it is a *move of type 3* to substitute $\overline{S}_k \cdot T_{n-k}$ for $S_k \cdot \overline{T}_{n-k}$ in Γ_n . (The case $k = n$ is included).

If Δ can be obtained from Γ by a succession of moves of any of the three types we write $\Gamma \rightarrow \Delta$ ⁹⁾; if only moves of types p and q are required we write $\Gamma \xrightarrow{pq} \Delta$.

If $\Gamma_n \rightarrow S_n$, Γ_n is an n -element (E_n). The boundary of an $(n+1)$ -element is an n -sphere (Σ_n). The necessary and sufficient condition that Γ_n should be an n -sphere is that $\Gamma_n \xrightarrow{3} \overline{S}_{n+1}$ (FI 9 and 14).

If S_k and Σ_{n-k-1} have no common vertex $S_k \Sigma_{n-k-1}$ is a *complete* $(n:k)$ -cluster; if S_k and E_{n-k-1} have no common vertex $S_k E_{n-k-1}$ is an *incomplete* $(n:k)$ -cluster. S_k is the *core*, Σ_{n-k-1} and E_{n-k-1} are the *shells*. An $(n:0)$ -cluster is an n -star.

An n -manifold (M_n) is a connected n -array such that the sum of the units at each vertex is an n -star. If all the stars are complete the manifold is unbounded, otherwise it is bounded.

⁹⁾ $d(\Gamma)$ means "the dimension-number of Γ ".

⁹⁾ The letters Γ and Δ will be used for arrays of uncertain character; S, T, U, V , for simplexes; small Greek letters for vertices. If a lower index is present, it denotes the dimension number; upper indices are merely distinguishing marks.

⁹⁾ This differs from the definition of FI in the case of unbounded arrays, but the two definitions are equivalent in the case of manifolds in view of FI 21 and FII Theorem 5. " $\Gamma \rightarrow \Delta$ " is no longer to be read " Γ is topologically equivalent to Δ " but, e. g., as " Γ leads to Δ ". Cf. F. Add., under "Topological equivalence".

The units of M_n containing the component S_k form an $(n:k)$ -cluster, complete or incomplete as S_k is internal to M_n or in \bar{M}_n . (F I 19 and 31.) This $(n:k)$ -cluster is called the S_k -cluster in M_n .

The numerous small theorems — $(n:k)$ -clusters are n -elements, n -spheres a ren-manifolds, etc. — required to set in motion the theory based on these definitions will be found in F I. The following results from F II may here be mentioned:

F II Theorems 1, 3a, 3b: If $M^1 \rightarrow M^2$ then $M^1 \xrightarrow{12} M^2$, $M^1 \xrightarrow{23} M^2$, $M^1 \xrightarrow{13} M^2$.

F II Theorem 2. If $M^1 \xrightarrow{3} M^2$ and $\Gamma + M^1$ is a manifold, then $\Gamma + M^1 \xrightarrow{3} \Gamma + M^2$, provided Γ contains no internal component of M^2 .

F II Theorem 4. If \bar{E}^1 is \bar{E}^2 , then $E^1 \xrightarrow{3} E^2$.

F II Theorem 5. If the unbounded manifolds M^1 and M^2 contain units S and T , respectively, such that $M^1 - S \rightarrow M^2 - T$, then $M^1 \xrightarrow{3} M^2$ ¹⁰.

F II Theorem 6. If \bar{E}_n^1 is \bar{E}_n^2 and M_n contains E_n^1 , then $M_n \rightarrow (M_n - E_n^1) + E_n^2$, provided $M_n - E_n^1$ contains no internal component of E_n^2 .

(If M_n is unbounded $M_n \xrightarrow{3} (M_n - E_n^1) + E_n^2$.)

Corollary 1. If \bar{E}_n^1 is \bar{E}_n^2 , but E_n^1 and E_n^2 have no common internal component, $E_n^1 + E_n^2$ is an n -sphere.

Corollary 2. $\Sigma_n - E_n$ is an n -element.

F II Theorem 7. Suppose M_n contains E_n and \bar{E}_n contains E_{n-1} , and let E_n^* be a second n -element whose boundary contains E_{n-1} . Then if all components of $M - E_n$ belonging to E_n or E_n^* belong to E_{n-1} , $M_n \rightarrow (M_n - E_n) + E_n^*$.

F II Theorem 8a. If the common part of M_n and E_n is an $(n-1)$ -element in the boundary of each, $M_n \rightarrow M_n + E_n$.

F II Theorem 8b. If M_n contains E_n , and the part of E_n in \bar{M}_n is an $(n-1)$ -element, then $M_n \rightarrow M_n - E_n$.

F II Lemma 7a. If E and Π have no common vertex, Π being a sphere or an element, then $E\Pi$ is an element.

F II Lemma 7b. If Σ and Σ' have no common vertex, $\Sigma\Sigma'$ is a sphere.

¹⁰ Theorems 1, 3, 5, give all the general relations between the three moves. The only other plausible suggestion — “If $M^1 \rightarrow M^2$ and \bar{M}^1 is \bar{M}^2 then $M^1 \xrightarrow{3} M^2$ ” — is false. See F. Add. p. 671.

§ 2.

Assemblies of pieces.

An n -dimensional assembly of pieces is a collection of elements, (the "pieces" of the assembly), of all dimension numbers from 0 to n such that (denoting by G_i an i -dimensional piece):

P(I) if G_i contains a unit or internal component of G_j , \bar{G}_i contains G_j unless G_i and G_j are identical;

P(II) if $i > 0$ every unit of \bar{G}_i belongs to an $(i - 1)$ -dimensional piece;

P(III) if $i < n$ G_i is contained in an $(i + 1)$ -dimensional piece.

Two assemblies Γ and Γ' have the same structure if a (1, 1) correspondence \mathfrak{A} can be set up between their pieces so that (a) $d(\mathfrak{A}G_i) = i$; (b) G_i (of Γ) contains G_j (of Γ') if, and only if, $\mathfrak{A}G_i$ contains $\mathfrak{A}G_j$. \mathfrak{A} is the structural relation.

If Γ denotes an n -assembly, Γ' denotes the sum of the n -dimensional pieces of Γ ¹¹).

Lemma 1. If all j -pieces of the assembly Γ_n for which $j > k$ are complete stars, and G_k is any k -piece, the units of Γ_n that contain a unit of G_k form an array $G_k A_{n-k-1}$.

It must be shewn that if U_k and V_k are units of G_k , and $U_k X_{n-k-1}$ is a unit of Γ_n , then $V_k X_{n-k-1}$ is a unit of Γ_n . Let G_n be the n -piece containing $U_k X_{n-k-1}$. The centre, a , of G_n does not, by P(I), belong to G_k , and therefore $U_k X_{n-k-1}$ is $\alpha U_k Y_{n-k-2}$. The pieces of Γ_n contained in \bar{G}_n constitute an $(n - 1)$ -assembly and so, using an inductive hypothesis, \bar{G}_n , which contains $U_k Y_{n-k-2}$, contains $V_k Y_{n-k-2}$. Hence G_n contains $V_k X_{n-k-1}$.

The array $G_k A_{n-k-1}$ is called the G_k -cluster in Γ_n , A_{n-k-1} is its shell.

If Γ_n is a manifold A_{n-k-1} is a sphere or an element. (For if U_k is a unit of G_k , $U_k A_{n-k-1}$ is the U_k -cluster in Γ_n .)

Lemma 2. Let G_k^0 be a k -piece of the n -assembly M , (M being a manifold), and let H_k^0 be a k -element, whose boundary is identical with \bar{G}_k^0 but of which no unit or component not contained in G_k^0 belongs to M . Then there is an assembly Γ with these properties:

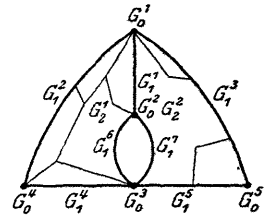


Fig. 1.
2-dimensional assembly with two 2-pieces.

¹¹) The conventions governing the use of letters may be extended to assemblies; e. g., the use of M to denote an n -assembly implies that the sum of its n -pieces is a manifold.

- (a) $M \rightarrow \Gamma$;
 (b) M and Γ have the same structure (with structural relation \mathfrak{A} , say);
 (c) if $j > k$ and G_j is a j -piece of M , $\mathfrak{A}G_j$ is a complete star;
 (d) if $j \leq k$ $\mathfrak{A}G_j$ is G_j , save that $\mathfrak{A}G_k^0$ is H_k^{012} .

The lemma being true when $k = n$, in virtue of *FII*, Theorems 4 and 2¹³), suppose it true when k is replaced by number greater than itself. By applying this hypothesis to all the $(k+1)$ -dimensional pieces of M in turn, (taking H_{k+1}^0 itself to be a complete star), M may be transformed into an assembly M' with the same structure, in which all pieces with $k+1$ or more dimensions are complete stars, while other pieces are unchanged, and $M \rightarrow M'$. We may then suppose, without loss of generality, that in M itself all pieces of more than k dimensions are complete stars.

The boundary of any piece, G_j , containing G_k^0 has, (by Lemma 1, and *FII* Lemma 7a and Corollary 2 to Theorem 6), the form $E_{j-1} + G_k^0 \Sigma_{j-k-2}$. If G_h (with centre α) is contained \bar{G}_j and contains G_k^0 , then $\alpha \Sigma_{h-k-2}$ is contained in Σ_{j-k-2} and αE_{h-1} in E_{j-1} . (Remark A.) For $\alpha \Sigma_{h-k-2} G_k^0$ must be contained in $\Sigma_{j-k-2} G_k^0$, the set of all units of \bar{G}_j containing a unit of G_k^0 ; and if αE_{h-1} did not belong entirely to E_{j-1} it would contain an internal component of $G_k^0 \Sigma_{j-k-2}$, i. e. an internal component of G_k^0 .

Consider the set of j -arrays, $\mathfrak{A}G_j$, defined as follows (G_j being a typical cell of M):

- $\mathfrak{A}G_k^0$ is H_k^0 ;
 if G_j is $\alpha^j(E_{j-1} + G_k^0 \Sigma_{j-k-2})$, $\mathfrak{A}G_j$ is $\alpha^j(E_{j-1} + H_k^0 \Sigma_{j-k-2})$;
 if G_j does not contain G_k^0 , $\mathfrak{A}G_j$ is G_j .

This set of arrays may be called $\mathfrak{A}M$. The following properties are evident: All vertices of $\mathfrak{A}G_i$ not in H_k^0 belong to G_j (Remark B). The necessary and sufficient condition that $\mathfrak{A}\bar{G}_h$ should contain an array $H_k^0 \Gamma_q$ is that \bar{G}_h should contain $G_k^0 \Gamma_q$. (Remark C.) If a unit U_j of G_j contains no unit of G_k^0 , U_j belongs to $\mathfrak{A}G_j$; and conversely, a unit of $\mathfrak{A}G_j$ containing no unit of H_k^0 belongs to G_j . (Remark D.)

It can now be shewn that $\mathfrak{A}M$ is an n -assembly having the properties required of Γ .

- (a) $\mathfrak{A}G_j$ is a j -element.

¹³) The invariance of the property of being a manifold was proved by Weyl (Revista di Matem. Hisp.-Amer., 1923) by a process somewhat similar to that here adopted. Weyl's fundamental definitions are, however, of a radically different character.

¹⁴) Quoted on p. 402.

Since α^j does not (by definition) belong to E_{j-1} or Σ_{j-k-2} , nor (by hypothesis) to H_k^0 , it is sufficient to shew that $E_{j-1} + H_k^0 \Sigma_{j-k-2}$ is a $(j-1)$ -sphere, (FI 12). $B(H_k^0 \Sigma_{j-k-2})$ is $\bar{H}_k^0 \Sigma_{j-k-2}$, which is $\bar{G}_k^0 \Sigma_{j-k-2}$, which is \bar{E}_{j-1} ; units and internal components of $H_k^0 \Sigma_{j-k-2}$ contain units or internal components of H_k^0 , and so do not belong to E_{j-1} . Hence (FII, Theorem 6, Corollary 1), $E_{j-1} + H_k^0 \Sigma_{j-k-2}$ is a $(j-1)$ -sphere.

(b) Every unit of $\overline{\mathfrak{A}G_j}$ belongs to an $\mathfrak{A}G_{j-1}$.

A unit, U_{j-1} , of $\mathfrak{A}G_j$ not containing a unit of H_k^0 belongs to \bar{G}_j , and therefore to some piece G_{j-1} , i. e. (by Remark D) to $\mathfrak{A}G_{j-1}$. A unit V_{j-1} that contains a unit of H_k^0 belongs to an array $H_k^0 X_{j-k-2}$ contained in $\overline{\mathfrak{A}G_j}$; $G_k^0 X_{j-k-2}$ is contained in \bar{G}_j , and therefore in a piece G_{j-1} ; and $\mathfrak{A}G_{j-1}$ contains V_{j-1} .

(c) If \bar{G}_j contains G_h , $\overline{\mathfrak{A}G_j}$ contains $\mathfrak{A}G_h$.

This follows at once from Remark A unless G_h contains G_k^0 . In that case (with the notation of Remark A) αE_{h-1} is contained in E_{j-1} , $H_k^0 \alpha \Sigma_{h-k-2}$ in $H_k^0 \Sigma_{j-k-2}$, and therefore $\mathfrak{A}G_h$ in $\overline{\mathfrak{A}G_j}$.

(d) If $\mathfrak{A}G_h$ contains a unit or internal component of $\mathfrak{A}G_j$, $\overline{\mathfrak{A}G_h}$ contains $\mathfrak{A}G_j$.

By (c) it is sufficient to shew that \bar{G}_h contains \bar{G}_j . If $j \leq k$ the result follows from Remark D (cf. (b)). If $j > k$ $\mathfrak{A}G_j$ is a complete star, whose centre, β , is contained in $\mathfrak{A}G_h$. Since β is also the centre of G_j it is not interior to H_k^0 and therefore (Remark A) belongs to G_h . Hence, by P(II), \bar{G}_h contains G_j .

From these results it follows that $\mathfrak{A}M$ is an n -assembly. The complete symmetry of the relations between M and $\mathfrak{A}M$ being thus established it may be inferred that

(c') If $\overline{\mathfrak{A}G_j}$ contains $\mathfrak{A}G_h$, then \bar{G}_j contains G_h , shewing that M and $\mathfrak{A}M$ have the same structure. Finally,

(e) $M \rightarrow \mathfrak{A}M$.

If $G_k^0 \Pi_{n-k-1}$, (say E_n^1), is the sum of all units of M containing a unit of G_k^0 , the corresponding sum in $\mathfrak{A}M$ is, by Remark C, $H_k^0 \Pi_{n-k-1}$, (say E_n^2). The arrays $M - E_n^1$ and $\mathfrak{A}M - E_n^2$ are identical, and so are the $(n-1)$ -elements $\bar{G}_k^0 \cdot \Pi_{n-k-1}$ and $\bar{H}_k^0 \cdot \Pi_{n-k-1}$, which are respectively the part of \bar{E}_n^1 interior to M and the part of \bar{E}_n^2 interior to $\mathfrak{A}M$. Hence (by F II Theorem 6, if Π_{n-k-1} is a sphere, by Theorem 7 if it is an element) $M \rightarrow (M - E_n^1) + E_n^2$, which is $\mathfrak{A}M$.

Theorem 1. *If the assemblies M and Γ have the same structure and M is a manifold, then Γ is a manifold and $M \rightarrow \Gamma$.*

The 1-dimensional pieces in M can, by Lemma 2, be changed one by one into their assigned correlates in Γ , while all other pieces are

changed into complete stars. The right boundaries having thus been provided for the 2-pieces these may next be changed, without disturbing the 1-pieces; and so on. To ensure that this process can be continued until Γ is finally reached it is only necessary to shew that the final condition imposed on H_k^0 in Lemma 2 is fulfilled at every stage.

Suppose then that M has been successfully changed into $M^{(k)}$, in which some k -dimensional pieces, and all j -pieces for which $j < k$, are "regulated", (i. e. congruent¹⁴) to their assigned Γ -correlates), while the remaining "unregulated" pieces are complete stars. Let the operation to be justified be the replacement of the k -piece, G_k^0 , of $M^{(k)}$ by a k -element, with boundary \bar{G}_k^0 , congruent to G_k^* of Γ , (\bar{G}_k^0 and \bar{G}_k^* being congruent by hypothesis). It can certainly be arranged that no internal *vertex* of the new k -piece belongs to $M^{(k)}$, for the choice of the internal vertices is unrestrained. If, on the other hand, there is an internal component, U^* , of G_k^* , with vertices lying in \bar{G}_k^* , such that the simplex with the corresponding vertices in \bar{G}_k^0 belongs to $M^{(k)}$, then U must be a unit or internal component of some piece, G_j , of $M^{(k)}$ ¹⁵. If G_j is not G_k^0 itself it is one of the pieces already regulated, for the unregulated pieces are complete stars whose centres do not belong to \bar{G}_k^0 . Hence G_j is congruent to its Γ -correlate, G_j^* , and $j > k$. It follows that U^* , which is a unit or internal component of G_k^* , belongs to G_j^* ; and this is incompatible with $P(I)$.

§ 3.

Complexes.

(In this section the objects so far called "elements", "spheres", "manifolds", "pieces", will be called " Δ elements", " Δ spheres", " Δ manifolds", and " Δ pieces", to distinguish them from the " \circ elements", " \circ spheres" etc. now to be introduced¹⁶). "Array", "unit", "component", having no new analogues, are still used, without prefix, in the same sense as before.)

A *cell* is simply an object associated with a positive integer, its *dimension number*. A collection of cells, of dimension numbers from 0 to n , becomes an *n-set of cells*, if certain pairs of cells with dimension numbers differing by 1 are associated together. The association of an i -cell,

¹⁴) Two arrays are *congruent* if they are completely similar.

¹⁵) If Γ is any assembly every component of Γ is a unit or internal component of some piece.

¹⁶) The connection between the Δ - and \circ -entities having once been cleared up it should be possible to drop the prefixes in most contexts.

a_i , with an $(i+1)$ -cell, a_{i+1} , is expressed in the statement " a_i bounds a_{i+1} ", and more generally, if a_i bounds a_j and a_j bounds a_k , a_i is said to bound a_k .

A *proper* n -set is one in which if $i < n$ every i -cell bounds at least one $(i+1)$ -cell, and if $j > 0$ every j -cell is bounded by at least one $(j-1)$ -cell. From any proper n -set, $\circ\Gamma$, an n -array $\Delta\Gamma$, called the *skeleton* of $\circ\Gamma$, is formed by taking as vertices the cells of $\circ\Gamma$ and as units all sets of $n+1$ cells, $a^{(0)}, a^{(1)}, \dots, a^{(n)}$, such that $a^{(i)}$ bounds $a^{(i+1)}$ ($i = 0, \dots, n$). Evidently $a^{(i)}$ must be an i -cell¹⁷.

If a_k is a cell of a proper n -set, the proper k -set formed by a_k and all the cells bounding it is denoted by $\circ\{a_k\}$. The array $\Delta\{a_k\}$ is seen to be the sum of all k -components of $\Delta\Gamma$ whose vertices form a sequence descending from a_k . Clearly if a_n^1, a_n^2, \dots , are the n -cells of $\circ\Gamma$, $\Delta\Gamma$ is $\sum_{(i)} \Delta\{a_n^i\}$.

The *boundary* $(n-1)$ -cells of an n -set are those that bound only one n -cell; the *boundary* k -cells, ($k < n-1$), are those that bound a boundary $(n-1)$ -cell. Other cells, including all n -cells, are *internal*. The *boundary* of $\circ\Gamma$ (denoted by $\circ\bar{\Gamma}$ or $B(\circ\Gamma)$) is the set of all boundary cells, (if any). If $\circ\Gamma$ is a proper n -set $\circ\bar{\Gamma}$ is evidently a proper $(n-1)$ -set. If there is no boundary $\circ\Gamma$ is *unbounded*¹⁸.

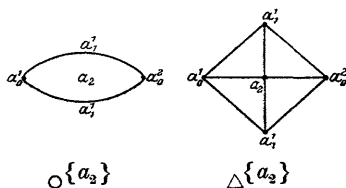


Fig. 2.

Theorem 2. *If a_k is a boundary cell of the proper n -set $\circ\Gamma$, $\Delta\{a_k\}$ belongs to $B(\Delta\Gamma)$.*

By hypothesis there exists in $\circ\Gamma$ a sequence of cells ascending from a_k to a cell a_{n-1} bounding only one n -cell, a_n . If then $a_k a_{k-1} \dots a_0$ is a unit of $\Delta\{a_k\}$, $a_{n-1} a_{n-2} \dots a_k \dots a_0$ is a face of $\Delta\Gamma$ containing it and belonging to only one unit of $\Delta\Gamma$, viz. $a_n a_{n-1} \dots a_k \dots a_0$.

Corollary 1. *The skeleton of $\circ\bar{\Gamma}$ is contained in $B(\Delta\Gamma)$.*

(Set $k = n - 1$ in Theorem 2).

Corollary 2. *If $\Delta\Gamma$ is unbounded $\circ\Gamma$ is unbounded.*

(From Corollary 1).

An n osphere is a proper n -set whose skeleton is an $n\Delta$ sphere.

From Corollary 2 to Theorem 2 it follows that n ospheres are unbounded.

¹⁷) A sequence of cells a_i, a_{i+1}, \dots, a_j , in which each cell bounds its successor is called a sequence ascending from a_i or descending from a_j .

¹⁸) The margin of an n -set of cells may be defined analogously to the margin of an array (p. 400) and has properties similar to those of the boundary.

An n -complex is an n -set of cells (not necessarily "proper") such that the boundary of every k -cell¹⁹) is a $(k-1)$ -sphere, ($k=1, 2, \dots, n$).

Theorem 3. If $\circ\Gamma$ is a proper n -complex the skeleton of $\circ\bar{\Gamma}$ is $B(\Delta\Gamma)$.

In view of Theorem 2, Corollary 1, it is sufficient to shew that every unit of $B(\Delta\Gamma)$ belongs to the skeleton of $\circ\bar{\Gamma}$.

In the first place the vertices of any boundary face of $\Delta\Gamma$ form a sequence descending from a boundary $(n-1)$ -cell. For if $a_n \dots a_{j+1} a_{j-1} \dots a_0$ were a bounding face there would be only one j -cell, a_j , such that a_{j+1}, a_j, a_{j-1} is a descending sequence. But \bar{a}_{j+1} is unbounded and therefore there are at least two j -cells in \bar{a}_{j+1} bounded by a_{j-1} . Again if $a_{n-1} a_{n-2} \dots a_0$ is a bounding face of $\Delta\Gamma$, a_{n-1} is a boundary cell, for if it bounded both a_n and a'_n , $a_n a_{n-1} \dots a_0$ and $a'_n a_{n-1} \dots a_0$ would be two units of $\Delta\Gamma$ containing the given face.

From this Theorem 3 follows at once, for if a_{n-1} is a bounding $(n-1)$ -cell all members of any sequence descending from it are by definition boundary cells, i. e. the members of any such sequence are the vertices of a unit of the skeleton of $\circ\Gamma$ ²⁰).

Both $B(\Delta\Gamma)$ and the skeleton of $\circ\bar{\Gamma}$ may, then, be denoted by $\Delta\bar{\Gamma}$.

Corollary. If a_k is a boundary vertex of $\Delta\Gamma$ it is a boundary cell of $\circ\Gamma$.

Theorem 4. Every n -sphere is an n -complex.

Let the theorem be assumed true of \circ spheres of less than n dimensions.

If a_n is an n -cell of the n -sphere $\circ\Sigma$, the array $\Delta\{a_n\}$, being the sum of all units of $\Delta\Sigma$ containing the vertex a_n , is a complete star with centre a_n (FI 18). Hence the skeleton of \bar{a}_n , whose units are those of $\Delta\{a_n\}$ with the vertex a_n left out, is an $(n-1)$ -sphere: the boundaries of n -cells of $\circ\Sigma$ are $(n-1)$ -spheres. It now follows from the inductive hypothesis that $\circ\Sigma$ is an n -complex.

An n -element is a proper n -complex whose skeleton is an n -element²¹). By Theorem 3 its boundary is an $(n-1)$ -sphere.

Theorem 5. If $\circ\Gamma$ is a proper n -complex the arrays $\Delta\{a_k\}$ formed from all the cells of $\circ\Gamma$ are the pieces of an n -assembly.

¹⁹) "The boundary of $\circ\{a_k\}$ " may be abbreviated to "the boundary of a_k " and denoted by \bar{a}_k .

²⁰) It will be noticed that the only property of \circ spheres used in this proof is their unboundedness. For a systematic investigation of the properties that the boundaries of cells must be assumed to possess see the paper of Weyl already cited.

²¹) That an n -element cannot be defined to be any n -set whose skeleton is an n -element, and then proved to be an n -complex, is shewn by the simple example of a 1-cell bounded by a single 0-cell.

The arrays $\Delta\{a_k\}$, being complete stars, are elements (FI 12). If a_j bounds a_k , $\Delta\{a_j\}$ is contained in $\Delta\{a_k\}$: $P(\text{III})$ is satisfied. If an internal component of $\Delta\{a_j\}$ belongs to $\Delta\{a_k\}$ the vertex a_j belongs to $\Delta\{a_k\}$; i. e., in $\circ\Gamma$, a_j bounds a_k . Hence $\Delta\{a_j\}$, which does not contain a_k , belongs to $B \cdot \Delta\{a_k\}$. This is $P(\text{I})$. The vertices of a boundary face, U_{k-1} , of $\Delta\{a_k\}$ form a sequence descending from some $(k-1)$ -cell, a_{k-1} , and U_{k-1} is a unit of $\Delta\{a_{k-1}\}$. This is $P(\text{II})$.

Since all the pieces in this n -assembly are complete stars the " $\Delta\{a_k\}$ -cluster in $\Delta\Gamma$ " exists. (cf. Lemma 1).

Let $\circ\Gamma$ be a complex and a_k one of its cells. If the dimension numbers of all cells bounded by a_k are diminished by $k+1$ while the bounding relations between the cells are maintained, the modified cells form an $(n-k-1)$ -set called the *neighbour complex* of a_k in $\circ\Gamma$, ($\text{NC}(a_k)$ in $\circ\Gamma$)²². (The name "complex" will be justified presently.)

Lemma 3. *The necessary and sufficient condition that a_k should belong to $\circ\bar{\Gamma}$ is that $\text{NC}(a_k)$ should be bounded.*

(Obvious.)

Lemma 4. *If $\circ\Gamma$ is a proper complex, the shell of the $\Delta\{a_k\}$ -cluster in $\Delta\Gamma$ is the skeleton of $\text{NC}(a_k)$ in $\circ\Gamma$.*

Let $\Delta\Gamma_{n-k-1}$ be the shell of the $\Delta\{a_k\}$ -cluster in $\Delta\Gamma$. A unit, U_{n-k-1} , of $\Delta\text{NC}(a_k)$ has for vertices a descending sequence of cells $a_n, a_{n-1}, \dots, a_{k+1}$ all bounded by a_k . If then V_k is a unit of $\Delta\{a_k\}$, i. e. if the vertices of V_k form a sequence descending from a_k , $U_{n-k-1}V_k$ is a unit of $\Delta\Gamma$, and therefore U_{n-k-1} is a unit of $\Delta\Gamma_{n-k-1}$.

Conversely the vertices of a unit, U_{n-k-1} , of $\Delta\Gamma_{n-k-1}$ form a sequence descending from an n -cell a_n ; for if V_k is a unit of $\Delta\{a_k\}$ the vertices of $U_{n-k-1}V_k$ descend from a_n through a_k to a_0 , and V_k exhausts the first k dimension numbers. Hence U_{n-k-1} is a unit of $\Delta\{a_{n(-k-1)}\}$, where $a_{n(-k-1)}$ denotes a_n considered as an $(n-k-1)$ -cell of $\text{NC}(a_k)$.

An n ◦manifold is a connected n -complex in which the NC of every 0-cell is an $(n-1)$ ◦sphere or $(n-1)$ ◦element. From Lemma 3 it follows that an n ◦manifold is unbounded if, and only if, the NC of every 0-cell is an n ◦sphere. An n ◦manifold is clearly a proper set, and so Lemma 4 gives

Theorem 6a. *The skeleton of a n ◦manifold is a Δ manifold.*

Theorem 7a. *In an unbounded n ◦manifold $\text{NC}(a_k)$ is an $(n-k-1)$ ◦sphere.*

²² cf. H. Kneser, Proc. Amsterdām 27 (1924), p. 601.

It is now clear that if $\circ\Gamma$ is any complex $\text{NC}(a_k)$ is an $(n - k - 1)$ -complex²³). For if \bar{a}_j contains a_k the boundary of $a_{j(-k-1)}$ is all the cells $a_{i(-k-1)}$, where a_i is any cell bounding a_j and bounded by a_k ; i. e. $B(a_{j(-k-1)})$ is $\text{NC}(a_k)$ in \bar{a}_j . Since \bar{a}_j is an unbounded manifold this NC is a $(j - k - 2)$ \circ sphere.

Hence (by Lemma 4):

Theorem 6b. *A proper complex $\circ\Gamma$ is an n \circ manifold if $\Delta\Gamma$ is an n Δ manifold.*

Theorem 7b. *In a bounded manifold $\text{NC}(a_k)$ is an $(n - k - 1)$ \circ sphere or $(n - k - 1)$ \circ element.*

\circ spheres and \circ elements are \circ manifolds.

The definitions of n \circ assemblies and of *structural similarity* are derived from those on p. 403 by substituting " \circ element" for "element", " \circ piece" for "piece", "internal cell" for "unit or internal component" in $P(I)$, and " $(i - 1)$ -cell" for "unit" in $P(II)$.

Theorem 8. *The necessary and sufficient condition that the collection of \circ elements ($\circ G_i$) should be an n \circ assembly is that the collection (ΔG_i) should be an n Δ assembly.*

Suppose ($\circ G_i$) is known to be an n \circ assembly. Then the arrays ΔG_i are Δ elements.

$P(I)$: If ΔG_i has a component with vertices descending, say, from a_k which is a unit or internal component of ΔG_j , a_k is an internal cell of $\circ G_j$ (Theorem 2). Hence $\circ \bar{G}_i$ contains $\circ G_j$ and so $\Delta \bar{G}_i$ contains ΔG_j .

$P(II)$: If $a_{i-1} \dots a_0$ is a unit of $\Delta \bar{G}_i$, where $i > 0$, a_{i-1} is a boundary cell of $\circ G_i$, (proof of Theorem 3), and so belongs to a piece $\circ G_{j-1}$; and $a_{i-1} \dots a_0$ belongs to ΔG_{j-1} .

$P(III)$: If ΔG_i is any i Δ piece and $i < n$, $\circ G_i$ is contained in some $\circ G_{i+1}$, and ΔG_{i+1} contains ΔG_i .

The proof of the converse is precisely similar, using Theorem 3 (Corollary) and Theorem 2.

Theorem 9. *The necessary and sufficient condition that $\circ\Gamma$ and $\circ\Gamma'$ should have the same structure is that their skeleton Δ assemblies, $\Delta\Gamma$ and $\Delta\Gamma'$, should have the same structure.*

(Now obvious.)

A generalised n \circ assembly is a collection of \circ elements of dimension numbers from 0 to n , satisfying $P(I)$ and $P(II)$, (modified as above), but not necessarily $P(III)$.

²³ H. Kneser, *op. cit.*, pointed out that this need not be postulated in the definition.

If the generalised n -assemblies $\circ\Gamma^1$ and $\circ\Gamma^2$ have the same structure, and all the pieces of $\circ\Gamma^1$ are single cells, $\circ\Gamma^2$ is said to be obtained from $\circ\Gamma^1$ by *subdivision*. This is most conveniently denoted by using a small letter (α, β, \dots) for the structural relation. If $\circ\alpha\Gamma^1$ is $\circ\Gamma^2$, and $\circ\Gamma_k$ is any k -set in $\circ\Gamma^1$, $\circ\alpha\Gamma_k$ denotes the sum of the corresponding k -pieces in $\circ\Gamma^2$. In particular $\circ\beta\Gamma^1$ denotes a complex "constituted similarly" to the skeleton of $\circ\Gamma^{124}$, and $\circ\beta^r\Gamma^1$ is written for $\circ\beta\beta^{r-1}\Gamma^{125}$.

Topological Equivalence. Let $\circ\Gamma$ and $\circ\Gamma'$ be two n -complexes. If there exist n -complexes $\circ\Gamma^1, \circ\Gamma^2, \dots, \circ\Gamma^q$, ($\circ\Gamma^1$ is $\circ\Gamma$, $\circ\Gamma^q$ is $\circ\Gamma'$) with the property that $\circ\Gamma^i$ and $\circ\Gamma^{i+1}$ can be organised into generalised n -assemblies with the same structure, then $\circ\Gamma$ and $\circ\Gamma'$ are said to be *topologically equivalent*²⁶).

Theorem 10. *If $\circ M^1$ and $\circ M^2$ are n -manifolds, the necessary and sufficient condition for their topological equivalence is that $\Delta M^1 \xrightarrow{123} \Delta M^2$.*

That the condition is necessary follows from Theorem 1. For if a manifold is organised into an assembly satisfying $P(I)$ and $P(II)$, $P(III)$ is necessarily satisfied also.

If it is known that $\Delta M^1 \rightarrow \Delta M^2$, then Δ -subdivision-processes $\Delta\alpha$ and $\Delta\beta$ exist such that $\Delta\alpha M^1$ is $\Delta\beta M^2$ ²⁷). If $\circ M^3$ is a complex constituted similarly²⁸) to $\Delta\alpha M^1$ the Δ -processes $\Delta\alpha$ and $\Delta\beta$ serve in an obvious way as patterns for \circ -processes for dividing $\circ M^1$ and $\circ M^2$ into $\circ M^3$.

With the help of this theorem Theorems 6, 7, 8a, 8b, and Lemmas 7a and 7b of *FII*, quoted on p. 402, can be extended to complexes, provided that "is topologically equivalent to" is read for " \rightarrow "; also the following theorems from *S*:

Theorem 11²⁹). *If $\circ E_n^1$ and $\circ E_n^2$ are \circ -elements, and $\circ\alpha\bar{E}_n^1$ is $\circ\beta^r\bar{E}_n^2$, there is a division process $\circ\beta$ and an integer s such that $\circ\beta E_n^1$ is $\circ\beta^{r+s}E_n^2$, and $\circ\beta$ is $\circ\beta^s\alpha$ in \bar{E}_n^1 .*

²⁴) i. e., to each k -component, U_k , of $\Delta\Gamma^1$, there corresponds a k -cell, u_k , of $\circ\beta\Gamma^1$, and if U_j is a component of U_k , u_j bounds u_k .

²⁵) The corresponding notation for arrays was introduced in an earlier paper "On the superposition of n -dimensional manifolds", *Journal Lond. Math. Soc.* 2 (1927), p. 56—64, referred to as *S*. The only modification is that the statement " $\Delta\Gamma^1$ is $\Delta\Gamma^2$ " implies that all the pieces in $\Delta\Gamma^1$ are *simplexes*.

²⁶) cf. Weyl, *l. c.* The additional transformations allowed by Weyl (his axioms C and D) lead to no increased generality, with the present definitions, in view of *FII* Lemmas 7a and 7b (quoted on p. 402.).

²⁷) This is Theorem 2 of *S*.

²⁸) cf. f. n. ²⁴) above.

²⁹) *S*, Theorem 2, Case 2.

Theorem 12³⁰⁾. If $\circ M_n^1$ and $\circ M_n^2$ are topologically equivalent \circ manifolds, and if $\circ E_n^1$ and $\circ E_n^2$ are \circ elements, contained respectively in $\circ M_n^1$ and $\circ M_n^2$ but having no cell in their boundaries (if any), then $\circ M_n^1$ can be superposed on $\circ M_n^2$ so that $\circ E_n^1$ falls on $\circ E_n^2$; i. e. there exist a division process $\circ a$ and an integer r such that $\circ a M_n^1$ is $\circ \tilde{s}^r M_n^2$ and $\circ a E_n^1$ is $\circ \tilde{s}^r E_n^2$.

From Theorem 11 follows the general

Theorem on Superposition: If $\circ \Gamma$ and $\circ \Gamma'$ are any two equivalent n -complexes there is a division process $\circ a$ and an integer r such that $\circ a \Gamma$ is $\circ \tilde{s}^r \Gamma'$.

For if $\circ \Gamma$ and $\circ \Gamma'$ can themselves be organised into assemblies with the same structure, it is only necessary to "superpose" corresponding 1-, 2-, ..., n -pieces successively, making the structural relation between k -pieces agree with that already set up between their boundaries.

If $\circ \Gamma$ and $\circ \Gamma'$ are the end members of the chain $\circ \Gamma^1, \circ \Gamma^2, \dots, \circ \Gamma^q$, where $\circ \Gamma^i$ and $\circ \Gamma^{i+1}$ can be organised as assemblies with the same structure, $\circ \Gamma^{q-1}$ can first be superposed on $\circ \Gamma'$, giving $\circ \tilde{s}^{r_1} \Gamma'$; then $\circ \Gamma^{q-2}$ may be superposed on $\circ \tilde{s}^{r_1} \Gamma'$, giving $\circ \tilde{s}^{r_2} \Gamma'$; and so on.

Corollary. At most one intermediate complex is required to exhibit topological equivalence between two complexes.

³⁰⁾ S., Theorem 3.

(Eingegangen am 4. 4. 1927.)

Parallelverschiebung und Krümmungstensor.

Von

Ludwig Schlesinger in Gießen.

Der Zusammenhang, der zwischen dem Begriff der Parallelverschiebung, wie ihn Levi-Civita¹⁾ eingeführt hat, und dem Riemannschen Krümmungstensor für den Fall einer allgemeinen Riemannschen Mannigfaltigkeit besteht, ist schon von Levi-Civita selbst und nach ihm von zahlreichen anderen Autoren nachgewiesen und behandelt worden. Die einschlägigen Untersuchungen sind in fast alle Lehrbücher der Riemannschen Geometrie und der Relativitätstheorie übergegangen; ich nenne nur die bekannten Werke von Levi-Civita, Weyl, Eddington, Appell. Bei allen diesen Untersuchungen wird der Zuwachs, den die Komponenten des parallelverschobenen Vektors beim Durchlaufen einer geschlossenen Kurve erfahren, nur für eine unendlich kleine solche Kurve durch das Doppelintegral über den Krümmungstensor dargestellt, während die Untersuchung einer Kurve von endlicher Ausdehnung nur in den Arbeiten von Tietze²⁾ in Angriff genommen wird. Die Schwierigkeit besteht immer darin, daß bei den genannten Untersuchungen die gewöhnliche Integration (Quadratur) zur Anwendung gelangt, die aber, da es sich um die Lösung eines Systems von totalen linearen Differentialgleichungen handelt, nicht das geeignete Hilfsmittel für die Behandlung sein kann. Nimmt man dagegen an Stelle der Quadratur das Produktintegral, wie ich es im Anschluß an Volterra (1887; 1899) in älteren Arbeiten³⁾ eingeführt habe, so zeigt sich, daß mit Hilfe einer nahe liegenden Verallgemeinerung bzw. Übertragung des Stokesschen Satzes die Änderung der Vektorkomponenten bei Parallelverschiebung längs einer beliebigen geschlossenen Kurve direkt durch das Doppelintegral über den Krümmungstensor dargestellt werden kann.

¹⁾ T. Levi-Civita, Rendiconti Palermo 42 (1917), S. 173 ff.

²⁾ H. Tietze, Math. Zeitschr. 16 (1923), S. 308 ff.; Crelles Journal 153 (1924), S. 141 ff.

³⁾ Von 1903 ab, siehe besonders „Vorlesungen“, Leipzig 1908; vgl. auch „Differentialgleichungen“, 3. Aufl., Berlin 1922. Ich schließe mich im folgenden der an der letztgenannten Stelle benutzten Bezeichnungsweise für den Matrizenkalkül an.

I.

Die Differentialgleichungen der Parallelverschiebung eines Vektors in kontravarianten Komponenten u^1, \dots, u^m lauten bekanntlich:

$$(1) \quad du^k + \sum_{l=1}^m u^l \sum_{r=1}^n \Gamma_{lr}^k dt^r = 0 \quad (k = 1, 2, \dots, m).$$

Dabei haben wir die Christoffelschen Symbole Γ_{lr}^k als hinreichend oft differenzierbare Funktionen der n unabhängigen Veränderlichen t^1, \dots, t^n zu denken, für die aber vorerst weder ein Zusammenhang mit einem metrischen Tensor noch die sogenannte Symmetriebedingung $\Gamma_{lr}^k = \Gamma_{rl}^k$ vorausgesetzt werden soll, so daß also die Differentialgleichungen im Sinne von Weyl⁴⁾ als eine affine Übertragung definierend gelten können. Daß auch $m \geq n$ vorausgesetzt werden kann, ist eine naheliegende Verallgemeinerung, auf die kürzlich besonders Schouten⁵⁾ aufmerksam gemacht hat und für die er eine Note von R. König⁶⁾ zitiert. Wir gehen im Anschluß an Schouten von der Vorstellung aus, daß jedem Punkte (t^1, \dots, t^n) der amorphen n -fach ausgedehnten Mannigfaltigkeit \mathfrak{M}_n eine m -fach ausgedehnte Mannigfaltigkeit \mathfrak{G}_m von Vektoren mit den kontravarianten Komponenten u^1, \dots, u^m zugeordnet sei; der Zusammenhang zwischen den zu verschiedenen Punkten von \mathfrak{M}_n gehörigen Vektormannigfaltigkeiten \mathfrak{G}_m werde durch die Differentialgleichungen (1) in Verbindung mit einer Führungskurve \mathfrak{C} hergestellt, die wir uns durch Gleichungen $t^k = \varphi_k(s)$ ($k = 1, 2, \dots, n$) gegeben denken, wo die φ_k stetig differenzierbare Funktionen des Parameters s bedeuten. Um diesen Zusammenhang bequem analytisch formulieren zu können, denken wir uns in \mathfrak{G}_m ein System von m linear-unabhängigen Vektoren (kurz: eine *Vektorbasis*), deren Komponenten u_i^k eine Matrix U von m^2 Elementen bilden. Durch diese Basis ist dann jeder Vektor mit den Komponenten u^k von \mathfrak{G}_m in der Form $u^k = \sum_{i=1}^m c_i^k u_i^k$ darstellbar und der Übergang von der Basis mit der Komponentenmatrix U zu einer andern Basis mit der Komponentenmatrix \bar{U} erfolgt durch Gleichungen

$$\bar{u}_i^k = \sum_{\lambda=1}^m c_{i\lambda} u_\lambda^k,$$

die wir in die Matrizengleichung $\bar{U} = CU$ zusammenfassen, wo die $c_{i\lambda}$ Konstante (d. h. von den t^1, \dots, t^n unabhängige Größen) bedeuten, deren Determinante nicht verschwindet. — Es sei nun Γ_r die Matrix der n^2 Größen Γ_{lr}^k

⁴⁾ H. Weyl, *Raum, Zeit, Materie*, 5. Aufl., Berlin 1923, S. 113 ff.

⁵⁾ J. A. Schouten, *Rendiconti Palermo* 50 (1926), S. 142 ff.

⁶⁾ R. König, *Jahresbericht der Deutsch. Math.-Ver.* 28 (1919), S. 213 ff.

und $A = - \sum_{r=1}^n \Gamma_r \frac{dt^r}{ds}$; dann kann das Differentialsystem (1) für die Führungskurve \mathfrak{C} in der Form

$$(2) \quad \frac{dU}{ds} = - U \sum_{r=1}^n \Gamma_r \frac{dt^r}{ds} = UA$$

als Matrizengleichung oder, wenn wir wie üblich $U^{-1} \frac{dU}{ds} = D_s U$ setzen (zu lesen: derivierte Matrix von U in bezug auf s), noch konzentrierter in der Form

$$(2a) \quad D_s U = A$$

geschrieben werden. — Wird dann im Anfangspunkte der Führungskurve \mathfrak{C} , der dem Parameterwerte s_0 entspricht, die Vektorbasis, d. h. also ihre Komponentenmatrix U_0 willkürlich vorgeschrieben, so vollzieht sich die Übertragung dieser Basis nach dem Punkte s von \mathfrak{C} durch das bekannte Interpolationsverfahren (von Cauchy-Lipschitz), indem wir das Intervall (s_0, s) der Führungskurve durch die Zwischenwerte $s_1, \dots, s_{p-1}, s = s_p$ teilen und die Differentialgleichungen (2) durch die Differenzgleichungen

$$\begin{aligned} \frac{U_1 - U_0}{s_1 - s_0} &= U_0 A_0, & U_1 &= U_0 (A_0 (s_1 - s_0) + I), \\ \frac{U_2 - U_1}{s_2 - s_1} &= U_1 A_1, & U_2 &= U_1 (A_1 (s_2 - s_1) + I), \\ &\dots & & \dots \\ \frac{U_p - U_{p-1}}{s_p - s_{p-1}} &= U_{p-1} A_{p-1}, & U_p &= U_{p-1} (A_{p-1} (s_p - s_{p-1}) + I) \end{aligned}$$

approximieren, wo I die Einheitsmatrix ist und der untere Index k bedeuten soll, daß man die betreffende Matrix im Punkte s_k der Führungskurve zu nehmen hat. Es ergibt sich

$$(3) \quad U_p = U_0 \prod_{v=1, 2, \dots, p} \{A_{v-1} (s_v - s_{v-1}) + I\}.$$

Bei Verdichtung der Teilung strebt bekanntlich die Matrix U_p einer bestimmten endlichen Grenzmatrix U zu, die das Differentialsystem (2) befriedigt und sich für s_0 auf U_0 reduziert. Diese Grenzmatrix stellt die Komponentenmatrix der längs \mathfrak{C} von s_0 nach s übertragenen Vektorbasis dar. — Wir setzen

$$(4) \quad \lim_{p \rightarrow \infty} \prod_{v=1, 2, \dots, p} \{A_{v-1} (s_v - s_{v-1}) + I\} = \int_{s_0}^s (A ds + I)$$

(zu lesen: Produktintegral), dann ist also

$$(4a) \quad U = U_0 \int_{s_0}^s (A ds + I)$$

und wir können das Produktintegral (4) selbst als die *Änderung*

$$(4b) \quad \int_{s_0}^s (A ds + I) = U_0^{-1} U$$

auffassen, die die Komponentenmatrix erfahren hat, wenn die Vektorbasis von s_0 nach s hin längs \mathbb{C} sich selbst parallel verschoben worden ist. Für $m = 1$ reduziert sich das Produktintegral (4) auf

$$e^{s_0} \int A ds$$

Es ist nun die Untersuchung derjenigen Änderungen von besonderer Wichtigkeit, die eine Komponenten- oder, wie wir auch sagen, eine Integralmatrix erfährt, wenn die Übertragung längs einer *geschlossenen Kurve* zum Ausgangspunkte zurück erfolgt. Das erhellt ja schon daraus, daß gerade diese, geschlossenen Wegen entsprechenden Änderungen Aufschluß darüber geben, wie eine dem Übergang zwischen zwei verschiedenen Punkten entsprechende Änderung von der Wahl der Führungskurve abhängt. Um die gedachte Untersuchung durchführen zu können, bedürfen wir des folgenden Begriffs.

Wir betrachten eine τ^1, τ^2 -Ebene und in dieser eine geschlossene Kurve C , die von einer Parallelen zur τ^1 -Achse in höchstens zwei Punkten mit den Abszissen $\psi_1(\tau^2)$, $\psi_2(\tau^2)$ geschnitten wird. A bedeute eine Matrix von m^2 innerhalb und auf C integrierbaren Funktionen von τ^1, τ^2 , und es werde:

$$(5) \quad \int_{\tau_0^2}^{\tau_1^2} \left\{ \int_{\psi_1(\tau^2)}^{\psi_2(\tau^2)} A d\tau^1 d\tau^2 + I \right\} \\ = \lim \prod_k \sum_i \left\{ A(\sigma_{i-1}^1, \sigma_{k-1}^2) \cdot (\tau_i^1 - \tau_{i-1}^1) (\tau_k^2 - \tau_{k-1}^2) + I \right\}$$

betrachtet, wo τ_k^2 Teilpunkte des Intervalls $\tau_0^2 \dots \tau_1^2$, innerhalb dessen die Kurve C gelegen ist, bedeuten, τ_i^1 ebenso Teilpunkte des Intervalls $\psi_1(\tau_k^2) \dots \psi_2(\tau_k^2)$ und $\sigma_{i-1}^1, \sigma_{k-1}^2$ die Koordinaten eines beliebigen Punktes in dem Rechteck (i, k) . Wir bezeichnen dieses Gebilde^{?)} als *gemischtes Doppelintegral*:

$$\int_{(C)} \int (A d\tau^1 d\tau^2 + I).$$

Auch dieses reduziert sich dann für $m = 1$ auf

$$e^C \int \int A d\tau^1 d\tau^2$$

^{?)} Vgl. das von V. Volterra, *Memorie Soc. Ital. dei XL 6* (1887), S. 79 ff., eingeführte „integrale doppio di una sostituzione“.

Man kann diese Definition des gemischten Doppelintegrals in üblicher Weise auch auf den Fall ausdehnen, wo die Kurve C von einer Parallelen zur τ^1 -Achse in mehr als zwei Punkten getroffen wird.

Bei Einführung von neuen Veränderlichen durch die Gleichungen

$$\tau^i = \chi_i(\bar{\tau}^1, \bar{\tau}^2)$$

gilt genau wie für gewöhnliche Doppelintegrale die Transformationsformel:

$$(6) \quad \iint_{(C)} \{A d\tau^1 d\tau^2 + I\} = \iint_{(\bar{C})} \left\{ A \frac{\partial(\tau^1, \tau^2)}{\partial(\bar{\tau}^1, \bar{\tau}^2)} d\bar{\tau}^1 d\bar{\tau}^2 + I \right\},$$

wo \bar{C} die dem C entsprechende Kurve in der $\bar{\tau}^1, \bar{\tau}^2$ -Ebene und $\frac{\partial(\tau^1, \tau^2)}{\partial(\bar{\tau}^1, \bar{\tau}^2)}$ die Funktionaldeterminante bedeutet. Der Beweis kann ebenso geführt werden wie für das gewöhnliche Doppelintegral, da es sich ja nur um die Transformation des Produkts der Differentiale handelt.

II.

Wir stellen nun das Analogon des sog. *Greenschen Satzes* der gewöhnlichen Integralrechnung

$$(7) \quad \int_C (H_1 d\tau^1 + H_2 d\tau^2) = \iint_{(C)} \left(\frac{\partial H_2}{\partial \tau^1} - \frac{\partial H_1}{\partial \tau^2} \right) d\tau^1 d\tau^2$$

auf. Dabei können wir uns, was die Kurve C anlangt, auf die Betrachtung des durch die Ungleichungen $\tau_0^1 \leq \tau^1 \leq \tau_1^1$; $\tau_0^2 \leq \tau^2 \leq \tau_1^2$ definierten „Rechtecks“ \mathfrak{R} beschränken, da ja τ^1, τ^2 beliebige krummlinige Koordinaten bedeuten können und die bei dieser Wahl auftretenden Ecken durch Approximation stets abgerundet werden können⁸⁾. Die Schwierigkeit bei dem Beweise der der Greenschen Formel (7) analogen Gleichung besteht hauptsächlich darin, daß dem gemischten Doppelintegral ebenso wie dem Produktintegral die distributive Eigenschaft abgeht. Es mögen nun H_1, H_2 Matrizen von m^2 in dem Rechteck \mathfrak{R} stetig differenzierbaren Funktionen von τ^1, τ^2 bedeuten, dann handelt es sich also darum, das über die Begrenzung von \mathfrak{R} im positiven Sinne erstreckte Produktintegral

$$(8) \quad \begin{aligned} & \int \{H_1 d\tau^1 + H_2 d\tau^2 + I\} \\ &= \int_{\tau_0^1}^{\tau_1^1} (H_1(\tau^1, \tau_0^2) d\tau^1 + I) \cdot \int_{\tau_0^2}^{\tau_1^2} (H_2(\tau_1^1, \tau^2) d\tau^2 + I) \\ & \quad \cdot \int_{\tau_0^1}^{\tau_1^1} (H_1(\tau^1, \tau_1^2) d\tau^1 + I) \cdot \int_{\tau_1^2}^{\tau_0^2} (H_2(\tau_0^1, \tau^2) d\tau^2 + I) \end{aligned}$$

⁸⁾ Vgl. z. B. A. S. Eddington, *Relativitätstheorie*, Berlin 1925, S. 92.

in ein über das Innere von \mathfrak{R} erstrecktes gemischtes Doppelintegral umzuwandeln⁹⁾.

Wir betrachten die Matrix:

$$(9) \quad T = \int_{\tau_0^2}^{\tau^2} \{H_2(\tau_0^1, \tau^2) d\tau^2 + I\} \cdot \int_{\tau_0^1}^{\tau^1} \{H_1(\tau^1, \tau^2) d\tau^1 + I\},$$

wo τ^1, τ^2 irgendeinen Punkt im Innern von \mathfrak{R} bedeuten soll, und bilden

$$(10) \quad H_2 - D_{\tau^2} T.$$

Aus den Derivationsformeln

$$(\alpha) \quad D_{\tau} UV = V^{-1} \cdot D_{\tau} U \cdot V + D_{\tau} V,$$

$$(\beta) \quad D_{\tau} U^{-1} = -U \cdot D_{\tau} U \cdot U^{-1}$$

ergibt sich

$$D_{\tau} T H_2 T^{-1} = T H_2^{-1} \cdot D_{\tau} T \cdot H_2 T^{-1} + T \cdot D_{\tau} H_2 \cdot T^{-1} - T \cdot D_{\tau} T \cdot T^{-1},$$

also

$$\frac{\partial}{\partial \tau} (T H_2 T^{-1}) = T \left\{ D_{\tau} T \cdot H_2 + \frac{\partial H_2}{\partial \tau} - H_2 \cdot D_{\tau} T \right\} T^{-1}$$

und somit

$$\begin{aligned} \frac{\partial}{\partial \tau^1} (T (H_2 - D_{\tau^2} T) T^{-1}) &= T \left\{ D_{\tau^1} T \cdot H_2 - D_{\tau^1} T \cdot D_{\tau^2} T + \frac{\partial H_2}{\partial \tau^1} - \frac{\partial}{\partial \tau^1} D_{\tau^2} T \right. \\ &\quad \left. - H_2 \cdot D_{\tau^1} T + D_{\tau^2} T \cdot D_{\tau^1} T \right\} T^{-1}. \end{aligned}$$

Nun besteht für die Matrix T , deren Elemente ja Funktionen von τ^1, τ^2 sind, offenbar die Integrabilitätsbedingung:

$$\frac{\partial}{\partial \tau^1} D_{\tau^2} T + D_{\tau^1} T \cdot D_{\tau^2} T = \frac{\partial}{\partial \tau^2} D_{\tau^1} T + D_{\tau^2} T \cdot D_{\tau^1} T,$$

also folgt

$$(11) \quad \begin{aligned} \frac{\partial}{\partial \tau^1} (T (H_2 - D_{\tau^2} T) T^{-1}) \\ = T \left\{ D_{\tau^1} T \cdot H_2 + \frac{\partial H_2}{\partial \tau^1} - H_2 \cdot D_{\tau^1} T - \frac{\partial}{\partial \tau^2} D_{\tau^1} T \right\} T^{-1}. \end{aligned}$$

Da aber τ^1 im ersten Faktor von T nicht vorkommt und demnach

$$(12) \quad D_{\tau^1} T = H_1(\tau^1, \tau^2)$$

ist, so folgt aus (11)

$$(13) \quad \frac{\partial}{\partial \tau^1} (T (H_2 - D_{\tau^2} T) T^{-1}) = T \cdot P_{12} \cdot T^{-1},$$

wo

$$(14) \quad P_{12} = \frac{\partial H_2}{\partial \tau^1} - \frac{\partial H_1}{\partial \tau^2} + H_1 H_2 - H_2 H_1$$

⁹⁾ Vgl. V. Volterra, am unter ⁷⁾ angeführten Orte.

gesetzt wurde; die Elemente der Matrix P_{12} denken wir uns mit $P_{\beta 12}^\alpha$ ($\alpha, \beta = 1, 2, \dots, m$) bezeichnet. Wir können also schreiben

$$\int TP_{12}T^{-1}d\tau^1 = T(H_2 - D_{\tau^2}T)T^{-1},$$

und folglich

$$(15) \quad \int_{(9t)} (TP_{12}T^{-1}d\tau^1d\tau^2 + I) = \int_{\tau_0^2}^{\tau_1^2} \left(\int_{\tau_0^1}^{\tau_1^1} TP_{12}T^{-1}d\tau^1d\tau^2 + I \right) \\ = \int_{\tau_0^2}^{\tau_1^2} \left([T(H_2 - D_{\tau^2}T)T^{-1}]_{\tau_0^1}^{\tau_1^1} d\tau^2 + I \right).$$

Nun ist nach der Derivationsregel (α):

$$(16) \quad D_{\tau^2}T = \left(\int_{\tau_0^1}^{\tau_1^1} (H_1 d\tau^1 + I) \right)^{-1} \cdot H_2(\tau_0^1, \tau^2) \cdot \int_{\tau_0^1}^{\tau_1^1} (H_1 d\tau^1 + I) \\ + D_{\tau^2} \int_{\tau_0^1}^{\tau_1^1} (H_1 d\tau^1 + I).$$

Da aber der bei der Differentiation nach τ^2 auszuführende Grenzübergang mit dem Grenzübergang $\tau^1 \rightarrow \tau_0^1$ vertauschbar ist, so gilt:

$$(17) \quad \lim_{\tau^1 \rightarrow \tau_0^1} D_{\tau^2} \int_{\tau_0^1}^{\tau_1^1} (H_1 d\tau^1 + I) = 0.$$

Es folgt also aus (16)

$$(18) \quad \lim_{\tau^1 \rightarrow \tau_0^1} D_{\tau^2}T = H_2(\tau_0^1, \tau^2),$$

so daß im dritten Gliede der Gleichung (15) der Ausdruck unter dem Produktintegralzeichen an der unteren Grenze τ_0^1 verschwindet. Wir finden somit

$$(19) \quad \int_{(9t)} (TP_{12}T^{-1}d\tau^1d\tau^2 + I) = \int_{\tau_0^2}^{\tau_1^2} \left(\lim_{\tau^1 \rightarrow \tau_0^1} T(H_2 - D_{\tau^2}T)T^{-1}d\tau^2 + I \right).$$

Wir bilden nun

$$(20) \quad U = \int_{\tau_0^2}^{\tau_1^2} (T(H_2 - D_{\tau^2}T)T^{-1}d\tau^2 + I) \cdot T,$$

dann ist nach der Derivationsregel (α):

$$D_{\tau^2}U = T^{-1}(T(H_2 - D_{\tau^2}T)T^{-1})T + D_{\tau^2}T = H_2,$$

also

$$(21) \quad U = C \cdot \int_{\tau_0^2}^{\tau_1^2} (H_2 d\tau^2 + I),$$

wo C eine von τ^2 unabhängige Matrix bedeutet. Nun ist nach (20):

$$\lim_{\tau^2 \rightarrow \tau_0^2} U = C = \lim_{\tau^2 \rightarrow \tau_0^2} T = \int_{\tau_0^1}^{\tau^1} (H_1(\tau^1, \tau_0^2) d\tau^1 + I),$$

also

$$U = \int_{\tau_0^1}^{\tau^1} (H_1(\tau^1, \tau_0^2) d\tau^1 + I) \cdot \int_{\tau_0^2}^{\tau^2} (H_2 d\tau^2 + I).$$

Wir erhalten folglich, da nach (19)

$$\int_{\mathfrak{R}} (TP_{12} T^{-1} d\tau^1 d\tau^2 + I) = \lim_{\substack{\tau^1 \rightarrow \tau_1^1 \\ \tau^2 \rightarrow \tau_1^2}} (U \cdot T^{-1})$$

ist, für das über das Rechteck \mathfrak{R} erstreckte gemischte Doppelintegral die Darstellung:

$$\int_{\mathfrak{R}} (TP_{12} T^{-1} d\tau^1 d\tau^2 + I) = \int_{\tau_0^1}^{\tau_1^1} (H_1(\tau^1, \tau_0^2) d\tau^1 + I) \cdot \int_{\tau_0^2}^{\tau_1^2} (H_2(\tau_1^1, \tau^2) d\tau^2 + I) \cdot \left\{ \int_{\tau_0^2}^{\tau_1^2} (H_2(\tau_0^1, \tau^2) d\tau^2 + I) \cdot \int_{\tau_0^1}^{\tau_1^1} (H_1(\tau^1, \tau_1^2) d\tau^1 + I) \right\}^{-1}.$$

Die rechte Seite dieser Gleichung stimmt nun vollständig mit der rechten Seite der Gleichung (8) überein, wir erhalten also endlich:

$$(22) \quad \int_{\mathfrak{R}} (H_1 d\tau^1 + H_2 d\tau^2 + I) = \int_{\mathfrak{R}} (TP_{12} T^{-1} d\tau^1 d\tau^2 + I),$$

wo T und P_{12} bzw. durch die Gleichungen (9) und (14) gegeben werden. Dies ist das Analogon der Greenschen Formel (7). Wir bemerken, daß die Anwendung der vorstehenden Rechnung auf das im ersten Gliede von (7) auftretende gewöhnliche Doppelintegral die Greensche Formel ergibt, ohne daß von der distributiven Eigenschaft des Doppelintegrals Gebrauch gemacht wird.

III.

Aus der Gleichung (22) lassen sich ganz analoge Folgerungen ziehen, wie aus der Greenschen Formel (7) der gewöhnlichen Integralrechnung. Für den besonderen Fall, wo $P_{12} = 0$ ist, d. h. wo die Integrabilitätsbedingung für das Differentialsystem

$$du^k = \sum_{l=1}^m u^l H_{l1}^k d\tau^1 + \sum_{l=1}^m u^l H_{l2}^k d\tau^2 \quad (k=1, 2, \dots, m)$$

erfüllt sind, kann ich im wesentlichen auf die Erörterungen Seite 51–73 der „Vorlesungen“ von 1908 verweisen. Hier wollen wir den allgemeinen

Fall ins Auge fassen, wo $P_{12} \neq 0$ ist, und aus der abgeleiteten Formel das Analogon des *Stokesschen Satzes* beweisen. Wir bemerken zuvörderst, daß sich die Formel (22) ohne Schwierigkeit auch auf eine geschlossene Kurve C ausdehnen läßt, die von einer Parallelen zur τ^1 -Achse in mehr als zwei Punkten geschnitten wird.

Wir nehmen nun die allgemeinen Gleichungen (1) und betrachten eine zweifach ausgedehnte Mannigfaltigkeit

$$(23) \quad t^r = \varphi_r(\tau^1, \tau^2) \quad (r = 1, 2, \dots, n)$$

und in dieser wieder das Rechteck \mathfrak{R} . Dann ist;

$$(24) \quad dt^r = \alpha_r d\tau^1 + \beta_r d\tau^2,$$

$$(25) \quad \alpha_r = \frac{\partial t^r}{\partial \tau^1}, \quad \beta_r = \frac{\partial t^r}{\partial \tau^2},$$

und demnach

$$(26) \quad \frac{\partial \alpha_r}{\partial \tau^2} = \frac{\partial \beta_r}{\partial \tau^1}.$$

Setzen wir also

$$(27) \quad H_1 = - \sum \Gamma_r \alpha_r, \quad H_2 = - \sum \Gamma_r \beta_r,$$

so ist

$$(28) \quad \int_{\mathfrak{R}} \left(- \sum_{r=1}^m \Gamma_r dt^r + I \right) = \int_{\mathfrak{R}} (H_1 d\tau^1 + H_2 d\tau^2 + I),$$

also nach (22):

$$(29) \quad \int_{\mathfrak{R}} \left(- \sum_{r=1}^m \Gamma_r dt^r + I \right) = \int_{\mathfrak{R}} \int_{\mathfrak{R}} (TP_{12} T^{-1} d\tau^1 d\tau^2 + I),$$

wo T und P_{12} durch die Gleichungen (9) bzw. (14) gegeben sind. Nun ist nach (27):

$$(30) \quad \begin{aligned} H_1 H_2 - H_2 H_1 &= \sum_{r,s=1}^m \alpha_r \beta_s (\Gamma_r \Gamma_s - \Gamma_s \Gamma_r) \\ &= \sum_{r < s} (\alpha_r \beta_s - \alpha_s \beta_r) (\Gamma_r \Gamma_s - \Gamma_s \Gamma_r), \end{aligned}$$

und mit Rücksicht auf (26)

$$(31) \quad \frac{\partial H_2}{\partial \tau^1} - \frac{\partial H_1}{\partial \tau^2} = \sum_{r < s} (\alpha_r \beta_s - \beta_r \alpha_s) \left(\frac{\partial \Gamma_r}{\partial \tau^s} - \frac{\partial \Gamma_s}{\partial \tau^r} \right).$$

Wir erhalten also, wenn wir im Anschluß an die übliche Bezeichnung¹⁰⁾

$$(32a) \quad \frac{\partial \Gamma_{\beta s}^{\alpha}}{\partial t^r} - \frac{\partial \Gamma_{\beta r}^{\alpha}}{\partial t^s} + \sum_{\lambda=1}^m (\Gamma_{\beta s}^{\lambda} \Gamma_{\lambda r}^{\alpha} - \Gamma_{\beta r}^{\lambda} \Gamma_{\lambda s}^{\alpha}) = R_{\beta r s}^{\alpha} \quad (\alpha, \beta = 1, 2, \dots, m)$$

¹⁰⁾ Vgl. etwa P. Appell, *Mécanique rationnelle* V, Paris 1926, S. 56.

oder in Matrizenbezeichnung

$$(32b) \quad \frac{\partial \Gamma_s}{\partial t^r} - \frac{\partial \Gamma_r}{\partial t^s} + \Gamma_s \Gamma_r - \Gamma_r \Gamma_s = R_{rs} = -R_{sr}$$

setzen,

$$(33) \quad P_{12} = \sum_{r < s} (\alpha_r \beta_s - \alpha_s \beta_r) R_{sr},$$

was in (29) eingesetzt

$$(34) \quad \widehat{\int}_{\mathfrak{R}} \left(- \sum_{r=1}^m \Gamma_r dt^r + I \right) = \widehat{\int}_{(\mathfrak{R})} \int_{r < s} \left(T \sum_{r < s} (\alpha_r \beta_s - \alpha_s \beta_r) R_{sr} \cdot T^{-1} dt^1 dt^2 + I \right)$$

ergibt. Mit Rücksicht auf die Gleichungen

$$(35) \quad \alpha_r \beta_s - \alpha_s \beta_r = \frac{\partial (t^r, t^s)}{\partial (\tau^1, \tau^2)}$$

und auf die Transformationsformel (6) des gemischten Doppelintegrals kann die Gleichung (34) auch in der Form

$$(34a) \quad \widehat{\int}_{\mathfrak{R}} \left(- \sum_{r=1}^m \Gamma_r dt^r + I \right) = \widehat{\int}_{(\mathfrak{R})} \int_{r < s} \left(T \cdot \sum_{r < s} R_{sr} \cdot T^{-1} dt^r dt^s + I \right)$$

geschrieben werden, in der sie das vollständige Analogon des *Stokesschen Satzes* der gewöhnlichen Integralrechnung darstellt. Sie zeigt, daß das über den Rand von \mathfrak{R} erstreckte Produktintegral unabhängig ist von der Wahl der durch diesen Rand hindurch gelegten zweifach ausgedehnten Mannigfaltigkeit (23). Für die Matrix T ergibt sich, indem man in den Ausdruck (9) die Werte (27) einsetzt, die Darstellung:

$$(36) \quad T = \widehat{\int}_{\tau_0^1}^{\tau^1} \left(- \sum_{r=1}^m (\Gamma_r)_{\tau^1=\tau_0^1} dt^r + I \right) \cdot \widehat{\int}_{\tau_0^1}^{\tau^1} \left(- \sum_{r=1}^m \Gamma_r dt^r + I \right).$$

Zunächst zeigt die Gleichung (34a), daß, wenn die Gleichungen

$$(37) \quad R_{rs} = 0 \quad (r, s = 1, 2, \dots, n)$$

bestehen, das über irgendeine geschlossene Kurve erstreckte Produktintegral gleich I wird; also ist in diesem Fall das zwischen zwei Punkten von \mathfrak{M}_n erstreckte Produktintegral unabhängig von der Führungskurve und somit eine bloße Funktion der Koordinaten des Anfangs- und Endpunkts. Das Differentialsystem (1) ist also, wenn die Gleichungen (37) bestehen, *unbeschränkt integrierbar*, d. h. es besitzt Lösungen, die Funktionen der n unabhängigen Veränderlichen t^1, \dots, t^n sind. Daß das Bestehen dieser sogenannten Integrabilitätsbedingungen für das Vorhandensein solcher Lösungen auch notwendig ist, ergibt sich unmittelbar aus der Vertauschbarkeit der Differentiationsfolgen. Wenn also die Bedingungen (37) erfüllt sind, so sind die partiellen Differentialgleichungen

$$(38) \quad \frac{\partial u^k}{\partial t^r} = - \sum_{l=1}^m u^l \Gamma_{lr}^k \quad \left(\begin{array}{l} k = 1, 2, \dots, m \\ r = 1, 2, \dots, n \end{array} \right)$$

miteinander verträglich und wir kommen so zu dem von Fuchs und mehreren anderen Mathematikern eingehend behandelten Fall simultaner linearer partieller Differentialsysteme¹¹⁾. *Geometrisch* bedeutet das Bestehen der Gleichungen (37), daß der sogenannte *Riemannsche Krümmungstensor* von \mathfrak{M}_n , dessen Komponenten ja nichts anderes sind, als die Größen (32a), verschwindet, d. h. daß wir es in diesem Fall mit einer Euklidischen oder ebenen Mannigfaltigkeit zu tun haben. Dabei ist freilich noch immer der allgemeine Gesichtspunkt einer Weylschen Mannigfaltigkeit, in der bloß ein affiner Zusammenhang ohne Metrik herrscht, festgehalten und sogar noch verallgemeinert, indem die Γ_{lr}^k weder der Bedingung $m = n$ noch der Symmetriebedingung unterworfen zu sein brauchen. — Wir kommen nachher noch einmal kurz auf diesen Euklidischen Fall zurück und wenden uns nun wieder dem allgemeinen Falle zu.

In der Gleichung (34a) steht links vom Gleichheitszeichen die *Änderung*, die die Komponentenmatrix U einer Vektorbasis erfährt, wenn diese Basis längs der geschlossenen Kurve \mathfrak{R} zu ihrem Anfangspunkte zurück sich selbst parallel verschoben wird. Diese Änderung ist also direkt dargestellt durch ein gemischtes Doppelintegral, das über eine an sich beliebige, in die Kurve \mathfrak{R} eingespannte Fläche (23) zu erstrecken ist, und das sich auf die zu den Indexpaaren r, s gehörigen Komponenten des Riemannschen Krümmungstensors bezieht. Wir wollen uns zuvörderst die Frage vorlegen, inwieweit der Änderung der Komponentenmatrix U eine *geometrische Bedeutung* innerhalb der Vektormannigfaltigkeit \mathfrak{E}_n als Änderung der Vektorbasis selbst unterlegt werden kann. Hierzu bedarf es einiger begrifflicher Feststellungen, denen wir uns jetzt zuwenden wollen.

IV.

In der zu einem Punkte t^1, \dots, t^m von \mathfrak{M}_n gehörigen Vektormannigfaltigkeit \mathfrak{E}_m definieren wir den Übergang von einem Komponentensystem u^k eines Vektors zu einem andern v^k durch Gleichungen von der Form

$$(39) \quad v^k = \sum_{i=1}^m u^i h_{ik} \quad (k = 1, 2, \dots, m),$$

wo die h_{ik} hinreichend oft differenzierbare Funktionen von t^1, \dots, t^m bedeuten mögen, deren Determinante nicht identisch verschwindet, so daß also die Gleichungen (39) nach den u^k auflösbar sind. Wir können diese Gleichungen als eine *Koordinatentransformation* in \mathfrak{E}_m und demgemäß alles, was für diese Transformationen invariant ist, als dem betreffenden

¹¹⁾ L. Fuchs (1892), Werke III, S. 117 ff.; vgl. J. Horn, Acta Mathematica 12 (1891), S. 113 ff; u. a.

Vektor *geometrisch* eigentümlich auffassen. — Für eine Vektorbasis geschrieben lauten die Gleichungen (39):

$$(39a) \quad v_i^k = \sum_{\lambda=1}^n u_i^\lambda h_{\lambda k} \quad (i, k = 1, \dots, m)$$

oder kürzer in **Matrizenform**

$$(39b) \quad V = UH.$$

Wenn wir die linken Seiten der Differentialgleichungen (1) in **Matrizenform** zusammenfassen, so lauten sie:

$$(40) \quad dU + U \sum_{r=1}^n \Gamma_r dt^r;$$

hierin setzen wir für U seinen aus (39b) entnommenen Wert $U = VH^{-1}$, dann erhalten wir

$$(41) \quad dU + U \sum_{r=1}^n \Gamma_r dt^r = dV \cdot H^{-1} + V \cdot d(H^{-1}) + V \cdot H^{-1} \sum_r \Gamma_r dt^r, \\ = (dV + V \sum_r B_r dt^r) H^{-1},$$

wo gesetzt wurde

$$\sum_r B_r dt^r = d(H^{-1}) \cdot H + H^{-1} \sum_r \Gamma_r dt^r \cdot H,$$

also nach der Derivationsregel (β)

$$(42) \quad \sum_r B_r dt^r = -H^{-1} dH + H^{-1} \sum_r \Gamma_r dt^r \cdot H.$$

Die B_r sind Matrizen von Funktionen der t^1, \dots, t^n vom selben Charakter wie die $\Gamma_{lr}^k, h_{\lambda k}$, die Gleichungen (42) können also auch in der Form

$$(42a) \quad B_r = H^{-1} \Gamma_r H - D_{t^r} H$$

oder

$$(42b) \quad \frac{\partial H}{\partial t^r} = \Gamma_r H - H B_r \quad (r = 1, \dots, n)$$

geschrieben werden. — Im Anschluß an die in der Theorie der linearen Differentialgleichungen übliche Poincarésche Terminologie sagen wir von den beiden Differentialsystemen

$$(43) \quad dU + U \sum_r \Gamma_r dt^r = 0,$$

$$(44) \quad dV + V \sum_r B_r dt^r = 0$$

oder auch von den ihre linken Seiten bildenden Differentialausdrücken, daß sie zu derselben *Art* gehören. Damit also zwei Differentialgleichungen von der Form (43), (44), wo die Γ_{lr}^k, B_{lr}^k Funktionen der t^1, \dots, t^n be-

deuten, zu derselben Art gehören, ist notwendig und hinreichend, daß die Differentialsysteme (42 b) für die Matrix H verträglich sind oder, was dasselbe heißt, daß das totale Differentialsystem

$$(42 c) \quad dH = \sum_r \Gamma_r dt^r \cdot H - H \sum_r B_r dt^r$$

komplett integrierbar sei. — Hiernach können wir sagen, daß die Frage nach den geometrischen Eigenschaften und Begriffen innerhalb der Mannigfaltigkeit \mathfrak{C}_m gleichbedeutend ist mit der Frage nach den *Artinvarianten*.

Da ist nun zunächst leicht einzusehen, daß die Gleichung (34 a) selbst eine Artinvariante ist, also geometrische Bedeutung besitzt. Man braucht dazu nur auszurechnen, wie sich ihre beiden Seiten verhalten, wenn man durch (39 b) den Übergang von dem Differentialsystem (43) zu (44) vollzieht. Dazu bemerken wir, daß für zwei Punkte p, q einer Führungskurve \mathfrak{C}

$$(45) \quad \int_p^q (A \cdot ds + I) = U_p^{-1} U_q,$$

so daß also

$$\int_p^q (-\sum_r B_r dt^r + I) = V_p^{-1} V_q$$

und mit Rücksicht auf (39 b) und (45)

$$(46) \quad \int_p^q (-\sum_r B_r dt^r + I) = H_p^{-1} \int_p^q (-\sum_r \Gamma_r dt^r + I) \cdot H_q$$

ist. Wenn sich die Führungskurve insbesondere auf eine geschlossene \mathfrak{H} reduziert, so daß die Punkte p, q in einen und denselben Punkt (t_0^1, \dots, t_0^n) oder kurz (0) zusammenfallen, so hat man also

$$(47) \quad \int_{\mathfrak{H}} (-\sum_r B_r dt^r + I) = H_0^{-1} \int_{\mathfrak{H}} (-\sum_r \Gamma_r dt^r + I) \cdot H_0,$$

d. h. das Produktintegral wird, wie man sagt, mit der Matrix H_0^{-1} transformiert. Läßt man die Integration in dem durch (36) gegebenen Produktintegral T von dem Punkte (0) ausgehen, so hat man in leichtverständlicher Bezeichnungsweise

$$T_B = H_0^{-1} T_\Gamma H_0,$$

und für die R_{rs} folgt direkt mit Rücksicht auf (42 a)

$$(48) \quad (R_{rs})_B = H_0^{-1} (R_{rs})_\Gamma H_0,$$

so daß sich also beim Übergang von (43) zu (44) durch die Transformation (39 b) tatsächlich beide Seiten von (34 a) mit der Matrix H_0^{-1} transformieren.

Für eine von dem Punkte (0) ausgehende geschlossene Kurve \mathfrak{K} ist die Änderung der Komponentenmatrix U eine bestimmte Matrix von m^2 Elementen; diese wird beim Übergang zu einem anderen Komponentensystem V mit H_0^{-1} transformiert, sie hat also selbst keine geometrische Bedeutung, wohl aber gilt das für ihre *Elementarteiler*, die ja die Invarianten für die Transformation mit irgendeiner Matrix sind. — Zu einem allgemeineren Begriffe gelangen wir, wenn wir die *Gesamtheit* aller von dem Punkte (0) ausgehenden geschlossenen Kurven ins Auge fassen.

Wir wollen zwei solche geschlossene Kurven als äquivalent betrachten, wenn für sie die rechte Seite der Gleichung (34 a) denselben Wert hat, d. h. mit anderen Worten, wenn ihnen dieselbe Änderung der Komponentenmatrix entspricht. Hiernach geht z. B. eine geschlossene Kurve in eine äquivalente über, wenn man sie durch einen an sie anstoßenden „ebenen“ Bereich von \mathfrak{M}_n hindurch stetig deformiert. Beachtet man dann noch, daß nach (45) die Änderung, die der im negativen Sinne beschriebenen Kurve \mathfrak{K} entspricht, die Inverse der Änderung ist, die bei positiver Durchlaufung von \mathfrak{K} entsteht, so erkennt man, daß die allen möglichen von (0) ausgehenden geschlossenen Kurven entsprechenden Änderungen der Matrix U eine Gruppe \mathfrak{G} mit paarweise inversen Elementen bilden. Nach Gleichung (47) wird bei einer Koordinatentransformation (39) jedes Element dieser Gruppe mit H_0^{-1} transformiert. Ohne die Allgemeinheit wesentlich einzuschränken, können wir es so einrichten, daß sowohl U als auch V sich im Punkte (0) auf die Einheitsmatrix I reduzieren, dann wäre also $H_0 = I$, d. h. die Gruppe \mathfrak{G} bliebe ungeändert; wir wollen darum übereinkommen, daß wir Gruppen, die durch Transformation aller ihrer Elemente durch eine und dieselbe konstante Matrix auseinander hervorgehen, als nicht wesentlich verschieden ansehen. Dann können wir also sagen, daß die Gruppe \mathfrak{G} eine Artinvariante darstellt, oder daß ihr in \mathfrak{G}_m geometrische Bedeutung zukommt.

Die analytische Bedeutung von \mathfrak{G} besteht im folgenden. Denken wir uns die Komponentenmatrix U im Ausgangspunkte (0) gleich I gewählt, so hängt im allgemeinen, d. h. wenn der Krümmungstensor von Null verschieden ist, ihr Wert U_t in einem Punkte (t^1, \dots, t^n) , oder kurz (t) , von dem Wege \mathfrak{C} ab, der von (0) nach (t) hinführt, und es ist

$$U_t = \mathfrak{C} \int_0^t \left(- \sum_{r=1}^n \Gamma_r \frac{dt^r}{ds} ds + I \right).$$

Auf einem anderen Wege \mathfrak{C}' erstreckt, ergibt sich offenbar in (t) für U der Wert CU_t , wenn wir durch C diejenige Matrix der Gruppe \mathfrak{G} bezeichnen, die der geschlossenen Kurve entspricht, die entsteht, wenn wir

erst auf \mathcal{C}' von (0) nach (t) und dann auf \mathcal{C} von (t) nach (0) zurückgehen. Hat man nun einen von den u_i^k abhängenden Ausdruck $\Phi(U)$, der die Eigenschaft besitzt, daß seine Werte in allen Punkten (t) von der Führungskurve \mathcal{C} unabhängig sind, der also eine Funktion der unabhängigen Veränderlichen t^1, \dots, t^n darstellt, so muß $\Phi(U)$ ungeändert bleiben, wenn U von links her mit irgendeiner Matrix von \mathcal{G} komponiert wird, und umgekehrt ist eine Funktion $\Phi(U)$, die in bezug auf alle Matrizen von \mathcal{G} diese Eigenschaft besitzt, eine bloße Funktion von t^1, \dots, t^n . Wir können demnach \mathcal{G} mit dem in der Theorie der linearen Differentialgleichungen für eine komplexe Variable üblichen Namen der *Monodromiegruppe* belegen, weil diese Gruppe in der Tat ganz analoge Eigenschaften besitzt, wie sie in jener Theorie der Monodromiegruppe zukommen. — Von maßgebender Bedeutung für die Aufstellung dieser Gruppe ist der topologische Charakter des Gebiets der Mannigfaltigkeit \mathfrak{M}_n , in dem sich die Untersuchung abspielt; man erhält ein instruktives Beispiel, wenn man das Differentialsystem in zwei unabhängigen Variablen nimmt, das entsteht, indem in einem linearen Differentialsystem für eine komplexe Veränderliche etwa die reellen Teile der abhängigen Variablen für sich betrachtet werden. Die Mannigfaltigkeit \mathfrak{M}_n hat dann zwar euklidischen Charakter, da die Integrabilitätsbedingungen erfüllt sind, sie ist aber im allgemeinen nicht einfach zusammenhängend¹²⁾.

V.

Wir wollen in dieser Note nur noch die vorstehend entwickelten Generalitäten auf den einfachen Fall der Gaußschen Flächentheorie anwenden, um an diesem Beispiel in Evidenz zu setzen, daß die von uns zur Anwendung gebrachten Methoden die allein dem Gegenstande angemessenen sind; die eingehende Behandlung des für die theoretische Physik wichtigsten Falles, wo $n = m = 4$ ist, muß einer späteren Gelegenheit vorbehalten bleiben. Um die Anwendung auf die Gaußsche Flächentheorie geben zu können, bedürfen wir nur noch einiger formaler Betrachtungen, die hier zunächst folgen mögen.

Neben den kontravarianten Komponenten u^k eines Vektors betrachtet man noch die kovarianten Komponenten u_k , die¹³⁾ dadurch definiert werden können, daß das Aggregat $\sum_{k=1}^m u^k u_k$ bei einer simultanen Parallel-

¹²⁾ L. Schlesinger, Vorlesungen, 1908, 4. und 5. Vorlesg., besonders S. 68 ff.

¹³⁾ H. Weyl, loc. cit. 4), S. 114. Die Bezeichnung als kontravariante bzw. kovariante Komponenten wird hier nur formal von dem Falle $m = n$ her übertragen; in diesem Falle kann nämlich $u^k = dt^k/ds$ gesetzt werden, so daß sich bei einer Änderung der Veränderlichen t^1, \dots, t^n die u^k ebenso transformieren wie die Differentiale dt^k .

verschiebung des Vektors mit den kontravarianten Komponenten u^k und des Vektors mit den kovarianten Komponenten u_k längs derselben Führungskurve ungeändert, also absolut konstant bleiben soll. Durch Differentiation der Gleichung

$$(49) \quad \sum_{k=1}^m u^k u_k = \text{konst.}$$

folgt mit Rücksicht auf die Differentialgleichungen (1) für die u^k das Differentialsystem

$$(50) \quad du_k - \sum_{l=1}^m u_l \sum_{r=1}^n \Gamma_{kr}^l dt^r = 0,$$

das man als das zu (1) *adjungierte* bezeichnet¹⁴⁾. Dabei haben wir vorerst noch kein Merkmal dafür, welche Lösungen von (50) als die kovarianten Komponenten desselben Vektors anzusehen sind, dessen kontravariante Komponenten durch ein bestimmtes Lösungssystem von (1) gegeben werden.

Ist $U = (u_i^k)$ die kontravariante Komponentenmatrix einer Vektorbasis, so bildet, wie eine sehr bekannte kleine Rechnung zeigt, die aus U^{-1} durch Transposition (d. h. Vertauschung von Zeilen und Spalten) hervorgehende Matrix eine Integralmatrix von (50), also die kovariante Komponentenmatrix einer Vektorbasis. Wenn wir den Übergang von einer Matrix zu ihrer Transponierten durch einen Überstrich andeuten, so können wir das System (50) in Matrizenform als

$$(51) \quad d\tilde{U} - \tilde{U} \sum_{r=1}^n \bar{\Gamma}_r dt^r = 0$$

schreiben, ähnlich wie (1) in der Form (43) geschrieben wurde, und es ist

$$(52) \quad \tilde{U} = C \cdot (\bar{U})^{-1},$$

wo C eine willkürliche konstante Matrix bedeutet.

Besteht zwischen zwei Systemen (43), (44) die Artbeziehung (39 b), so gehören auch die adjungierten Systeme (51) und

$$(53) \quad d\tilde{V} - \tilde{V} \sum_{r=1}^n \bar{B}_r dt^r = 0$$

zu derselben Art, indem nämlich die Beziehung gilt:

$$(54) \quad \tilde{V} = \tilde{U}(\bar{H})^{-1}.$$

Eine Verallgemeinerung der adjungierten Systeme ergibt sich, wenn man das System betrachtet, dem die sämtlichen $(m_p)^2$ Determinanten p -ter

¹⁴⁾ T. Levi-Civita, loc. cit. ¹⁾, § 5, S. 180 ff.

Ordnung genügen, die aus der rechteckigen Matrix von p unabhängigen Lösungssystemen von (1) gebildet werden können. Man nennt es das $(m - p)$ -te zu (1) assoziierte System¹⁵⁾; das erste assoziierte System geht aus dem adjungierten hervor, indem man alle u_k mit der Determinante

$$(55) \quad \Delta = |u_i^k| = e^{-\int \sum_{r,k} \Gamma_{rk}^k dt^r} \quad (i, k = 1, 2, \dots, m)$$

multipliziert. Man kann hiernach neben den kovarianten und kontravarianten Komponenten eines Vektors noch $m - 2$ Zwischenstufen einschalten, entsprechend den in der analytischen Geometrie gebräuchlichen Koordinaten der linearen Gebilde verschiedener Stufen. — In bezug auf die assoziierten Systeme bemerken wir nur noch, daß ein System, das aus dem adjungierten des $(m - p)$ -ten assoziierten Systems dadurch hervorgeht, daß man alle abhängigen Variablen mit einer geeigneten Potenz von Δ multipliziert, mit dem p -ten assoziierten zur selben Art gehört. Wenn also $m = 2q$ eine gerade Zahl ist, so gehört das q -te assoziierte System mit seinem adjungierten, nach Multiplikation aller abhängigen Veränderlichen mit einer Potenz von Δ , zu derselben Art¹⁶⁾.

Die letztere Bemerkung legt es nahe¹⁷⁾, überhaupt nach den Differentialsystemen der Form (1) zu fragen, die mit ihrem adjungierten zu derselben Art gehören. — Wir wollen also annehmen, daß die Differentialsysteme (43) und (51) zur selben Art gehören, d. h. daß ihre linken Seiten durch

$$(56) \quad \tilde{U} = U \cdot G$$

ineinander übergehen, wo G eine Matrix von Funktionen der t^1, \dots, t^n vom Charakter der Γ_{rk}^k bedeutet; es sollen also, mit anderen Worten, die Elemente der Matrix $\tilde{U}U^{-1} = C(\bar{U})^{-1}U^{-1} = C(U \cdot \bar{U})^{-1}$ so beschaffen sein, daß sie ungeändert bleiben, wenn man U von links her mit einer Matrix der Gruppe \mathfrak{G} komponiert. — Da ja die Beziehung zwischen adjungierten Systemen eine gegenseitige ist, so folgt, ebenso wie (54) aus (39 b), aus (56) die Beziehung $U = \tilde{U}(\bar{G})^{-1}$, und da andererseits (56) direkt $U = \tilde{U}G^{-1}$ ergibt, so schließen wir, daß $(\bar{G})^{-1} = G^{-1}$ oder $\bar{G} = G$ sein muß, d. h. es ist

$$(57) \quad g_{ik} = g_{ki} \quad (i, k = 1, \dots, m),$$

¹⁵⁾ L. Schlesinger, Crelles Journal 128 (1905), S. 296; daselbst ist auch die explizite Form der assoziierten Systeme gegeben, jedoch fehlt in der ersten Formel S. 269 rechts vom Gleichheitszeichen der Faktor $(-1)^{2+\mu}$.

¹⁶⁾ L. Schlesinger, Handbuch II, 1, Leipzig 1897, Nr. 171, S. 142 ff.

¹⁷⁾ L. Fuchs (1899), Werke III, S. 300 ff.; R. Fuchs, Crelles Journal 121 (1899), S. 205 ff., ebenda 123 (1900), S. 54 ff.; G. Fano, Math. Annalen 53 (1899), S. 568 ff.; A. Loewy, Münchener Berichte 32 (1902), S. 3 ff.

die Matrix G ist, wie man sagt, *symmetrisch*. Nimmt man in der Matrixgleichung (56) beiderseits die Determinanten, so erhält man die Gleichung

$$(58) \quad \Delta = |u_i^k| = \frac{1}{\sqrt{g}}, \quad g = |g_{ik}|,$$

die zeigt, daß die in den Aussagen über die Beziehungen der assoziierten Systeme zu ihren adjungierten auftretende Determinante Δ in dem jetzt betrachteten Falle einfach durch die Quadratwurzel aus g vertreten werden kann.

Man sagt, wenn das Differentialsystem der Parallelverschiebung (1) mit seinem adjungierten zu derselben Art gehört, daß ein *metrischer Tensor* mit den Komponenten g_{ik} vorhanden sei; nach der Gleichung (42) ist dafür notwendig und hinreichend, daß das totale Differentialsystem

$$(59) \quad dG = G \left(\sum_{r=1}^m \bar{\Gamma}_r dt^r \right) + \left(\sum_{r=1}^m \Gamma_r dt^r \right) G$$

komplett integrierbar, oder daß die partiellen Differentialsysteme

$$(59a) \quad \frac{\partial G}{\partial t^r} = G \bar{\Gamma}_r + \Gamma_r G \quad (r = 1, \dots, n)$$

miteinander verträglich seien. Wir können jetzt in üblicher Weise festsetzen, daß die kontravarianten und kovarianten Komponenten einer und derselben Vektorbasis durch die Gleichungen (56) miteinander verknüpft sein mögen, d. h. daß zwischen den kontravarianten und kovarianten Komponenten eines Vektors die Gleichungen

$$(56a) \quad u_k = \sum_{\lambda} u^{\lambda} g_{\lambda k}, \quad u^k = \sum_{\lambda} u_{\lambda} g^{\lambda k}$$

bestehen sollen, wo also $g^{\lambda k}$ die Elemente von G^{-1} bedeuten. — Setzen wir die aus den Gleichungen (56a) folgenden Ausdrücke in die Relation (49) ein, so finden wir für das Differentialsystem (1) und das adjungierte

$$(60) \quad \sum_{i,k} g_{ik} u^i u^k = \sum_{i,k} g^{ik} u_i u_k = \text{konst.},$$

eine Gleichung, die dem Integral der lebendigen Kraft analog ist. Man bezeichnet die nach (60) bei einer Parallelverschiebung invariante quadratische Form¹⁸⁾ der Komponenten u^k bzw. u_k als das Quadrat der Länge l_u des durch diese Komponenten bestimmten Vektors; damit hängt auch die Bezeichnung von g_{ik} als metrischer Tensor zusammen. Man kann jetzt auch durch die Formel

$$(61) \quad \cos \theta = \frac{1}{l_u l_v} \sum_{i,k} g_{ik} u^i v^k = \frac{1}{l_u l_v} \sum_{i,k} g^{ik} u_i v_k$$

¹⁸⁾ L. Fuchs (1888), Werke III, S. 7 ff.

den Neigungswinkel θ zweier Vektoren definieren, der dann auch für Parallelverschiebung invariant ist. — Wir stellen noch einige für das Folgende nötige Formeln¹⁹⁾ zusammen, die zwischen den g_{ik} und den Christoffelschen Symbolen Γ_{lr}^k bestehen. Setzt man wie üblich

$$(62) \quad \Gamma_{k,iv} = \sum_l \Gamma_{iv}^l g_{lk}, \quad \Gamma_{iv}^k = \sum_l \Gamma_{l,iv} g^{lk},$$

so schreibt sich die Gleichung (59a)

$$(59b) \quad \frac{\partial g_{ik}}{\partial t^r} = \Gamma_{i,kr} + \Gamma_{k,ir},$$

der dann auch die analoge Gleichung

$$(59c) \quad \frac{\partial g^{ik}}{\partial t^r} + \sum_l g^{il} \Gamma_{lr}^k + \sum_l \Gamma_{lr}^i g^{lk} = 0$$

hinzugesellen ist. Für die Determinante g ergibt sich

$$(63) \quad \frac{\partial \log g}{\partial t^r} = 2 \sum_k \Gamma_{kr}^k.$$

Es gilt nun, unter der Voraussetzung, daß ein metrischer Tensor vorhanden ist, der Satz von Fuchs, wonach für $m = 2q$ das q -te assoziierte System des Differentialsystems (1) *reduzibel* ist²⁰⁾. Der Beweis, den z. B. A. Loewy (a. a. O. 17)) für diesen Satz gegeben hat, läßt sich ohne weiteres auf die hier vorliegenden Verhältnisse übertragen; wir bemerken nur, daß man sowohl bei der Definition der Reduzibilität eines Differentialsystems als überhaupt bei allen hier in Frage kommenden formalen Betrachtungen den Gebrauch der Lösungen des Differentialsystems vermeiden kann, ähnlich wie wir es ja auch bei der Definition des Artbegriffs durch die Gleichung (42) getan haben. In diesem Sinne sind die Formulierungen von A. Loewy²¹⁾, die gerade in der angedeuteten Richtung liegen, für unsere Theorie von besonderer Wichtigkeit. — Die Anwendung des Fuchsschen Satzes auf den

¹⁹⁾ Vgl. für diese Formeln etwa P. Appell, loc. cit. 10), S. 53 ff.; T. Levi-Civita, *Calcolo differenziale assoluto*, Roma 1925, S. 128 ff.; es ist nur zu bemerken, daß in unserem Text $m \neq n$ und $\Gamma_{lr}^k \neq \Gamma_{rl}^k$ sein kann. Für $m = n$ gibt es zu einem metrischen Tensor, d. h. zu einem symmetrischen G , stets ein System Christoffelscher Symbole Γ_{lr}^k , die die Symmetriebedingung erfüllen, aber im allgemeinen noch andere Systeme von Γ_{lr}^k , ohne Symmetrie.

²⁰⁾ Siehe die unter 17) angeführten Stellen. — Wenn in einem einfach zusammenhängenden Gebiete von \mathfrak{M}_{2q} die Γ_{lr}^k beschränkt sind und die Integrabilitätsbedingungen erfüllen, d. h. also \mathfrak{M}_{2q} daselbst eben ist, so sind die u_i^k daselbst auch Funktionen der t^1, \dots, t^n vom Charakter der Γ_{lr}^k , es existiert also sicher ein metrischer Tensor und die q -te Assoziierte ist reduzibel (vgl. L. Fuchs (1888), Werke III, S. 28); in diesem Falle ist aber alles trivial.

²¹⁾ A. Loewy, *Math. Annalen* 78 (1918), S. 1, 343, 359 ff.

physikalisch bedeutsamen Fall $m = n = 4$ lehrt, daß bei Vorhandensein eines metrischen Tensors das zweite assoziierte System des Differentialsystems der Parallelverschiebung (das also für die 6 Subdeterminanten zweiter Ordnung aus der Matrix zweier Vektorkomponenten gilt) reduzibel ist. Für $m = n = 2$, den Fall der Gaußschen Flächentheorie, ist das Differentialsystem der Parallelverschiebung selbst reduzibel; wir wollen jetzt dazu übergehen, unsere Betrachtungen auf diesen Fall zu spezialisieren.

VI.

Levi-Civita hat schon in seiner ersten, auf die Parallelverschiebung bezüglichen Arbeit hervorgehoben²³⁾, daß man, wenn ein metrischer Tensor existiert, durch Übergang zu einem Differentialsystem derselben Art, d. h. also durch eine Koordinatentransformation in \mathfrak{E}_m erreichen kann, daß das Differentialsystem der Parallelverschiebung mit seinem adjungierten zusammenfällt, so daß also $G = I$ wird und die quadratische Form $\sum_{i,k} g_{ik} u^i u^k$ sich auf eine Summe von Quadraten reduziert. — Man hat, um dies zu erreichen, in der Transformation (39 a) die Matrix H nur so zu bestimmen, daß das Differentialsystem (44) mit seinem adjungierten (53) identisch wird, und dies ergibt $G = H\bar{H}$, es handelt sich also um die rein algebraische Aufgabe, die symmetrische Matrix G als Produkt einer Matrix und ihrer Transponierten darzustellen. — Für $m = n = 2$ soll also H so bestimmt werden, daß für

$$v^1 = u^1 h_{11} + u^2 h_{21},$$

$$v^2 = u^1 h_{12} + u^2 h_{22}$$

die quadratische Form $\sum_{i,k} g_{ik} u^i u^k = (v^1)^2 + (v^2)^2$ wird. Durch Zerlegung von $\sum_{i,k} g_{ik} u^i u^k$ in lineare Faktoren findet man in unserem Falle, wenn wir etwa $g_{11} \neq 0$ voraussetzen,

$$(64) \quad \begin{cases} v^1 = \sqrt{g_{11}} u^1 + \frac{g_{12}}{\sqrt{g_{11}}} u^2, \\ v^2 = \sqrt{\frac{g}{g_{11}}} u^2, \end{cases} \quad g = g_{11} g_{22} - g_{12}^2$$

und die Umformung des Systems (1) in die abhängigen Veränderlichen v^1, v^2 ergibt mit Zuhilfenahme der Formeln (59 b, c) und (63)

$$(65) \quad \begin{cases} dv^1 = v^2 \frac{\sqrt{g}}{g_{11}} (\Gamma_{11}^2 dt^1 + \Gamma_{12}^2 dt^2), \\ dv^2 = -v^1 \frac{\sqrt{g}}{g_{11}} (\Gamma_{11}^2 dt^1 + \Gamma_{12}^2 dt^2). \end{cases}$$

²³⁾ T. Levi-Civita, loc. cit. ¹⁾, S. 187.

Setzt man nun noch

$$(66) \quad v^1 + \sqrt{-1} v^2 = w^1, \quad v^2 = w^2, \quad \text{also} \quad v^1 = w^1 - \sqrt{-1} w^2,$$

so folgt aus (65) für w^1, w^2 das *evident reduzible* Differentialsystem

$$(67) \quad \begin{cases} dw^1 + \frac{\sqrt{-g}}{g_{11}} (\Gamma_{11}^2 dt^1 + \Gamma_{12}^2 dt^2) w^1 & = 0, \\ dw^2 + (w^1 - \sqrt{-1} w^2) \frac{\sqrt{g}}{g_{11}} (\Gamma_{11}^2 dt^1 + \Gamma_{12}^2 dt^2) & = 0, \end{cases}$$

das also durch die Transformation

$$(68) \quad \begin{cases} w^1 = u^1 \sqrt{g_{11}} + u^2 \frac{g_{12} + \sqrt{-g}}{\sqrt{g_{11}}}, \\ w^2 = & u^2 \sqrt{\frac{g}{g_{11}}} \end{cases}$$

aus dem ursprünglichen für u^1, u^2 hervorgeht. Die Koeffizientenmatrix des Systems (67) lautet

$$B_1 dt^1 + B_2 dt^2 = -\frac{\sqrt{-g}}{g_{11}} (\Gamma_{11}^2 dt^1 + \Gamma_{12}^2 dt^2) \begin{pmatrix} 1 & -\sqrt{-1} \\ 0 & -1 \end{pmatrix},$$

wir können also (67) in Matrizenform so schreiben:

$$(67a) \quad W^{-1} dW = -\frac{\sqrt{-g}}{g_{11}} (\Gamma_{11}^2 dt^1 + \Gamma_{12}^2 dt^2) \begin{pmatrix} 1 & -\sqrt{-1} \\ 0 & -1 \end{pmatrix},$$

und hiernach ist evident, daß die Integralmatrix dieses Differentialsystems durch ²³⁾

$$W = e^{-\int \frac{\sqrt{-g}}{g_{11}} (\Gamma_{11}^2 dt^1 + \Gamma_{12}^2 dt^2) \cdot \begin{pmatrix} 1 & -\sqrt{-1} \\ 0 & -1 \end{pmatrix}}$$

gegeben wird, wo die Exponentialfunktion einer Matrix in üblicher Weise zu verstehen ist. Wir finden also für die Änderung einer Vektorbasis längs einer Führungskurve \mathfrak{C}

$$\int_{\mathfrak{C}} (B_1 dt^1 + B_2 dt^2 + I) = e^{-\int_{\mathfrak{C}} \frac{\sqrt{-g}}{g_{11}} (\Gamma_{11}^2 dt^1 + \Gamma_{12}^2 dt^2) \cdot \begin{pmatrix} 1 & -\sqrt{-1} \\ 0 & -1 \end{pmatrix}}.$$

Der Krümmungstensor $(R_{12})_B$ für das System (67) erhält, da die Matrizen B_1, B_2 offenbar vertauschbar sind, die einfache Form

$$(R_{12})_B = \frac{\partial B_1}{\partial t^2} - \frac{\partial B_2}{\partial t^1} = \left\{ \frac{\partial}{\partial t^2} \left(\frac{\sqrt{g}}{g_{11}} \Gamma_{11}^2 \right) - \frac{\partial}{\partial t^1} \left(\frac{\sqrt{g}}{g_{11}} \Gamma_{12}^2 \right) \right\} \cdot \begin{pmatrix} -\sqrt{-1} & -1 \\ 0 & \sqrt{-1} \end{pmatrix},$$

also ²⁴⁾

$$(R_{12})_B = \frac{\sqrt{g}}{g_{11}} R_{112}^2 \begin{pmatrix} -\sqrt{-1} & -1 \\ 0 & \sqrt{-1} \end{pmatrix} = -\sqrt{g} K \begin{pmatrix} -\sqrt{-1} & -1 \\ 0 & \sqrt{-1} \end{pmatrix},$$

²³⁾ L. Schlesinger, Differentialgleichungen, 1922, S. 155.

²⁴⁾ Vgl. etwa die Formeln bei L. Bianchi-Lukat, Differentialgeometrie, Leipzig 1910, §§ 28, 29, S. 48 ff.

wo K das Gaußsche Krümmungsmaß, das zum Linienelement $\sum_{i,k} g_{ik} dt^i dt^k$ gehört, bedeutet. Da nun offenbar die Matrix $(R_{12})_B$ auch mit der Matrix T (siehe Gleichung (36)) vertauschbar ist, so schreibt sich in dem vorliegenden Falle die dem Stokesschen Satze entsprechende Gleichung (34 a) in der einfachen Form

$$(69) \quad e^{-\int_{\mathfrak{R}} \frac{\sqrt{-g}}{g_{11}} (\Gamma_{11}^2 dt^1 + \Gamma_{12}^2 dt^2)} \cdot \begin{pmatrix} 1 & -\sqrt{-1} \\ 0 & -1 \end{pmatrix} = e^{\sqrt{-1} \iint_{(\mathfrak{R})} K \sqrt{g} dt^1 dt^2} \cdot \begin{pmatrix} 1 & -\sqrt{-1} \\ 0 & -1 \end{pmatrix},$$

oder wenn wir logarithmieren und nur den skalaren Faktor beibehalten

$$(69 a) \quad \int_{\mathfrak{R}} \frac{\sqrt{g}}{g_{11}} (\Gamma_{11}^2 dt^1 + \Gamma_{12}^2 dt^2) = - \iint_{(\mathfrak{R})} K \sqrt{g} dt^1 dt^2.$$

Setzen wir nun noch

$$w^1 = e^{\theta \sqrt{-1}},$$

so schreibt sich die Gleichung (69 a) mit Rücksicht auf die erste der Differentialgleichungen (67)

$$(69 b) \quad \int_{\mathfrak{R}} d\theta = \iint_{(\mathfrak{R})} K \sqrt{g} dt^1 dt^2,$$

und dies ist nichts anderes als die *Formel von Gauß-Ossian Bonnet*. Unsere Formel (34 a) ist also die genuine Verallgemeinerung der Gauß-Bonnetschen Formel oder besser, der dieser äquivalenten Formel (69 a).

Gießen, Mathem. Seminar, den 17. April 1927.

(Eingegangen am 28. 4. 1927.)

Über eine neue geometrische Eigenschaft der Bianchischen Transformation der auf die Mittelpunktsflächen zweiten Grades abwickelbaren Flächen.

Von

Hans Jonas in Berlin-Steglitz.

Inhaltsübersicht.

	Seite
Einleitung	435
§ 1. Bestimmung einer allgemeinen Flächenklasse auf Grund einer Biegungseigenschaft	438
§ 2. Biegungsflächen des einschaligen Hyperboloids. Satz über die Asymptotenlinien. Die drei Achsenspurflächen	447
§ 3. Aufstellung der Bianchischen Transformation B_k	454
§ 4. Formeln für die transformierte Biegungsfläche. Inverse Transformation. Korrespondenz der Asymptotenlinien	461
§ 5. Simultane asymptotische Transformationen der Achsenspurflächen	469

Einleitung.

Aus der Reihe geometrischer Tatsachen, die sich an die Verbiegung der Flächen zweiten Grades und an die Bianchische Transformation B_k knüpfen, seien hier diejenigen hervorgehoben, bei denen es sich um *Hilfsflächen* handelt, die einer verbogenen F^2 mit Korrespondenz der Asymptotenlinien zugeordnet sind und bei Anwendung der Operation B_k ebenso wie die Biegungsfläche selber asymptotische Transformationen erfahren. Die letzteren sind, wie man weiß, dadurch gekennzeichnet, daß die gegebene und die transformierte Fläche von den Strahlen einer W -Kongruenz in entsprechenden Punkten berührt werden.

An erster Stelle stehen die eigenartigen und durch Mannigfaltigkeit der geometrischen Zusammenhänge ausgezeichneten Beziehungen zu den Flächen von konstanter Krümmung und ihren Transformationen, die für die Biegungsflächen der Rotationsflächen zweiten Grades und ihrer Verallgemeinerungen, der sogenannten Darboux'schen F^2 , sowie für die ver-

bogenen Paraboloiden mit Ausschluß gewisser Grenzfälle von einfacherer Natur gelten¹⁾. Es sei ferner daran erinnert, daß die auf die Flächen zweiten Grades abwickelbaren Flächen zur Klasse der R -Flächen gehören, auf denen ein Netz konjugierter Kurven, im gedachten Sonderfalle das permanente konjugierte System, existiert, dessen Tangenten zwei W -Kongruenzen bilden. Zwischen der verbogenen F^2 und den zweiten Brennflächenmänteln dieser Kongruenzen, wir können auch sagen: den beiden Laplaceschen Transformierten des permanenten konjugierten Systems, besteht also Zuordnung ihrer Asymptotenlinien. Überdies gilt der Satz: Wird auf die Biegungsfläche eine Bianchische Transformation B_k ausgeübt, so gehen die Laplaceschen Transformierten durch asymptotische Transformationen in das auf gleiche Weise mit der neuen Biegungsfläche verbundene Flächenpaar über²⁾.

Als weiteres Beispiel erwähne ich Ergebnisse einer eigenen unlängst veröffentlichten Untersuchung³⁾ über die Transformation der auf das gleichseitige hyperbolische Paraboloid abwickelbaren Flächen. Der Sachverhalt ist dort insofern neuartig, als die an die Spitze der Theorie gestellten simultanen asymptotischen Transformationen der beiden zur Biegungsfläche des Paraboloids $z = xy$ gehörigen Hauptevolventenflächen nicht mit einer einfachen Transformation B_k zusammenfallen, sondern einem Kompositum von der Form $B_k B_{-k}$ entsprechen, das übrigens, wie beiläufig bemerkt sei, von der Guichardschen Transformation verschieden ist. Daneben tritt noch eine *dritte Hilfsfläche*⁴⁾ auf, deren Punkte und Normalen ebenso wie die der Hauptevolventenflächen eine von den Biegungen unabhängige Lage in den Tangentialebenen der Biegungsfläche des Paraboloids besitzen, die aber im Gegensatz zu jenen bei Anwendung einer *einfachen* Transformation B_k eine asymptotische Transformation erfährt, die sie mit der entsprechenden Hilfsfläche der neuen Biegungsfläche verbindet. Wir stellen hier noch ergänzend fest, daß *bei der Abwicklung der Biegungsfläche auf*

¹⁾ Betreffs neuer Ergebnisse sei verwiesen auf: Jonas, Über den wahren geometrischen Zusammenhang zwischen der Bianchischen Transformation der auf die Paraboloiden abwickelbaren Flächen und der Bäcklund'schen Transformation der Flächen von konstanter Krümmung. *Math. Annalen* 97 (1927), S. 387.

²⁾ Bianchi, Sui sistemi coniugati permanenti nelle deformate delle quadriche. *Acc. dei Linc. Rend.* 222 (1913), p. 3. Bezüglich des entsprechenden allgemeinen Satzes aus der Theorie der R -Flächen vgl. Jonas, Über die Konstruktion der W -Kongruenzen zu einem gegebenen Brennflächenmantel und über die Transformation der R -Flächen. *Jahresber. d. Deutsch. Math.-Vereing.* 29 (1920), S. 40 (insbesondere § 6, 1).

³⁾ Jonas, Ricerche sulle trasformazioni delle superficie applicabili sul paraboloido iperbolico equilatero. *Annali di Mat.* (4) 2 (1924–25), p. 161.

⁴⁾ A. a. O. S. 166 und 191; vgl. auch den Schluß von § 3 der in Fußnote ¹⁾ genannten Arbeit.

das Paraboloid oder, was auf dasselbe hinauskläuft, beim Rollen des Paraboloids auf der Biegungsfläche die einzelnen Punkte der dritten Hilfsfläche die Schnittpunkte der sie enthaltenden Tangentialebenen mit der Achse des Paraboloids werden.

Die soeben angegebene geometrische Konstruktion läßt sich nun auf den Fall der *dreiaxigen Mittelpunktsflächen zweiten Grades* ausdehnen. Den Biegungsflächen derselben werden damit je drei *Achsenpurflächen*, wie wir sie nennen wollen, angegliedert. Es soll im folgenden gezeigt werden, daß nach wie vor Korrespondenz der Asymptotenlinien zwischen Biegungsfläche und Achsenpurfläche besteht und daß das Verhalten der Achsenpurflächen der Bianchischen Transformation B_k gegenüber dasselbe ist wie in dem erwähnten Spezialfall.

Der erste Teil dieses Satzes wird in § 1 im Anschluß an ein allgemeineres Problem erledigt, bei dem es sich um die Verbiegung einer Strahlenkongruenz im Ribaucourschen Sinne und um die Bestimmung eines Typus des Linienelements auf Grund einer geforderten Biegungseigenschaft handelt. Der Beweis des zweiten Teils, demzufolge also bei Anwendung der Transformation B_k die Achsenpurflächen *simultane asymptotische Transformationen* erfahren, findet sich erst am Schluß der vorliegenden Abhandlung in § 5. Die voraufgeschickten, §§ 2—4 umfassenden Entwicklungen betreffen die Grundlagen und den Aufbau der von Bianchi geschaffenen Transformationstheorie. Mit der an sich naheliegenden Aufgabe, nun auch die prinzipiellen Gesichtspunkte der unter S. 436, Fußnote ³⁾, genannten Arbeit zu übertragen und unter besonderer Heranziehung der Achsenpurflächen die Transformation B_k der Moutard-Guichardschen Theorie der W -Kongruenzen unterzuordnen, werden wir uns dabei nicht beschäftigen; es sollen vielmehr unter Wahrung engen Anschlusses an die von Bianchi im dritten Bande seines Werkes⁵⁾ gegebene Darstellung der Transformationstheorie *Vereinfachungen in rechnerischer Hinsicht* gewonnen werden, die den Zugang dazu wesentlich erleichtern und deswegen auch ein selbständiges Interesse beanspruchen. Ein bemerkenswerter Vorteil wird dadurch erzielt, daß wir von vornherein mit der *quadrilinearen Relation* $\varphi(u, v, u_1, v_1) = 0$ operieren, die von den Parameterpaaren u, v und u_1, v_1 erfüllt wird, denen auf der gegebenen und der transformierten Fläche die Biegungslinien der Erzeugenden entsprechen. Wir sind so in der Lage, die Mehrzahl der benötigten Ausdrücke *formal* darzustellen, anstatt sie mit Bianchi *in extenso* schreiben zu müssen.

Geboten schien es, die Ausführungen auf das *einschalige Hyperboloid* zu beschränken, da die Diskussion der im Bereiche der Flächen zweiten

⁵⁾ Bianchi, *Lezioni di Geom. diff.* 3 (1909), insbes. Kap. 1—3.

Grades sich bietenden Fälle⁶⁾ und der damit verbundenen Realitätsfragen einen zu breiten Raum einnehmen würde. Unerörtert bleibt im Rahmen der gegenwärtigen Untersuchung alles, was mit der sukzessiven Anwendung der Transformation B_k und dem *Vertauschbarkeitssatz* zusammenhängt. Der kundige Leser erkennt übrigens als unmittelbare Folge dieses die Bianchische Theorie krönenden Theorems die Tatsache, daß im Verein mit den Transformationen B_k sich auch die simultanen asymptotischen Transformationen der drei Achsenspurflächen zu *viergliedrigen Zyklen* zusammenfügen.

§ 1.

Bestimmung einer allgemeinen Flächenklasse auf Grund einer Biegungseigenschaft.

1. Wir sprechen mit Bianchi⁷⁾ von der *Verbiegung einer Strahlenkongruenz*, wenn ihre Strahlen bei den isometrischen Deformationen einer Fläche von deren Flächenelementen in starrer Koppelung mitgeführt werden. Es gibt zwei Fälle, in denen die besondere Eigenschaft einer Kongruenz, ein Normalsystem darzustellen, d. h. ∞^1 Orthogonalflächen zuzulassen, bei beliebiger Biegung der Bezugsfläche erhalten bleibt: die Verbiegung der Kongruenz *im Beltramischen Sinne*, bei der die Strahlen durch die ihnen zugeordneten Punkte der Fläche gehen, und die Verbiegung *im Ribaucourschen Sinne*, bei der sie in den Tangentialebenen liegen. An die erste Operation knüpfen sich die von Guichard entdeckten Eigenschaften der Rotationsflächen zweiten Grades, an die zweite, als eine Art Gegenstück dazu, die Beziehungen zwischen den Biegungsflächen der Darboux'schen F^2 und den Flächen von konstanter Krümmung.

Die Fragestellung, von der wir ausgehen wollen, bezieht sich auf die Verbiegung der Kongruenzen im Ribaucourschen Sinne, bedeutet aber gleichzeitig auch einen Sonderfall eines anderen allgemeinen Problems, hinsichtlich dessen ich auf eine Mitteilung in den *Comptes Rendus* von 1913⁸⁾ verweise. Es handelt sich dabei um die Bestimmung der Paare von Flächen, die einander mit Orthogonalität der Normalen und Korrespondenz der Asymptotenlinien zugeordnet sind. Wir fragen hier nach der *Form*

⁶⁾ Dazu gehören auch die zahlreichen, z. T. interessanten Typen reeller ds^2 , die imaginären F^2 oder idealen Teilen reeller F^2 entsprechen. Erwähnt sei z. B. der besonders in den Anfängen der Theorie bedeutsame Fall, wo die acht isothermen Flächen, deren Punkte in den Tangentialebenen der Biegungsfläche durch die isotropen Erzeugenden der rollenden F^2 bestimmt werden, sämtlich oder teilweise reell sind.

⁷⁾ *Lezioni* 2 (1903), § 253.

⁸⁾ Jonas, *Sur une transformation qui dépend d'une équation aux dérivées partielles du 3^{me} ordre*. C. R. de l'Ac. des Sc. 156 (1913), p. 1816.

des Linienelements einer beliebigen Biegungen zu unterwerfenden Fläche S , in deren mit den Flächenelementen gekoppelten Tangentialebenen sich die Punkte und die Normalen einer (bei den Biegungen veränderlichen) Fläche S' derart festlegen lassen, daß die jeweiligen Asymptotenlinien von S und S' sich stets entsprechen.

Sehen wir von dem bei früherer Gelegenheit⁹⁾ berücksichtigten Fall ab, daß S' eine durch ein System geodätischer Linien definierte Evolventenfläche von S wird — das Linienelement entspricht dann einem geraden Konoid —, so können wir die Parameterkurven auf S so wählen, daß die Punkte von S' auf den Tangenten der Kurven $u = \text{konst.}$ liegen und die Normalen von S' parallel zu den Tangenten der Kurven $v = \text{konst.}$ werden. Es sei unter dieser Voraussetzung

$$\sum dx^2 = E du^2 + 2F du dv + G dv^2$$

das Quadrat des Linienelements der Fläche $S(x, y, z)$, und

$$- \sum dx dX = L du^2 + 2M du dv + N dv^2$$

ihre für die Gestalt im Raume charakteristische zweite Fundamentalform, wobei X, Y, Z die Richtungskosinus der Normalen bedeuten. Für die laufenden Koordinaten x', y', z' von S' gilt der folgende Ansatz:

$$(1) \quad x' = x - \lambda x_v^{10)},$$

und für die Normalenkosinus X', Y', Z' :

$$X' = \frac{1}{\sqrt{E}} x_u.$$

Wir differenzieren (1) mit Benutzung der Formeln der allgemeinen Flächentheorie, die die zweiten Ableitungen von x linear durch x_u, x_v, X ausdrücken, beachten, daß

$$\sum x'_u x_u = 0, \quad \sum x'_v x_u = 0$$

sein soll, und finden so zunächst die beiden Bedingungsgleichungen:

$$(2) \quad \begin{cases} (1 - \lambda \begin{Bmatrix} 12 \\ 1 \end{Bmatrix}) E - (\lambda_u + \lambda \begin{Bmatrix} 12 \\ 2 \end{Bmatrix}) F = 0, \\ \lambda \begin{Bmatrix} 22 \\ 1 \end{Bmatrix} E + (\lambda_v + \lambda \begin{Bmatrix} 22 \\ 2 \end{Bmatrix} - 1) F = 0, \end{cases}$$

die dem Ribaucourschen Satz entsprechend L, M, N nicht enthalten.

⁹⁾ Jonas, Untersuchungen über die als Gewebe bezeichneten Kurvennetze usw. Math. Annalen 87 (1922), S. 157; siehe daselbst § 2.

¹⁰⁾ Der Hinweis auf das Bestehen analoger Formeln bezüglich der y - und der z -Achse soll der Kürze wegen überall unterdrückt werden. Buchstabenindizes deuten partielle Differentiationen an.

Die Annahme $F = 0$ führt auf Bekanntes. Es wird dann nämlich:

$$\left\{ \begin{matrix} 22 \\ 1 \end{matrix} \right\} = 0, \quad \lambda = \frac{1}{\left\{ \begin{matrix} 12 \\ 1 \end{matrix} \right\}} = \frac{2E}{E_v}.$$

Die Kurven $u = \text{konst.}$ sind hiernach geodätische Linien; S' ist die zugehörige Ergänzungsfläche, d. h. der zweite Brennflächenmantel des von ihren Tangenten gebildeten Normalensystems. Dieses muß mit Rücksicht auf die außerdem noch geforderte Korrespondenz der Asymptotenlinien ein W -Normalensystem sein. Einem Satz von Weingarten zufolge ist S also Biegungsfläche einer Rotationsfläche, auf der die Kurven $u = \text{konst.}$ die Meridiane vorstellen.

Für $F \neq 0$ erhalten wir unter Anwendung der einsteilen als erfüllt vorausgesetzten Beziehungen (2) die Formeln:

$$(3) \quad \begin{cases} x'_u = \frac{1}{F} \left(1 - \lambda \left\{ \begin{matrix} 12 \\ 1 \end{matrix} \right\} \right) (F x_u - E x_v) - \lambda M X, \\ x'_v = -\frac{1}{F} \lambda \left\{ \begin{matrix} 22 \\ 1 \end{matrix} \right\} (F x_u - E x_v) - \lambda N X, \end{cases}$$

deren wir uns zur Berechnung der zweiten Fundamentalgrößen von S' :

$$L' = -\frac{1}{\sqrt{E}} \sum x'_u x_{uu}, \quad M' = -\frac{1}{\sqrt{E}} \sum x'_u x_{uv}, \quad N' = -\frac{1}{\sqrt{E}} \sum x'_v x_{vv},$$

bedienen. Es ergibt sich so:

$$(4) \quad \begin{cases} \sqrt{E} L' = \frac{1}{F} (EG - F^2) \left\{ \begin{matrix} 11 \\ 2 \end{matrix} \right\} \left(1 - \lambda \left\{ \begin{matrix} 12 \\ 1 \end{matrix} \right\} \right) + \lambda M L, \\ \sqrt{E} M' = \frac{1}{F} (EG - F^2) \left\{ \begin{matrix} 12 \\ 2 \end{matrix} \right\} \left(1 - \lambda \left\{ \begin{matrix} 12 \\ 1 \end{matrix} \right\} \right) + \lambda M^2, \\ \sqrt{E} N' = -\frac{1}{F} (FG - F^2) \left\{ \begin{matrix} 12 \\ 2 \end{matrix} \right\} \left\{ \begin{matrix} 22 \\ 1 \end{matrix} \right\} \lambda + \lambda M N. \end{cases}$$

Damit nun die Asymptotenlinien der Flächen S und S' (und gleichzeitig auch ihre konjugierten Systeme) einander entsprechen, muß

$$L' : M' : N' = L : M : N$$

sein. Dieser Bedingung kann aber, wie eine von analogen Fällen her geäußerte Überlegung zeigt, allgemein nur in der Weise genügt werden, daß

$$\left\{ \begin{matrix} 11 \\ 2 \end{matrix} \right\} \left(1 - \lambda \left\{ \begin{matrix} 12 \\ 1 \end{matrix} \right\} \right) = 0, \quad \left\{ \begin{matrix} 12 \\ 2 \end{matrix} \right\} \left(1 - \lambda \left\{ \begin{matrix} 12 \\ 1 \end{matrix} \right\} \right) = 0, \quad \left\{ \begin{matrix} 12 \\ 2 \end{matrix} \right\} \left\{ \begin{matrix} 22 \\ 1 \end{matrix} \right\} = 0,$$

wird. Von den sich bietenden Möglichkeiten ist

$$1 - \lambda \left\{ \begin{matrix} 12 \\ 1 \end{matrix} \right\} = 0, \quad \left\{ \begin{matrix} 22 \\ 1 \end{matrix} \right\} = 0$$

im Hinblick auf (3) zu verwerfen. Ebenso wenig kommt

$$\left\{ \begin{matrix} 11 \\ 2 \end{matrix} \right\} = 0, \quad \left\{ \begin{matrix} 12 \\ 2 \end{matrix} \right\} = 0$$

in Betracht, da alsdann auf Grund der Formel

$$KE = \left\{ \begin{matrix} 11 \\ 2 \end{matrix} \right\}_v - \left\{ \begin{matrix} 12 \\ 2 \end{matrix} \right\}_u + \left\{ \begin{matrix} 11 \\ 1 \end{matrix} \right\} \left\{ \begin{matrix} 12 \\ 2 \end{matrix} \right\} + \left\{ \begin{matrix} 11 \\ 2 \end{matrix} \right\} \left\{ \begin{matrix} 22 \\ 2 \end{matrix} \right\} - \left\{ \begin{matrix} 12 \\ 1 \end{matrix} \right\} \left\{ \begin{matrix} 11 \\ 2 \end{matrix} \right\} - \left\{ \begin{matrix} 12 \\ 2 \end{matrix} \right\}^2$$

entweder E oder das Krümmungsmaß K identisch verschwinden müßte. Es bleibt die Annahme

$$(5) \quad 1 - \lambda \left\{ \begin{matrix} 12 \\ 1 \end{matrix} \right\} = 0, \quad \left\{ \begin{matrix} 12 \\ 2 \end{matrix} \right\} = 0$$

übrig, die im Verein mit der ersten Relation (2) $\lambda_u = 0$, also $\lambda = \lambda(v)$ liefert. Die Wahl dieser Funktion $\lambda(v)$ ist aber, wie ein Blick auf (1) lehrt, völlig belanglos, so daß man, ohne die Allgemeinheit zu beschränken, $\lambda = 1$ setzen könnte. Als zweckmäßiger erweist es sich,

$$(6) \quad \lambda = v$$

zu wählen. Der zweiten Relation (2) entnimmt man dann:

$$E \left\{ \begin{matrix} 22 \\ 1 \end{matrix} \right\} + F \left\{ \begin{matrix} 22 \\ 2 \end{matrix} \right\} = 0,$$

und findet an Hand der Formel

$$F_v = E \left\{ \begin{matrix} 22 \\ 1 \end{matrix} \right\} + F \left(\left\{ \begin{matrix} 22 \\ 2 \end{matrix} \right\} + \left\{ \begin{matrix} 12 \\ 1 \end{matrix} \right\} \right) + G \left\{ \begin{matrix} 12 \\ 2 \end{matrix} \right\}$$

mit Benutzung von (5) und (6):

$$\frac{F_v}{F} = \frac{1}{v}, \quad \text{d. i. } F = vU',$$

wobei U' eine mit Rücksicht auf das Folgende als Ableitung geschriebene Funktion von u bedeutet. Mittels des für F gewonnenen Ausdrucks muß schließlich noch den Gleichungen

$$\left\{ \begin{matrix} 12 \\ 1 \end{matrix} \right\} = \frac{1}{v}, \quad \left\{ \begin{matrix} 12 \\ 2 \end{matrix} \right\} = 0$$

genügt werden. Diese lauten ausführlich:

$$\frac{GE_v - FG_u}{2(EG - F^2)} = \frac{1}{v}, \quad EG_u - FE_v = 0.$$

Wird

$$(7) \quad G_u = \frac{F}{E} E_v$$

in die erste eingesetzt, so folgt

$$\frac{E_v}{E} = 2,$$

also mit Unterdrückung eines offenbar entbehrlichen, von u allein abhängigen Faktors:

$$E = v^2.$$

Aus (7) erhält man:

$$G = 2(U + V),$$

wobei V Funktion von v ist. Das Quadrat des Linienelements der Fläche S hat demnach die charakteristische Form:

$$(8) \quad ds^2 = v^2 du^2 + 2vU' dudv + 2(U+V)dv^2.$$

Die zugehörige Fläche S' und ihre Normale sind durch

$$(9) \quad x' = x - vx_v, \quad X' = \frac{1}{v} x_u$$

gegeben. Der Ausdruck (8) umfaßt für $U' = 0$ auch den Rotationsflächentypus, der bei der vorstehenden Rechnung nur mit Rücksicht auf die Schreibweise der Formeln (3) ausgeschlossen wurde.

2. Das Ergebnis ist einer geometrischen Deutung fähig, durch die es den bekannten Sätzen über die Verbiegung der Rotationsflächen noch näher gerückt wird. Wir betrachten nämlich die Fläche

$$(10) \quad \xi = U_1 v, \quad \eta = U_2 v, \quad \zeta = V_0;$$

U_1, U_2 seien dabei Funktionen von u , V_0 sei Funktion von v . Im besonderen liegt für $U_1^2 + U_2^2 = \text{konst.}$ eine Rotationsfläche vor. Da allgemein über u so verfügt werden kann, daß

$$U_1'^2 + U_2'^2 = 1$$

wird, findet man als Quadrat des Linienelements der durch (10) definierten Fläche:

$$ds^2 = v^2 du^2 + 2v(U_1 U_1' + U_2 U_2') dudv + (U_1^2 + U_2^2 + V_0'^2) dv^2.$$

Dieser Ausdruck geht, wenn $U_1^2 + U_2^2 = 2U$, $V_0'^2 = 2V$ gesetzt wird, in (8) über.

Ist umgekehrt ein in dem Typus (8) enthaltenes ds^2 gegeben, wobei u, v reelle Parameter sein sollen, so ist die Existenz dazugehöriger Biegungsflächen S an die Bedingung

$$2U - U'^2 + 2V > 0$$

gebunden, die wir mithin für den zu betrachtenden Wertebereich von u, v als erfüllt voraussetzen. Dann läßt sich aber, wie ersichtlich, noch auf ∞^1 Weisen eine positive oder negative Konstante c so wählen, daß in einem Teilbereich nebeneinander die Ungleichheiten

$$2U - U'^2 + c > 0, \quad 2V - c > 0$$

gelten. Man kann also U_1, U_2, V_0 aus den Beziehungen

$$U_1^2 + U_2^2 = 2U + c, \quad U_1'^2 + U_2'^2 = 1, \quad V_0'^2 = 2V - c$$

bestimmen und erhält dadurch zu dem gegebenen ds^2 eine Schar von ∞^1 Biegungsflächen der Klasse (10). Es wird nämlich:

$$(11) \begin{cases} U_1 = \sqrt{2U+c} \cdot \cos \omega, & U_2 = \sqrt{2U+c} \cdot \sin \omega, & \omega = \int \frac{\sqrt{2U-U'^2+c}}{2U+c} du, \\ & & V_0 = \int \sqrt{2V-c} dv. \end{cases}$$

Für diese speziellen Biegungsflächen artet die Hilfsfläche S' aus; man findet mittels (9):

$$(12) \quad \begin{cases} \xi' = \eta' = 0, & \zeta' = V_0 - \frac{1}{v} V'_0, \\ \mathfrak{X}' = U'_1, & \mathfrak{Y}' = U'_2, & \mathfrak{Z}' = 0. \end{cases}$$

Das leicht zu deutende Resultat sprechen wir in dem folgenden Satze aus: *Eine Fläche, deren Schnitte mit den Ebenen $z = \text{konst.}$ (Höhenlinien) sich auf die xy -Ebene als ein System ähnlicher und bezüglich des Koordinatenanfangs ähnlich gelegener Kurven projizieren, besitzt die Eigenschaft, daß die in den Tangentialebenen markierten Spurrpunkte der z -Achse nach einer beliebigen Verbiegung der Fläche stets eine Fläche S' bilden, deren (reelle oder imaginäre) Asymptotenlinien denjenigen der Biegungsfläche S entsprechen. Die Normalen von S' gehen bei der Biegung aus den Schnittlinien der Tangentialebenen mit den Ebenen $z = \text{konst.}$ hervor.*

Es soll noch der folgende Zusatz bewiesen werden: *Das Krümmungsmaß von S' bleibt beim Verbiegen der Fläche S ungeändert.*

Man findet nämlich mit Berücksichtigung von (8):

$$x'_u = -v M X, \quad x'_v = \frac{V'(U' x_u - v x_v)}{2U - U'^2 + 2V} - v N X,$$

so daß sich für die Fundamentalgrößen erster Ordnung von S' die Ausdrücke

$$E' = v^2 M^2, \quad F' = v^2 M N, \quad G' = \frac{v^2 V'^2}{2U - U'^2 + 2V} + v^2 N^2$$

ergeben. Die Formeln (4) gehen in die folgenden über:

$$L' = M L, \quad M' = M^2, \quad N' = M N.$$

Es wird demnach:

$$K' = \frac{L' N' - M'^2}{E' G' - F'^2} = \frac{K(2U - U'^2 + 2V)^2}{v^3 V'^2}$$

und, wenn das Krümmungsmaß K von S durch den aus den Koeffizienten der Differentialform (8) berechneten Wert

$$(13) \quad K = - \frac{V'(U'' - 1)}{v(2U - U'^2 + 2V)^2}$$

ersetzt wird:

$$(14) \quad K' = - \frac{U'' - 1}{v^3 V'}.$$

In der Tat hängt K' von L, M, N nicht ab.

3. Wir bemerken noch, daß das Linienelement der von uns betrachteten Flächen leicht auf diejenige Form gebracht werden kann, die den Ausgangspunkt für die Weingartensche Methode zur Behandlung des Biegungsproblems bildet. Werden \bar{u} und ψ durch die Formeln

$$(15) \quad \bar{u} = uv, \quad \psi = vU - \frac{1}{2}u^2v + \int Vdv$$

definiert, so geht (8) in

$$(16) \quad ds^2 = d\bar{u}^2 + 2d\psi dv$$

über.

An die geometrische Deutung im Reellen sei zunächst erinnert: die Reduktion des Linienelements einer beliebigen Fläche S auf die Form (16) liefert ein normales System von Kreisen in den Tangentialebenen von S . Eine der Orthogonalflächen, deren Krümmungslinien dem sog. zyklischen konjugierten System von S entsprechen, ist die von Darboux¹¹⁾ mit Σ' bezeichnete Hilfsfläche mit den laufenden Koordinaten:

$$x - \bar{u} \frac{\partial x}{\partial \bar{u}} - v \frac{\partial x}{\partial v}, \dots$$

Die Konstruktion, durch die Darboux sie einführt, knüpft sich an die ebenfalls durch (16) definierte imaginäre Biegungsfläche S_0 :

$$(17) \quad x_0 = \bar{u}, \quad y_0 + iz_0 = v, \quad y_0 - iz_0 = 2\psi.$$

Rollt nämlich S_0 auf der reellen isometrischen Fläche S , so ist Σ' Ort der Schnittpunkte der Tangentialebenen mit der Minimalgeraden $x = 0$, $y + iz = 0$, die von dem bewegten Koordinatensystem der rollenden S_0 mitgenommen wird. Für das von uns betrachtete Linienelement (8) wird, wie man leicht erkennt, die Darboux'sche Hilfsfläche Σ' mit der durch (9) gegebenen Fläche S' identisch¹²⁾.

4. Bevor wir uns dem eigentlichen Gegenstand unserer Untersuchung, den Biegungsflächen der allgemeinen Mittelpunktsflächen zweiten Grades, zuwenden, möge der Fall, daß S' *konstantes Krümmungsmaß* hat, kurz erörtert werden. Wir erhalten dabei die Typen des Linienelements, die Bianchi, an eine Note von Darboux¹³⁾ anknüpfend, in seiner Abhandlung

¹¹⁾ Darboux, Leçons sur la théorie gén. des surf. 4 (1896), p. 310.

¹²⁾ Die Frage nach dem Typus des Linienelements, für den allgemein zwischen der Fläche S und der Darboux'schen Hilfsfläche Σ' Korrespondenz der Asymptotenlinien besteht, fällt also, von der spezielleren Formulierung abgesehen, mit dem in Art. 1 erledigten Problem zusammen.

¹³⁾ Darboux, Sur la déformation des surfaces du second degré. C. R. de l'Ac. des Sc. 129 (1899), p. 760.

über die Verbiegung der Kongruenzen im Ribaucourschen Sinne¹⁴⁾ aufgestellt hat. Der Umstand, daß wir im Gegensatz zu Bianchi zunächst Korrespondenz der Asymptotenlinien zwischen der Biegungsfläche S und der Orthogonalfläche S' der Kongruenz verlangt haben, vereinfacht die Herleitung, die an der Hand der Formel (14) unmittelbar gelingt. Aus

$$K' = \frac{1}{A} = \text{konst.}$$

folgt nämlich, wenn γ , γ_1 , γ_2 und C Konstanten sind:

$$U = \frac{\gamma+1}{2} u^2 + \gamma_1 u + \gamma_2, \quad V = \frac{\gamma A}{2v^3} + \frac{1}{2} C.$$

Für $\gamma \neq -1$ findet man durch Einsetzen in (8) nach Unterdrückung der entbehrlichen Konstanten γ_1 , γ_2 :

$$(18) \quad ds^2 = v^2 du^2 + 2(\gamma+1)uv du dv + \left[(\gamma+1)u^2 + \frac{\gamma A}{v^2} + C \right] dv^2.$$

Da man, ohne die Allgemeinheit zu beschränken, für $C \neq 0$ die Substitution $u \left| u \sqrt{|C|} \right|$, $v \left| \frac{v}{\sqrt{|C|}} \right|$ anwenden kann, so genügt es, der Konstanten C die Werte $+1$, -1 , 0 zu geben. Als Vertreter des Typus (18) im Bereiche der Flächen zweiten Grades erhalten wir mit Hilfe von (17) die imaginäre Darboux'sche F^2 , deren Schnitt mit der unendlich-fernen Ebene den Kugelkreis berührt:

$$(19) \quad \begin{aligned} x_0 &= uv, & y_0 + iz_0 &= v, & y_0 - iz_0 &= \gamma u^2 v - \frac{\gamma A}{v} + Cv, \\ y_0^2 + z_0^2 &= \gamma x_0^2 + C(y_0 + iz_0)^2 - Av. \end{aligned}$$

Im besonderen Falle $C=0$ entspricht das vorliegende Linienelement einer Rotationsfläche¹⁵⁾. Die Reduktion auf die Form (18) kann dann, wie schon aus Gründen der Symmetrie zu ersehen ist, auf doppelte Weise erfolgen; mit der (als reell vorausgesetzten) Biegungsfläche S verbinden sich also zwei Flächen von konstanter Krümmung als Flächen S' . Wir bemerken, daß die Normalen derselben sich paarweise in den Punkten der Ergänzungsfläche treffen, die die Fläche S in ihrer Eigenschaft als verbogene Rotationsfläche besitzt¹⁶⁾.

¹⁴⁾ Bianchi, Sulla deformazione delle congruenze e sopra alcune classi di superficie applicabili. Annali di Mat. (3) 6 (1901), p. 117.

¹⁵⁾ Man hat zu beachten, daß reelles ds^2 an die Bedingung $\gamma[A - (\gamma+1)u^2v^2] > 0$ gebunden ist. Reelle Vertreter der 4 sich ergebenden Typen sind: für $A > 0$ die Komplementärflächen des verlängerten Rotationsellipsoids und des zweischaligen Rotationshyperboloids, für $A < 0$ das verkürzte Katenoid und die Komplementärfläche des hyperbolischen Sinusoids.

¹⁶⁾ Bianchi, Lezioni 2, § 309. Betreffs der Konstruktion dieser Gebilde durch Zusammensetzung reeller oder imaginärer Bäcklund'scher Transformationen vgl. daselbst §§ 390, 402—405.

Für $\gamma = -1$ können durch geeignete Wahl von u die Konstanten γ_2 und C unterdrückt werden; außerdem darf man, wie die Substitution $u | u \gamma_1, v | \frac{v}{\gamma_1}$ erkennen läßt, $\gamma_1 = 1$ setzen. Man findet so den Typus:

$$(20) \quad ds^2 = v^2 du^2 + 2v du dv + \left(2u - \frac{A}{v^2}\right) dv^2$$

und dazu die imaginäre Darboux'sche F^2 , die den Kugelkreis oskuliert:

$$(21) \quad \begin{aligned} x_0 &= uv, & y_0 + iz_0 &= v, & y_0 - iz_0 &= 2uv - u^2v + \frac{A}{v}, \\ x_0^2 + y_0^2 + z_0^2 - 2x_0(y_0 + iz_0) &= A. \end{aligned} \quad (17)$$

Wir weisen schließlich darauf hin, daß man die Formeln (10), (11) zur Bestimmung gewisser reeller Biegungsflächen der Darboux'schen Flächen zweiten Grades (19) und (21) benutzen kann.

5. Bezüglich der *dreiachsigen Mittelpunktsflächen zweiten Grades* — die weiteren Entwicklungen sollen auf den Fall des einschaligen Hyperboloids beschränkt werden — stellen wir hier zunächst fest, daß bei diesen die in dem Satz des Art. 2 vorausgesetzte Eigenschaft der zur Achse senkrechten ebenen Schnitte in *dreifacher* Weise verwirklicht ist. Die in den Tangentialebenen markierten Spurpunkte der drei Hauptachsen bilden nach einer Verbiegung der F^2 , bei der die Flächenelemente die Tangentialebenen in starrer Koppelung mitführen, drei zur Biegungsfläche S gehörige Flächen S' . Wir wollen sie als die drei *Achsenpurflächen* der verbogenen F^2 bezeichnen. Nach dem Voraufgehenden gilt der Satz: *Zwischen der Biegungsfläche S einer Mittelpunktsfläche zweiten Grades und ihren drei Achsenflächen besteht Korrespondenz der Asymptotenlinien und damit sämtlicher konjugierten Systeme; die Normalen der Achsenpurflächen sind parallel zu den gegenüberliegenden Seiten des Achsenpurdreiecks*¹⁸⁾.

¹⁷⁾ Bei Bianchi (ebenda § 308) hat die Gleichung der Fläche die Form:

$$y^2 + z^2 + (x - y + iz)^2 = A.$$

Man erhält sie aus: $x = \left(u + \frac{1}{2}\right)v$, $y - iz = v$, $y + iz = -\left(u - \frac{1}{2}\right)v + \frac{A}{v}$. Erwähnt sei noch die Erzeugung der Flächen vom Typus (20) durch die Lamé'schen Scharen von Flächen konstanter Krümmung (s. daselbst § 447).

¹⁸⁾ Ein bemerkenswertes Beispiel außerhalb der Biegungstheorie der Flächen zweiten Grades liefern die mit den Affinsphären zusammenhängenden Biegungsflächen des Typus

$$ds^2 = \frac{9}{4}(u du^2 - 2 du dv + v dv^2).$$

Sie lassen als Flächen S' im Sinne des Satzes von Art. 2 eine Schar von ∞^1 Flächen $S^{(v)}$ zu [Jonas, Aufstellung einer Transformationstheorie für eine neue Klasse aufeinander abwickelbarer Flächen. Math. Annalen 92 (1924), S. 214]. Bei der auf ∞^2 Weisen möglichen Verbiegung in Flächen der tetraedralen Klasse $Ax^{2/3} + By^{2/3} + Cz^{2/3} = 1$ gehen die Punkte je dreier dieser Flächen $S^{(v)}$ in die Schnittpunkte der Tangentialebenen mit den Koordinatenachsen über.

§ 2.

Biegungsflächen des einschaligen Hyperboloids. Satz über die Asymptotenlinien. Die drei Achsenspurflächen.

1. Das einschalige Hyperboloid

$$(22) \quad \frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$$

werde wie bei Bianchi¹⁹⁾ auf die Parameter u, v ²⁰⁾ der Erzeugenden bezogen. Es gilt dann die folgende Darstellung:

$$(23) \quad x = a \frac{1+uv}{u+v}, \quad y = b \frac{u-v}{u+v}, \quad z = c \frac{1-uv}{u+v}.$$

Für die Fundamentalgrößen erster Ordnung erhält man die Formeln:

$$(24) \quad \begin{cases} (u+v)^4 E = a^2(1-v^2)^2 + 4b^2v^2 + c^2(1+v^2)^2, \\ (u+v)^4 F = a^2(1-u^2)(1-v^2) - 4b^2u^2v^2 + c^2(1+u^2)(1+v^2), \\ (u+v)^4 G = a^2(1-u^2)^2 + 4b^2u^2 + c^2(1+u^2)^2 \end{cases}$$

und findet:

$$EG - F^2 = \frac{4a^2b^2c^2H}{(u+v)^6},$$

wobei

$$(25) \quad H = \frac{1}{a^2}(1+uv)^2 + \frac{1}{b^2}(u-v)^2 + \frac{1}{c^2}(1-uv)^2$$

ist. Angemerkt seien die Werte der Christoffelschen Symbole:

$$(26) \quad \begin{cases} \begin{Bmatrix} 11 \\ 1 \end{Bmatrix} = \begin{Bmatrix} 22 \\ 2 \end{Bmatrix} = -\frac{2}{u+v}, & \begin{Bmatrix} 11 \\ 2 \end{Bmatrix} = \begin{Bmatrix} 22 \\ 1 \end{Bmatrix} = 0, \\ \begin{Bmatrix} 12 \\ 1 \end{Bmatrix} = \frac{1}{2} \frac{H_v}{H} - \frac{1}{u+v}, & \begin{Bmatrix} 12 \\ 2 \end{Bmatrix} = \frac{1}{2} \frac{H_u}{H} - \frac{1}{u+v}. \end{cases}$$

\sqrt{H} werde stets positiv genommen; $\sqrt{EG - F^2}$ dagegen möge entsprechend der Formel

$$(27) \quad \sqrt{EG - F^2} = \frac{2abc\sqrt{H}}{(u+v)^3}$$

das Vorzeichen von $u+v$ führen. Mit den Ausdrücken

$$(28) \quad \mathfrak{X} = \frac{\eta_u \delta v - \delta u \eta_v}{\sqrt{EG - F^2}} \quad \text{usw.}$$

für die Richtungskosinus der Normalen sind dann die folgenden auch hinsichtlich des Vorzeichens gleichwertig:

$$(29) \quad \mathfrak{X} = -\frac{(u+v)x}{a^2\sqrt{H}}, \quad \mathfrak{Y} = -\frac{(u+v)y}{b^2\sqrt{H}}, \quad \mathfrak{Z} = \frac{(u+v)z}{c^2\sqrt{H}}.$$

¹⁹⁾ Bianchi, Lezioni 3, § 16.

²⁰⁾ Die Einführung der weiterhin beizubehaltenden Bezeichnungen erfolgt unabhängig von § 1.

Die zweite Fundamentalform wird:

$$(30) \quad - \sum dx_i d\mathfrak{X} = - \frac{4 du dv}{(u+v)\sqrt{H}};$$

ihre Koeffizienten sind also $\mathfrak{L} = \mathfrak{N} = 0$ und

$$(31) \quad \mathfrak{M} = - \frac{2}{(u+v)\sqrt{H}}.$$

Das Krümmungsmaß hat den Wert

$$K = - \frac{\mathfrak{M}^2}{EG - F^2} = - \frac{(u+v)^4}{a^2 b^2 c^2 H^2}.$$

2. Um eine Biegungsfläche $S(x, y, z)$ des Hyperboloids (23) zu definieren, kann man zwei Wege einschlagen. Der eine besteht in der Einführung der Fundamentalgrößen zweiter Ordnung L, M, N , die der Gaußschen Relation und den Codazzischen Differentialgleichungen genügen müssen. Letztere sind mit Benutzung von (26) zu bilden. Ihre Gestalt vereinfacht sich, wenn wir uns der drei Größen

$$(32) \quad l = \frac{L}{\mathfrak{M}H}, \quad m = \frac{M}{\mathfrak{M}}, \quad n = \frac{N}{\mathfrak{M}H}$$

bedienen. Die drei zu erfüllenden Beziehungen lauten dann:

$$(33) \quad \begin{cases} H^2 ln = m^2 - 1, \\ m_u = H l_v, \quad m_v = H n_u. \end{cases}$$

Bei der zweiten Methode werden u und v als Funktionen zweier Variablen α und β dargestellt bzw. aufgefaßt, denen auf der Biegungsfläche S die Asymptotenlinien entsprechen. Die beiden Differentialgleichungen zweiter Ordnung, denen die Funktionen $u(\alpha, \beta)$ und $v(\alpha, \beta)$ unterworfen sind, erhält man aus den Christoffelschen Invarianzen für die gemischten zweiten Ableitungen. Die Eigentümlichkeit dieses Gleichungspaares, zwei Integrale erster Ordnung zuzulassen, hängt eng mit der Tatsache zusammen, daß sich auf der verbogenen F^2 die Asymptotenlinien, also die Integralkurven der Differentialgleichung

$$(34) \quad L du^2 + 2 M du dv + N dv^2 = 0$$

durch Quadraturen bestimmen lassen. Es handelt sich dabei um eine aus dem System (33) zu ziehende Folgerung. Wiewohl wir mit Rücksicht auf die von Bianchi gegebene Darstellung seiner Transformationstheorie den erstgenannten Weg bevorzugen werden, benutzen wir die Gelegenheit, den bislang nicht hinreichend geklärten Zusammenhang beider Methoden zu erörtern. An die zweite knüpft sich übrigens eine Bemerkung am Schluß von § 4, die geeignet erscheint, einen neuen Einblick in die analytische Natur der Transformation B_k zu eröffnen.

Zum Beweise des *Satzes über die Asymptotenlinien der Biegungsfläche* geht man zweckmäßig von den beiden durch (34) gegebenen Werten von $\frac{dv}{du}$ aus, die den asymptotischen Richtungen entsprechen. Sie seien mit p und q bezeichnet, so daß die Differentialgleichungen der beiden Scharen von Asymptotenlinien jetzt

$$(35) \quad p \, du - dv = 0, \quad q \, du - dv = 0$$

lauten. Wir finden, indem wir die Beziehung

$$M^2 - LN = \mathfrak{M}^2$$

beachten:

$$(36) \quad p = -\frac{M + \mathfrak{M}}{N} = -\frac{L}{M - \mathfrak{M}}; \quad q = -\frac{M - \mathfrak{M}}{N} = -\frac{L}{M + \mathfrak{M}}$$

und umgekehrt:

$$(37) \quad L = -\frac{2 \mathfrak{M} p q}{p - q}, \quad M = \frac{\mathfrak{M}(p + q)}{p - q}, \quad N = -\frac{2 \mathfrak{M}}{p - q}.$$

Bildet man nach (32)

$$l = -\frac{2 p q}{H(p - q)}, \quad m = \frac{p + q}{p - q}, \quad n = -\frac{2}{H(p - q)}$$

und setzt diese Ausdrücke in die beiden Differentialgleichungen des Systems (33) ein, so erhält man:

$$p(q_u + p q_v) - q(p_u + q p_v) = p q (p - q) \frac{H_v}{H},$$

$$q_u + p q_v - (p_u + q p_v) = (p - q) \frac{H_u}{H}$$

und hieraus schließlich:

$$(38) \quad p_u + q p_v = \frac{p}{H}(q H_v - H_u), \quad q_u + p q_v = \frac{q}{H}(p H_v - H_u).$$

Zu den gefundenen Formeln, denen man auch die einfachere Gestalt

$$(39) \quad (p H)_u + q H^2 \left(\frac{p}{H}\right)_v = 0, \quad (q H)_u + p H^2 \left(\frac{q}{H}\right)_v = 0$$

geben kann, sei zunächst folgendes bemerkt. Werden p und q als gesuchte Funktionen aufgefaßt, so bringen die simultanen Differentialgleichungen (39), die (33) ersetzen, von neuem das Biegungsproblem für die allgemeine F^2 analytisch zum Ausdruck. Wichtig erscheint an denselben der Umstand, daß eine der beiden Unbekannten leicht eliminiert werden kann; für die andere ergibt sich dabei eine partielle Differentialgleichung zweiter Ordnung vom Ampèreschen Typus.

Die Relationen (38) ermöglichen nun den Nachweis, daß

$$(40) \quad d\alpha = \frac{\sqrt{|q|}(p \, du - dv)}{\sqrt{H}(p - q)}, \quad d\beta = \frac{\sqrt{|p|}(q \, du - dv)}{\sqrt{H}(p - q)}$$

exakte Differentiale sind. Um nämlich das Verschwinden des Ausdrucks

$$\frac{\sqrt{H}(p-q)}{\sqrt{q}} \left[\frac{\partial}{\partial v} \left(\frac{\sqrt{q} p}{\sqrt{H}(p-q)} \right) + \frac{\partial}{\partial u} \left(\frac{\sqrt{q}}{\sqrt{H}(p-q)} \right) \right]$$

$$= p_v + p \left(\frac{1}{2} \frac{q_v}{q} - \frac{1}{2} \frac{H_v}{H} - \frac{p_v - q_v}{p-q} \right) + \frac{1}{2} \frac{q_u}{q} - \frac{1}{2} \frac{H_u}{H} - \frac{p_u - q_u}{p-q}$$

(nachträglich können p und q vertauscht werden) zu bestätigen, braucht man nur p_u und q_u durch die den Gleichungen (38) entnommenen Werte zu ersetzen. Die Integrale $\alpha = \text{konst.}$ und $\beta = \text{konst.}$ der beiden Differentialgleichungen (35) gewinnt man also in der Tat durch zwei Quadraturen²¹⁾.

Von (40) gelangt man zu den *Differentialgleichungen der zweiten, die Parameter α , β der Asymptotenlinien benutzenden Methode*²²⁾. Wird

$$(41) \quad q = \varepsilon_1 |q|, \quad p = \varepsilon_2 |p|, \quad (|\varepsilon_1| = |\varepsilon_2| = 1)$$

gesetzt, so ergibt die Auflösung nach du und dv die Formeln:

$$(42) \quad du = \sqrt{H} \left(\frac{d\alpha}{\sqrt{|q|}} - \frac{d\beta}{\sqrt{|p|}} \right), \quad dv = \sqrt{H} (\varepsilon_1 \sqrt{|q|} d\alpha - \varepsilon_2 \sqrt{|p|} d\beta),$$

aus denen hervorgeht, daß u und v als Funktionen von α und β den simultanen Differentialgleichungen

$$(43) \quad u_\alpha v_\alpha = \varepsilon_1 H, \quad u_\beta v_\beta = \varepsilon_2 H$$

genügen. Andererseits findet man, wenn ein Lösungspaar u, v von (43) als bekannt vorausgesetzt wird,

$$(44) \quad p = \frac{v_\beta}{u_\beta}, \quad q = \frac{v_\alpha}{u_\alpha}$$

und erkennt unschwer, daß diese Größen wiederum die Differentialrelationen (39) erfüllen, die jetzt auf dem umgekehrten Wege aus dem Bestehen der Formeln (42) folgen. Mit der Zurückführung auf das System (43), an dessen Stelle man

$$\frac{du dv}{H} = \varepsilon_1 d\alpha^2 + \varepsilon_2 d\beta^2 + 2\mathfrak{F} d\alpha d\beta$$

²¹⁾ Die von Bianchi in *Lezioni 3, Kap. VI* gegebenen Entwicklungen, die an Ergebnisse von Darboux, Servant und Calapo anknüpfen, beziehen sich auf die Parameter des permanenten konjugierten Systems, die mit denjenigen der Asymptotenlinien durch eine einfache lineare Substitution zusammenhängen (siehe den folgenden Artikel 3). Betreffs des hier vorgetragenen Beweises vgl. Jonas, Über eine neue partielle Differentialgleichung des Deformationsproblems usw. *Berl. Math. Ges. Ber.* 13 (1914), S. 52.

²²⁾ Bezüglich der oben erwähnten Herleitung aus den Christoffelschen Invarianten vgl. Bianchi, *Lezioni 3, § 82* und Jonas, Über gewisse mit der Verbiegung des orthogonalen Hyperboloids zusammenhängende Flächen usw. *Berl. Math. Ges. Ber.* 24 (1925), S. 54, daselbst Art. 5.

schreiben kann, erweist sich das Biegungsproblem für die Flächen zweiten Grades als gleichbedeutend mit dem *Tschebyscheffschen Problem für die quadratische Differentialform* $\frac{du dv}{H}$.

3. Der Vollständigkeit halber mögen hier noch zwei Bemerkungen Platz finden, von denen wir im folgenden keinen Gebrauch machen werden. Die erste betrifft das *permanente, d. h. der Biegungsfläche S und dem Hyperboloid gemeinsame konjugierte System*²³⁾, das durch das Verschwinden der quadratischen Kovariante ihrer zweiten Fundamentalformen (34) und (30) definiert ist. Seine Differentialgleichung

$$L du^2 - N dv^2 = 0$$

geht, auf die Parameter α, β der Asymptotenlinien von S bezogen, in

$$(45) \quad d\alpha^2 - \varepsilon_1 \varepsilon_2 d\beta^2 = 0$$

über, wobei ε_1 und ε_2 durch (41) gegeben sind. Auf S ist dieses Kurvennetz, wie (45) erkennen läßt, gleichzeitig *isotherm-konjugiert*. Es fällt reell oder imaginär aus, je nachdem ob $\frac{L}{N} = pq > 0$ oder < 0 ist, und wird, da $\varepsilon_1 \cdot \varepsilon_2$ entsprechend den Wert $+1$ oder -1 hat, im ersten Falle durch $\alpha \pm \beta = \text{konst.}$, im zweiten durch $\alpha \pm i\beta = \text{konst.}$ dargestellt. Die betreffenden Biegungen werden nach Bianchi als *Biegungen erster und zweiter Art* unterschieden.

Der zweite, hier gleichfalls in aller Kürze zu erörternde Punkt, dessen prinzipielle Bedeutung für einen weiteren Ausbau der Bianchischen Theorie indessen nicht verkannt werden soll, ist ein transformatorischer Zusammenhang zwischen den Differentialgleichungen der Biegung, durch den zwei verschiedene Flächen zweiten Grades, im besonderen *zwei einschalige Hyperboloide*, einander als *biegungs-konjugiert* (quadriche coniugate in deformazione) zugeordnet werden²⁴⁾. Wir gehen dabei auf die geometrischen Betrachtungen, deren sich Bianchi bedient, nicht ein, sondern wählen einen einfachen rechnerischen Weg im Anschluß an die Form (39) der Biegungsgleichungen. Der durch (25) definierten Größe H geben wir zunächst, wobei wir die stets zulässige Annahme $a > b$ machen, mit Hilfe der Identität

$$(u+v)^2 - (1+uv)^2 - (u-v)^2 + (1-uv)^2 = 0$$

die folgende Gestalt:

$$H = \frac{1}{a^2}(u+v)^2 + \left(\frac{1}{a^2} + \frac{1}{c^2}\right)(1-uv)^2 + \left(\frac{1}{b^2} - \frac{1}{a^2}\right)(u-v)^2,$$

²³⁾ Bianchi, *Lezioni* 3, Kap. VI.

²⁴⁾ Bianchi, *Lezioni* 3, Kap. V.

substituieren

$$(46) \quad u = \frac{1}{\bar{u}}$$

und erhalten den zu (25) analogen Ausdruck

$$\bar{H} = \frac{H}{u^2} = \frac{1}{a^2}(1 + \bar{u}v)^2 + \frac{1}{b^2}(\bar{u} - v)^2 + \frac{1}{c^2}(1 - \bar{u}v)^2,$$

in dem

$$\bar{b} = \frac{ac}{\sqrt{a^2 + c^2}}, \quad \bar{c} = \frac{ab}{\sqrt{a^2 - b^2}}$$

ist. \bar{H} gehört zu einem neuen einschaligen Hyperboloid:

$$(47) \quad \begin{cases} \bar{x} = a \frac{1 + \bar{u}v}{\bar{u} + v}, & \bar{y} = \bar{b} \frac{\bar{u} - v}{\bar{u} + v}, & \bar{z} = \bar{c} \frac{1 - \bar{u}v}{\bar{u} + v}, \\ \bar{x}^2/a^2 + \bar{y}^2/\bar{b}^2 - \bar{z}^2/\bar{c}^2 = 1. \end{cases}$$

Setzt man nun noch

$$(48) \quad \bar{p} = -pu^2, \quad \bar{q} = -qu^2,$$

so verwandeln sich die Differentialgleichungen (39) in

$$(\bar{p}\bar{H})_{\bar{x}} + \bar{q}\bar{H}^2 \left(\frac{\bar{p}}{\bar{H}} \right)_v = 0, \quad (\bar{q}\bar{H})_{\bar{x}} + \bar{p}\bar{H}^2 \left(\frac{\bar{q}}{\bar{H}} \right)_v = 0,$$

also in diejenigen, von denen die Verbiegung des Hyperboloids (47) abhängt. Vermöge der Formeln (46), (48) bestimmt demnach eine bekannte Biegungsfläche S des Hyperboloids (22) gleichzeitig, allerdings nur *intrinsek*, d. h. durch die flächentheoretischen Fundamentalgrößen, eine Biegungsfläche \bar{S} des *biegungs-konjugierten* Hyperboloids (47). Seiner rein analytischen Natur entsprechend ist dieser von Bianchi als *Transformation H* bezeichnete Übergang einer geometrischen Deutung in Form einer Konstruktion nicht zugänglich. Die Biegungsflächen S und \bar{S} sind mit Korrespondenz der Asymptotenlinien aufeinander bezogen; in der Tat gehen die Differentialgleichungen (35) durch die Substitution (46), (48) in

$$\bar{p}d\bar{u} - dv = 0, \quad \bar{q}d\bar{u} - dv = 0$$

über²⁵⁾.

4. Wir wollen schließlich, bevor wir die Transformation B_k behandeln, die in § 1, 5 eingeführt, mit der Biegungsfläche S verbundenen *drei Achsenspurflächen* für den Fall des einschaligen Hyperboloids in einer für unsere Zwecke geeigneten Weise darstellen. Sie mögen $S^{(1)}$, $S^{(2)}$, $S^{(3)}$ ge-

²⁵⁾ Das durch die Relation $\frac{1}{b^2} = \frac{1}{a^2} + \frac{1}{c^2}$ gekennzeichnete *orthogonale* Hyperboloid fällt, wie Bianchi (a. a. O. S. 213) bemerkt hat, mit dem *biegungs-konjugierten* zusammen.

nannt werden. Der Punkt $(x^{(1)}, y^{(1)}, z^{(1)})$ von $S^{(1)}$ ist also Schnittpunkt der Tangentialebene von S mit der ξ -Achse des auf S rollenden Hyperboloids (23). Wir setzen dementsprechend:

$$x^{(1)} = x + \sigma x_u + \tau x_v$$

und haben die Koeffizienten σ und τ auf Grund der Bedingung

$$(49) \quad \eta + \sigma \eta_u + \tau \eta_v = 0, \quad \xi + \sigma \xi_u + \tau \xi_v = 0$$

zu bestimmen. Es empfiehlt sich, die dritte Gleichung

$$(50) \quad \xi + \sigma \xi_u + \tau \xi_v = \frac{a^2}{\xi}$$

hinzuzufügen, die man leicht als Folge der beiden vorausgehenden erkennt, indem man sie mit $\frac{\xi}{a^2}$ multipliziert, dazu die mit $\frac{\eta}{b^2}$ und $-\frac{\xi}{c^2}$ multiplizierten Gleichungen (49) addiert und dann die Gleichung (22) des Hyperboloids berücksichtigt. Die gewünschten Ausdrücke für σ und τ ergeben sich jetzt, wenn wir die drei Relationen (50), (49) der Reihe nach mit $\frac{\xi v}{a^2}$, $\frac{\eta v}{b^2}$, $-\frac{\xi v}{c^2}$ bzw. mit $\frac{\xi u}{a^2}$, $\frac{\eta u}{b^2}$, $-\frac{\xi u}{c^2}$ multiplizieren und unter Anwendung der unschwer zu bestätigenden Formeln

$$(51) \quad \left\{ \begin{array}{l} \frac{(\xi u)^2}{a^2} + \frac{(\eta u)^2}{b^2} - \frac{(\xi u)^2}{c^2} = 0, \quad \frac{(\xi v)^2}{a^2} + \frac{(\eta v)^2}{b^2} - \frac{(\xi v)^2}{c^2} = 0, \\ \frac{\xi u \xi v}{a^2} + \frac{\eta u \eta v}{b^2} - \frac{\xi u \xi v}{c^2} = -\frac{2}{(u+v)^2} \end{array} \right.$$

addieren. Wir finden

$$\sigma = -\frac{(u+v)^2 \xi v}{2\xi}, \quad \tau = -\frac{(u+v)^2 \xi u}{2\xi},$$

also

$$(52) \quad x^{(1)} = x - \frac{(u+v)^2}{2\xi} (\xi_v x_u + \xi_u x_v).$$

Die Richtungskosinus der Normalen von $S^{(1)}$ verhalten sich wie

$$(53) \quad (\xi_v x_u - \xi_u x_v) : (\xi_v y_u - \xi_u y_v) : (\xi_v z_u - \xi_u z_v).$$

Für $S^{(2)}$ und $S^{(3)}$ hat man entsprechend:

$$(54) \quad \left\{ \begin{array}{l} x^{(2)} = x - \frac{(u+v)^2}{2\eta} (\eta_v x_u + \eta_u x_v), \\ x^{(3)} = x - \frac{(u+v)^2}{2\xi} (\xi_v x_u + \xi_u x_v). \end{array} \right.$$

Der Deutlichkeit halber sei darauf hingewiesen, daß beim Übergang von (52), (54) zu den analogen Formeln für die y - und z -Koordinaten die deutschen Buchstaben unverändert bleiben.

Wiewohl wir der Darstellung der Achsenspurflächen durch (52), (54) im Hinblick auf § 5 den Vorzug geben — der gleiche Bau der Formeln gestattet uns dort, die Untersuchung auf $S^{(1)}$ zu beschränken —, mögen hier noch die einzeln jedenfalls bequemer zu handhabenden Formeln folgen, in denen die Koeffizienten durch die Parameter u, v ausgedrückt sind:

$$(55) \quad \begin{cases} x^{(1)} = x + \frac{u+v}{2(1+uv)} [(1-u^2)x_u + (1-v^2)x_v], \\ x^{(2)} = x + \frac{u+v}{u-v} [u x_u - v x_v], \\ x^{(3)} = x + \frac{u+v}{2(1-uv)} [(1+u^2)x_u + (1+v^2)x_v]. \end{cases}$$

§ 3.

Aufstellung der Bianchischen Transformation B_k .

1. Die geometrischen Überlegungen, auf Grund deren es Bianchi gelang, die durch W -Kongruenzen vermittelten Transformationen für die Biegungsflächen der allgemeinen Flächen zweiten Grades zu entwickeln, sollen auch uns hier als Ausgangspunkt dienen. Wichtige Vereinfachungen erfährt die rechnerische Behandlung des Problems.

Neben dem Hyperboloid (22), bei dem weiterhin $a > b$ vorausgesetzt ist, betrachten wir ein zweites der gleichen Schar konfokaler F^2 angehöriges einschaliges Hyperboloid:

$$(56) \quad \frac{x'^2}{a^2+k} + \frac{y'^2}{b^2+k} - \frac{z'^2}{c^2-k} = 1.$$

Die *charakteristische Konstante* k der Bianchischen Transformation B_k ist von Null verschieden und im Intervalle

$$c^2 \geq k \geq -b^2$$

zu wählen²⁶⁾. Wir setzen mit Bianchi

$$(57) \quad \sqrt{a^2+k} = a', \quad \sqrt{b^2+k} = b', \quad \sqrt{c^2-k} = c'.$$

Zur leichteren Unterscheidung sei das ursprüngliche Hyperboloid mit \mathfrak{H} oder \mathfrak{H}_1 und ebenso das konfokale Hyperboloid (56) mit \mathfrak{H}' oder \mathfrak{H}'_1 bezeichnet, je nachdem es auf die Parameter u, v bezogen wird oder Träger des durch ein neues Parameterpaar u_1, v_1 definierten Punktes sein soll. Es gelten dann also die Formeln:

²⁶⁾ Für $k=c^2$ und $k=-b^2$ (Fall der *singulären Transformationen*) tritt an die Stelle des konfokalen Hyperboloids (56) die Ebene des einen oder des anderen Fokalkegelschnitts, soweit sie von dessen Tangenten, die als Erzeugende anzusehen sind, überdeckt wird.

$$(58) \quad \left\{ \begin{array}{l} \mathfrak{S}) \quad \xi = a \frac{1+uv}{u+v}, \quad \eta = b \frac{u-v}{u+v}, \quad \zeta = c \frac{1-uv}{u+v}, \\ \mathfrak{S}_1) \quad \xi_1 = a \frac{1+u_1v_1}{u_1+v_1}, \quad \eta_1 = b \frac{u_1-v_1}{u_1+v_1}, \quad \zeta_1 = c \frac{1-u_1v_1}{u_1+v_1}, \\ \mathfrak{S}') \quad \xi' = a' \frac{1+uv}{u+v}, \dots, \quad \mathfrak{S}'_1) \quad \xi'_1 = a' \frac{1+u_1v_1}{u_1+v_1}, \dots \end{array} \right.$$

Die Größen u_1, v_1 , denen die Rolle der Parameter u, v in bezug auf die transformierte Biegungsfläche zukommen wird, sind als gesuchte Funktionen der zwei Variablen u und v aufzufassen. Zunächst werden die beiden Parameterpaare u, v und u_1, v_1 durch eine die charakteristische Konstante k enthaltende *endliche* Gleichung verbunden, die der Forderung entspricht, daß der Punkt (u_1, v_1) von \mathfrak{S}'_1 in der Tangentialebene des Punktes (u, v) von \mathfrak{S} liege. Sie lautet:

$$\frac{\xi \xi'_1}{a^2} + \frac{\eta \eta'_1}{b^2} - \frac{\zeta \zeta'_1}{c^2} = 1$$

und geht durch Einsetzen der Ausdrücke (58) und durch Fortschaffen der Nenner in

$$(59) \quad \varphi = \frac{a'}{a}(1+u_1v_1)(1+uv) + \frac{b'}{b}(u_1-v_1)(u-v) - \frac{c'}{c}(1-u_1v_1)(1-uv) - (u_1+v_1)(u+v) = 0$$

über. Wir nennen (59) die *quadrilineare Relation*²⁷⁾ der Transformation B_k . Die Einführung eines Funktionszeichens φ für die linke Seite der Gleichung erweist sich in der Folge als bedeutungsvoll, da sie die Möglichkeit bietet, die erforderlichen Rechnungen durch *formale* Operationen wesentlich abzukürzen. Die durch Indizes angedeuteten partiellen Ableitungen von φ wie $\varphi_u, \varphi_{u_1}, \varphi_{uv}$ usw. sind durchweg so zu verstehen, daß φ dabei als Funktion der vier Argumente u, v, u_1, v_1 gilt. Im Hinblick auf die inverse Transformation sei noch die Symmetrie des Ausdrucks φ bezüglich der beiden Parameterpaare hervorgehoben; $\varphi = 0$ besagt also auch, daß der Punkt (u, v) von \mathfrak{S}' in der Tangentialebene des Punktes (u_1, v_1) von \mathfrak{S}_1 liegt²⁸⁾.

²⁷⁾ Gleichwertig mit dem Formelpaar (89) auf S. 60 in Bd. 3 der Lezioni.

²⁸⁾ Die mit a', b', c' bezeichneten Quadratwurzeln (57) können als positiv gewählt vorausgesetzt werden; prinzipielle Gründe für den Ausschluß negativer Werte liegen nicht vor. Die Beseitigung von Minuszeichen gelingt mittels der Substitutionen der \mathfrak{S}_3 , die aus den Kongruenzen und Symmetrien des Hyperboloids \mathfrak{S}'_1 bzw. \mathfrak{S}_1 in sich selbst besteht. Um zwei Minuszeichen verschwinden zu lassen, hat man (u_1, v_1) durch $(-u_1, -v_1)$, $\left(\frac{1}{u_1}, \frac{1}{v_1}\right)$ oder $\left(-\frac{1}{u_1}, -\frac{1}{v_1}\right)$ zu ersetzen, um eines oder drei zu beseitigen, durch (v_1, u_1) , $(-v_1, -u_1)$, $\left(\frac{1}{v_1}, \frac{1}{u_1}\right)$ oder $\left(-\frac{1}{v_1}, -\frac{1}{u_1}\right)$. Die vier letzt-

2. Wir benutzen jetzt, um den Punkt $\xi'_1, \eta'_1, \zeta'_1$ von \mathfrak{S}'_1 als Punkt der Tangentialebene von $\mathfrak{S}(\xi, \eta, \zeta)$ darzustellen, den (gleicherweise in η und ζ gültigen) Ansatz:

$$(60) \quad \xi'_1 = \xi + \lambda \xi_u + \mu \xi_v,$$

der uns dazu dienen soll, die am Hyperboloid \mathfrak{S} vorzunehmende Konstruktion auf die in § 2 betrachtete Biegungsfläche $S(x, y, z)$ zu übertragen. Um die beiden Koeffizienten λ, μ durch die zunächst nur der Relation $\varphi = 0$ unterworfenen Größen u, v, u_1, v_1 auszudrücken, gehen wir von der folgenden ohne Rücksicht auf das Verschwinden von φ bestehenden Gleichung aus, deren Richtigkeit man mit Hilfe von (29), (58), (59) leicht bestätigt:

$$(61) \quad \sum (\xi'_1 - \xi) \mathfrak{X} = - \frac{\varphi}{(u_1 + v_1) \sqrt{H}}.$$

Sie darf als Identität nach v differenziert werden. Wird nachträglich $\varphi = 0$ gesetzt, so folgt:

$$\sum \xi'_1 \mathfrak{X}_v = - \frac{\varphi_v}{(u_1 + v_1) \sqrt{H}}.$$

Tragen wir für ξ'_1 den Ausdruck (60) ein und beachten, daß wegen (30)

$$\sum \xi_u \mathfrak{X}_v = - \mathfrak{M} = \frac{2}{(u+v) \sqrt{H}}, \quad \sum \xi_v \mathfrak{X}_v = 0$$

ist, so finden wir λ und haben unter Hinzufügung der analog gebildeten Formel für μ :

$$(62) \quad \lambda = - \frac{u+v}{2(u_1+v_1)} \varphi_v, \quad \mu = - \frac{u+v}{2(u_1+v_1)} \varphi_u,$$

so daß (60) in

$$(63) \quad \xi'_1 = \xi - \frac{u+v}{2(u_1+v_1)} (\varphi_v \xi_u + \varphi_u \xi_v)$$

übergeht.

Nehmen wir nun die Biegungsfläche $S(x, y, z)$, so wird durch die entsprechende Formelgruppe

$$(64) \quad x_1 = x - \frac{u+v}{2(u_1+v_1)} (\varphi_v x_u + \varphi_u x_v)$$

ein Punkt x_1, y_1, z_1 in der Tangentialebene von S definiert, der eine bestimmte Fläche $S_1(x_1, y_1, z_1)$ beschreibt, sobald neben der Bedingung $\varphi = 0$ den beiden Parameterpaaren u, v und u_1, v_1 eine weitere Relation auferlegt wird. Diese kann, wie im folgenden Artikel gezeigt werden soll, noch auf ∞^1 Weisen, nämlich als Integral einer totalen Differential-

genannten, den Symmetrien entsprechenden Substitutionen ziehen in Folge des Buchstabenwechsels die Vertauschung der Vorzeichen $-\varepsilon$ und $+\varepsilon$ in den Differentialgleichungen (75) nach sich und ändern damit die Klasse der Transformation (siehe den Schluß des folgenden Art., sowie § 4, 1).

gleichung, so gewählt werden, daß S_1 eine neue Biegungsfläche des ursprünglichen Hyperboloids wird oder, wie wir, um die isometrische Abbildung als durch die Parameter u_1, v_1 bewirkt zu kennzeichnen, sagen wollen, daß $S_1(x_1, y_1, z_1)$ auf $\mathfrak{S}_1(\xi_1, \eta_1, \zeta_1)$ abwickelbar wird.

Den geometrischen Inhalt der Formeln (63), (64) sprechen wir wie folgt aus: Rollt das Hyperboloid $\mathfrak{S}(\xi, \eta, \zeta)$ unter Mitführung der mit den Flächenelementen gekoppelten Tangentialebenen auf der Biegungsfläche $S(x, y, z)$, so fällt der in der Tangentialebene von \mathfrak{S} markierte Punkt $\xi'_1, \eta'_1, \zeta'_1$ des konfokalen Hyperboloids auf den in der Tangentialebene von S gelegenen Punkt x_1, y_1, z_1 . Statt dessen können wir auch sagen: Wird die Biegungsfläche S (mit geeigneten Schnitten versehen) auf das Hyperboloid \mathfrak{S} abgewickelt, so gehen die Endpunkte x_1, y_1, z_1 der mitgeführten Tangentenabschnitte in die Punkte $\xi'_1, \eta'_1, \zeta'_1$ des konfokalen Hyperboloids über.

Es ist schon mit Rücksicht auf diese geometrischen Vorstellungen selbstverständlich, daß wir für \mathfrak{S} und S die beiden von Flächennormale und Parameterrichtungen gebildeten Dreikante als *gleichsinnig* voraussetzen, so daß sie beim Rollen oder bei der Abwicklung zur Deckung gelangen. Es gilt demnach, wenn X, Y, Z die Normalenkosinus von S bedeuten, die den Formeln (27), (28) entsprechende Festsetzung des Vorzeichens:

$$(65) \quad X = \frac{(u+v)^2}{2abc\sqrt{H}}(y_u z_v - z_u y_v),$$

so daß also

$$(66) \quad \begin{vmatrix} x_u & y_u & z_u \\ x_v & y_v & z_v \\ X & Y & Z \end{vmatrix} = \frac{2abc\sqrt{H}}{(u+v)^2}$$

wird.

3. Wir differenzieren jetzt die durch (64) und (63) gegebenen Koordinaten x_1 und ξ'_1 nach den Variablen u und v , wobei wir u_1 und v_1 als Funktionen von u, v behandeln und von den Formeln der allgemeinen Flächentheorie Gebrauch machen, vermöge deren sich die zweiten Ableitungen von x bzw. von ξ linear durch die ersten und durch die Normalenkosinus ausdrücken lassen:

$$(67) \quad \begin{cases} (x_1)_u = P x_u + Q x_v - \frac{u+v}{2(u_1+v_1)}(\varphi_v L + \varphi_u M) X, \\ (x_1)_v = R x_u + S x_v - \frac{u+v}{2(u_1+v_1)}(\varphi_v M + \varphi_u N) X, \end{cases}$$

$$(68) \quad \begin{cases} (\xi'_1)_u = P \xi_u + Q \xi_v - \frac{u+v}{2(u_1+v_1)} \varphi_u \mathfrak{M} \mathfrak{X}, \\ (\xi'_1)_v = R \xi_u + S \xi_v - \frac{u+v}{2(u_1+v_1)} \varphi_v \mathfrak{M} \mathfrak{X}. \end{cases}$$

Die umständlichen Ausdrücke für P, Q, R, S sind nicht erforderlich. Es genügt die bloße Feststellung, daß die beiden Formelsysteme in diesen vier Koeffizienten übereinstimmen. Man findet mittels (67) und (68):

$$\sum dx_1^2 = \sum d\xi_1'^2 + \frac{(u+v)^2}{4(u_1+v_1)^2} \{ [(\varphi_v L + \varphi_u M) du + (\varphi_v M + \varphi_u N) dv]^2 - \mathfrak{M}^2 (\varphi_u du + \varphi_v dv)^2 \}.$$

$\sum d\xi_1'^2$ kann durch u_1, v_1 ausgedrückt werden. Man hat in den Formeln (24) u, v durch u_1, v_1 und a^2, b^2, c^2 bezüglich durch $a^2 + k, b^2 + k, c^2 - k$ zu ersetzen. Nimmt man nun das ursprüngliche, aber auf u_1, v_1 bezogene Hyperboloid hinzu, das wir in (58) mit $\mathfrak{S}_1(\xi_1, \eta_1, \zeta_1)$ bezeichnet haben, so wird, wie man unmittelbar übersieht:

$$\sum dx_1^2 = \sum d\xi_1'^2 + \frac{4k du_1 dv_1}{(u_1 + v_1)^2}.$$

Soll demnach $S_1(x_1, y_1, z_1)$ auf $\mathfrak{S}_1(\xi_1, \eta_1, \zeta_1)$ abwickelbar, also

$$\sum dx_1^2 = \sum d\xi_1'^2$$

sein, so muß die Gleichung

$$(69) \quad 4k du_1 dv_1 = \frac{(u+v)^2}{4} \{ [(\varphi_v L + \varphi_u M) du + (\varphi_v M + \varphi_u N) dv]^2 - \mathfrak{M}^2 (\varphi_u du + \varphi_v dv)^2 \}$$

bestehen, die wir mit der aus $\varphi = 0$ folgenden Differentialrelation

$$(70) \quad \varphi_{u_1} du_1 + \varphi_{v_1} dv_1 = -\varphi_u du - \varphi_v dv$$

zu kombinieren haben.

Wir brauchen dazu die Beziehung

$$(71) \quad \varphi_{u_1} \varphi_{v_1} = -kH.$$

Um dieselbe zu bestätigen, hat man φ in der Form

$$(72) \quad \varphi = \mathfrak{A}u_1 v_1 + \mathfrak{B}u_1 + \mathfrak{C}v_1 + \mathfrak{D}$$

zu schreiben, wobei

$$(73) \quad \begin{cases} \mathfrak{A} = \frac{a'}{a}(1+uv) + \frac{c'}{c}(1-uv), & \mathfrak{B} = \frac{b'}{b}(u-v) - (u+v), \\ \mathfrak{C} = -\frac{b'}{b}(u-v) - (u+v), & \mathfrak{D} = \frac{a'}{a}(1+uv) - \frac{c'}{c}(1-uv) \end{cases}$$

ist, und

$$\varphi_{u_1} \varphi_{v_1} = \varphi_{u_1} \varphi_{v_1} - \varphi \varphi_{u_1 v_1} = \mathfrak{B}\mathfrak{C} - \mathfrak{A}\mathfrak{D}$$

mit Benutzung der Ausdrücke (72) zu bilden.

Quadriert man nun die beiden Seiten von (70), addiert dazu die mit H multiplizierte Relation (69), berücksichtigt links (71) und setzt rechts

den durch (31) gegebenen Ausdruck

$$H = \frac{4}{\mathfrak{R}^2(u+v)^2}$$

ein, so folgt, wenn man beiderseits die Quadratwurzel nimmt:

$$(74) \quad \varphi_{u_1} du_1 - \varphi_{v_1} dv_1 = -\frac{\varepsilon}{\mathfrak{R}} [(\varphi_v L + \varphi_u M) du + (\varphi_v M + \varphi_u N) dv] \\ (\varepsilon = \pm 1).$$

Unter abermaliger Heranziehung von (70) ergeben sich so die beiden mit der Bedingung $\varphi = 0$ jedenfalls vereinbaren simultanen totalen Differentialgleichungen für u_1 und v_1 , von denen die Transformation B_x abhängt:

$$(75) \quad \left\{ \begin{array}{l} du_1 = -\frac{1}{2\varphi_{u_1}}(\varphi_u du + \varphi_v dv) \\ \quad -\frac{\varepsilon}{2\mathfrak{R}\varphi_{u_1}}[(\varphi_v L + \varphi_u M) du + (\varphi_v M + \varphi_u N) dv], \\ dv_1 = -\frac{1}{2\varphi_{v_1}}(\varphi_u du + \varphi_v dv) \\ \quad +\frac{\varepsilon}{2\mathfrak{R}\varphi_{v_1}}[(\varphi_v L + \varphi_u M) du + (\varphi_v M + \varphi_u N) dv]. \end{array} \right.$$

Dem noch zu liefernden Nachweis, daß dieses System unbeschränkt integrierbar ist, lassen wir die folgenden Bemerkungen voraufgehen. In der von Bianchi gegebenen Darstellung seiner Theorie wird von vornherein der eine der beiden neuen Parameter bevorzugt. Es erscheint beachtenswert, daß wir die eine Größe, z. B. v_1 , auf formalem Wege eliminieren können, d. h. ohne genötigt zu sein, die bei Bianchi auftretenden Ausdrücke wirklich zu bilden. Verföhrt man nämlich wie bei der Aufstellung der Formel (71), so gelingt es, auch die übrigen Produkte, insbesondere $\varphi_u \varphi_{v_1}$ und $\varphi_v \varphi_{v_1}$, als Funktionen der beiden nicht als Ableitungsindizes fungierenden Größen darzustellen. Es wird:

$$(76) \quad \varphi_u \varphi_{v_1} = \varphi_u \varphi_{v_1} - \varphi \varphi_{u v_1} = \chi^{(1)}(v, u_1), \quad \varphi_v \varphi_{v_1} = \chi^{(2)}(u, u_1),$$

wobei die Polynome $\chi^{(1)}$ und $\chi^{(2)}$, ähnlich wie H , bezüglich jedes ihrer beiden Argumente vom zweiten Grade sind. Wir sind so in der Lage, die erste der Differentialgleichungen von v_1 zu befreien, indem wir die folgende Umformung vornehmen:

$$(77) \quad \frac{\varphi_u}{\varphi_{u_1}} = \frac{\varphi_u \varphi_{v_1}}{\varphi_{u_1} \varphi_{v_1}} = -\frac{\chi^{(1)}(v, u_1)}{kH}, \quad \frac{\varphi_v}{\varphi_{v_1}} = \frac{\varphi_v \varphi_{v_1}}{\varphi_{u_1} \varphi_{v_1}} = -\frac{\chi^{(2)}(u, u_1)}{kH}.$$

Damit reduziert sich das System (75) auf eine einzige totale Differentialgleichung vom Riccatischen Typus für u_1 , während sich v_1 nachträglich aus der Relation $\varphi = 0$ ergibt.

Auch aus den Formeln (64), die der zwischen S und S_1 vermittelnden geometrischen Konstruktion entsprechen, läßt sich v_1 ohne eine bis ins einzelne gehende Rechnung entfernen. Man hat dazu die in (62) mit λ und μ bezeichneten Brüche mit φ_{v_1} zu erweitern, auf die Zähler die Beziehungen (76) anzuwenden und im Nenner

$$(u_1 + v_1)\varphi_{v_1} = (u_1 + v_1)\varphi_{v_1} - \varphi = \mathfrak{A}u_1^2 - (\mathfrak{B} - \mathfrak{C})u_1 - \mathfrak{D}$$

zu schreiben. Der Übergang zu den grundlegenden Entwicklungen in Bd. 3 der *Lezioni* ist hiermit klargelegt.

Durch das Auftreten des doppelten Vorzeichens $\varepsilon = \pm 1$ zerfallen die Transformationen B_k in die von Bianchi unterschiedenen *beiden Klassen*. Eine in § 4, 1 aufzustellende, für das Endziel unserer Untersuchung wichtige Formelgruppe wird uns Gelegenheit geben, auf die geometrische Deutung dieses Gegensatzes einzugehen.

4. Um jetzt die *unbeschränkte Integrierbarkeit* des Systems (75) zu bestätigen, brauchen wir nur die *erste* der beiden totalen Differentialgleichungen ins Auge zu fassen, wofern wir uns v_1 mittels $\varphi = 0$ auf dem angegebenen Wege eliminiert denken. Wir ersetzen dieselbe durch die beiden partiellen Differentialgleichungen:

$$(78) \quad -2\varepsilon(u_1)_u = B H l + A(m + \varepsilon), \quad -2\varepsilon(u_1)_v = A H n + B(m + \varepsilon),$$

in denen

$$(79) \quad A = \frac{\varphi_u}{\varphi_{u_1}}, \quad B = \frac{\varphi_v}{\varphi_{u_1}}$$

ist und an Stelle von L, M, N mit Rücksicht auf die einfachere Gestalt der Codazzischen Gleichungen, die durch (32) definierten Größen l, m, n eingeführt sind. Nachzuweisen ist das Bestehen der Relation:

$$[B H l + A(m + \varepsilon)]_v - [A H n + B(m + \varepsilon)]_u = 0,$$

die sich als Identität darstellen muß, wenn beim Differentiieren A und B zunächst als Funktionen von u, v, u_1 behandelt werden (in diesem Sinne seien die Bezeichnungen $\frac{\partial}{\partial u}, \frac{\partial}{\partial v}, \frac{\partial}{\partial u_1}$ verstanden), und für die Ableitungen von u_1 nach den unabhängigen Variablen u und v wieder die den Differentialgleichungen (78) entnommenen Ausdrücke eintreten. Die linke Seite der fraglichen Relation wird auf diese Weise:

$$\begin{aligned} & \frac{\partial}{\partial v} [B H l + A(m + \varepsilon)] - \frac{\partial}{\partial u} [A H n + B(m + \varepsilon)] \\ & - \frac{\varepsilon}{2} \left[\frac{\partial B}{\partial u_1} H l + \frac{\partial A}{\partial u_1} (m + \varepsilon) \right] [A H n + B(m + \varepsilon)] \\ & + \frac{\varepsilon}{2} \left[\frac{\partial A}{\partial u_1} H n + \frac{\partial B}{\partial u_1} (m + \varepsilon) \right] [B H l + A(m + \varepsilon)]. \end{aligned}$$

Wir beachten, daß infolge von (77)

$$\frac{\partial(BH)}{\partial v} = 0, \quad \frac{\partial(AH)}{\partial u} = 0$$

wird und benutzen die Beziehungen (33) zur weiteren Reduktion. Es bleibt nach Unterdrückung des gemeinsamen Faktors $m + \varepsilon$ der Ausdruck:

$$\Omega = \frac{\partial A}{\partial v} - \frac{\partial B}{\partial u} + A \frac{\partial B}{\partial u_1} - B \frac{\partial A}{\partial u_1}.$$

Um auf kürzestem Wege das Verschwinden von Ω festzustellen, nehmen wir A und B jetzt wieder in der ursprünglichen Gestalt der auch v_1 explizite enthaltenden Brüche (79) an. Die Differentiationen sind dann so auszuführen, daß v_1 als eine durch $\varphi = 0$ gegebene Funktion der drei Argumente u, v, u_1 aufgefaßt und dementsprechend

$$\frac{\partial A}{\partial v} = A_v + A_{v_1} \frac{\partial v_1}{\partial v} = A_v - A_{v_1} \frac{\varphi_v}{\varphi_{v_1}} \text{ usw.}$$

eingesetzt wird. Dabei deuten wir durch Indizes die partiellen Ableitungen von A und B nach je einem der vier Argumente u, v, u_1, v_1 an, wie dies für φ allgemein verabredet wurde. Es ergibt sich so:

$$\Omega = A_v - B_u + AB_{u_1} - BA_{u_1} - \frac{1}{\varphi_{v_1}} [A_{v_1} \varphi_v - B_{v_1} \varphi_u + (AB_{v_1} - BA_{v_1}) \varphi_{u_1}],$$

und hieraus, da die eckige Klammer wegen (79) gleich Null ist:

$$\Omega = A_v - B_u + AB_{u_1} - BA_{u_1}.$$

Führt man schließlich die Differentiationen an den Ausdrücken (79) *formal* aus und berücksichtigt dabei, daß $\varphi_{u, u_1} = 0$ ist, so folgt $\Omega = 0$, w. z. b. w.

Das allgemeine Integral u_1 der hiermit als unbeschränkt integrabel erkannten Riccatischen Differentialgleichung enthält eine willkürliche Konstante, den Wert von u_1 für ein anfängliches Wertepaar $u = u^{(0)}, v = v^{(0)}$. Jede der beiden Klassen umfaßt demnach bei gegebenem k eine Schar von ∞^1 Transformationen B_k . Geometrisch entspricht dem Auftreten der Integrationskonstante die Tatsache, daß für *einen* Punkt $(u^{(0)}, v^{(0)})$ der Biegungsfläche S der zugehörige Punkt der transformierten Fläche S_1 auf dem in der Tangentialebene gelegenen, durch das konfokale Hyperboloid bestimmten Kegelschnitt beliebig angenommen werden darf.

§ 4.

Formeln für die transformierte Biegungsfläche. Inverse Transformation. Korrespondenz der Asymptotenlinien.

1. Die neue Biegungsfläche $S_1(x_1, y_1, z_1)$, die aus $S(x, y, z)$ mittels einer Transformation B_k hervorgeht, werde jetzt auf die Parameter u_1, v_1 bezogen, die den Erzeugenden des Hyperboloids \mathfrak{H}_1 entsprechen. Es han-

delt sich zunächst darum, die Ableitungen der Koordinaten x_1, y_1, z_1 nach den Variablen u_1 und v_1 auf die Form

$$(80) \quad \begin{cases} (x_1)_{u_1} = a_1 x_u + b_1 x_v + c_1 X, \\ (x_1)_{v_1} = a_2 x_u + b_2 x_v + c_2 X \end{cases}$$

zu bringen. Werden für das bei der Konstruktion benutzte konfokale Hyperboloid $\mathfrak{H}'_1(\xi'_1, \eta'_1, \zeta'_1)$ analoge Beziehungen angesetzt, so können sich diese, wie man leicht aus (67) und (68) schließt, nur durch die an dritter Stelle stehenden Koeffizienten von den Formeln (80) unterscheiden. Es wird also gleichzeitig:

$$(81) \quad \begin{cases} (\xi'_1)_{u_1} = a_1 \xi_u + b_1 \xi_v + c'_1 X, \\ (\xi'_1)_{v_1} = a_2 \xi_u + b_2 \xi_v + c'_2 X. \end{cases}$$

Wir schreiben nun

$$(x_1)_{u_1} du_1 + (x_1)_{v_1} dv_1 = (x_1)_u du + (x_1)_v dv$$

und drücken die Ableitungen von x_1 links durch (80) und rechts durch (67) aus. Identifizieren wir alsdann insbesondere die Faktoren von X , so folgt:

$$c_1 du_1 + c_2 dv_1 = -\frac{u+v}{2(u_1+v_1)} [(\varphi_u L + \varphi_u M) du + (\varphi_v M + \varphi_u N) dv].$$

Auf demselben Wege ergibt sich aus (81) und (68):

$$c'_1 du_1 + c'_2 dv_1 = -\frac{u+v}{2(u_1+v_1)} \mathfrak{M}(\varphi_u du + \varphi_v dv).$$

Wir ersetzen in den beiden letzten Gleichungen du_1 und dv_1 durch die Ausdrücke (75) und finden so mit Berücksichtigung von (31):

$$(82) \quad \begin{cases} c'_1 = -\frac{\varphi_{u_1}}{(u_1+v_1)\sqrt{H}}, & c'_2 = -\frac{\varphi_{v_1}}{(u_1+v_1)\sqrt{H}}, \\ c_1 = \varepsilon c'_1, & c_2 = -\varepsilon c'_2, \end{cases}$$

wobei $\varepsilon = \pm 1$ durch die Klasse der Transformation bestimmt ist.

Zwecks Ermittlung der übrigen vier Koeffizienten gehen wir von der Identität (61) aus, der wir, da

$$\sum \xi X = -\frac{u+v}{\sqrt{H}}$$

ist, die folgende Gestalt geben können:

$$\sum \xi'_1 X = -\frac{\varphi}{(u_1+v_1)\sqrt{H}} - \frac{u+v}{\sqrt{H}}.$$

Um a_1 zu berechnen, differenzieren wir sie, ohne auf die Abhängigkeit der vier Argumente Rücksicht zu nehmen, nacheinander nach u_1 und v

und beachten, daß infolge von (81)

$$\Sigma(\mathcal{E}'_1)_{u_1} \mathcal{X}_v = -a_1 \mathfrak{M}$$

wird. Wir erhalten damit für a_1 und auf entsprechende Weise für b_1, a_2, b_2 die Ausdrücke:

$$a_1 = \frac{1}{\mathfrak{M}} \left(\frac{\varphi}{(u_1 + v_1) \sqrt{H}} \right)_{u_1 v}, \quad b_1 = \frac{1}{\mathfrak{M}} \left(\frac{\varphi}{(u_1 + v_1) \sqrt{H}} \right)_{u_1 u},$$

$$a_2 = \frac{1}{\mathfrak{M}} \left(\frac{\varphi}{(u_1 + v_1) \sqrt{H}} \right)_{v_1 v}, \quad b_2 = \frac{1}{\mathfrak{M}} \left(\frac{\varphi}{(u_1 + v_1) \sqrt{H}} \right)_{v_1 u} \quad .^{29)}$$

Führt man die Differentiationen aus, setzt nachträglich $\varphi = 0$ und macht wieder Gebrauch von (31), so wird:

$$(83) \quad \left\{ \begin{array}{l} a_1 = -\frac{u+v}{2(u_1+v_1)} \left(\varphi_{u_1 v} - \frac{1}{2} \frac{H_v}{H} \varphi_{u_1} - \frac{\varphi_v}{u_1+v_1} \right), \\ b_1 = -\frac{u+v}{2(u_1+v_1)} \left(\varphi_{u_1 u} - \frac{1}{2} \frac{H_u}{H} \varphi_{u_1} - \frac{\varphi_u}{u_1+v_1} \right), \\ a_2 = -\frac{u+v}{2(u_1+v_1)} \left(\varphi_{v_1 v} - \frac{1}{2} \frac{H_v}{H} \varphi_{v_1} - \frac{\varphi_v}{u_1+v_1} \right), \\ b_2 = -\frac{u+v}{2(u_1+v_1)} \left(\varphi_{v_1 u} - \frac{1}{2} \frac{H_u}{H} \varphi_{v_1} - \frac{\varphi_u}{u_1+v_1} \right). \end{array} \right.$$

Die erhaltenen Formeln gestatten uns zunächst, eine bekannte geometrische Eigenschaft der Transformation B_k zu betätigen. Einem jeden Punkte des Kegelschnitts, der in der Tangentialebene von S mittels \mathcal{S}'_1 festgelegt wurde, sind, insofern er ein Punkt x_1, y_1, z_1 je einer transformierten Fläche S_1 innerhalb jeder der beiden Klassen wird, zwei durch ihn hindurchgehende Ebenen als Tangentialebenen zugeordnet. Diese Ebenen werden, da L, M, N in den Koeffizienten von (64) nicht auftreten, bei beliebiger Verbiegung der Fläche S von dieser in starrer Koppelung mit ihren Flächenelementen mitgenommen. Die Integration des Systems (75) im Verein mit $\varphi = 0$ löst also die Aufgabe, die $2 \cdot \infty^3$ Facetten, bestehend aus Punkt und Ebene, entsprechend der jeweils durch L, M, N bestimmten Gestalt der Fläche S als Flächenelemente zweier Scharen von ∞^1 transformierten Flächen S_1 anzuordnen.

Dazu kommt noch die folgende, aus (80), (81), (82) ersichtliche Tatsache, die dazu dient, die beiden Transformationsklassen geometrisch zu unterscheiden. Für $\varepsilon = +1$ wird nämlich $c_1 = c'_1, c_2 = -c'_2$, für $\varepsilon = -1$ dagegen $c_1 = -c'_1, c_2 = c'_2$. Demnach fällt von den Erzeugenden (u_1, v_1) des konfokalen Hyperboloids \mathcal{S}'_1 , das von dem ursprünglichen Hyperboloid \mathcal{S} beim Rollen auf der Biegungsfläche S mitgeführt wird, jedesmal die eine mit der Tangente an die entsprechende Parameterlinie von S_1 zusammen,

²⁹⁾ Es ist klar, daß H vor Ausführung der Differentiationen nur den Ausdruck (25) bedeutet und nicht etwa mittels der Formel (71) dargestellt werden darf.

und zwar $v_1 = \text{konst.}$ für $\varepsilon = +1$, $u_1 = \text{konst.}$ für $\varepsilon = -1$; die andere dagegen nimmt die zu der entsprechenden Tangente von S_1 symmetrische Lage in bezug auf die Tangentialebene von S ein.

2. Auf die Formeln des vorigen Artikels stützen wir uns ferner bei dem Nachweis, daß die inverse Operation, durch die S aus S_1 hervorgeht, gleichfalls eine Transformation B_k vorstellt. Wir beachten, daß $c_1 \varphi_{v_1} + c_2 \varphi_{u_1} = 0$ ist, und finden mittels (80):

$$(84) \quad \varphi_{v_1}(x_1)_{u_1} + \varphi_{u_1}(x_1)_{v_1} = (a_1 \varphi_{v_1} + a_2 \varphi_{u_1}) x_u + (b_1 \varphi_{v_1} + b_2 \varphi_{u_1}) x_v.$$

An Hand der Beziehungen (83) bilden wir:

$$a_1 \varphi_{v_1} + a_2 \varphi_{u_1} = -\frac{u+v}{2(u_1+v_1)} \left[(\varphi_{u_1} \varphi_{v_1})_v - \frac{H_v}{H} \varphi_{u_1} \varphi_{v_1} - \frac{\varphi_v (\varphi_{u_1} + \varphi_{v_1})}{u_1 + v_1} \right].$$

Differenziert man nun die Identität

$$\varphi_u \varphi_{v_1} - \varphi \varphi_{u_1 v_1} = -kH$$

partiell nach v , so folgt für $\varphi = 0$ nach einer leichten Umformung der rechten Seite:

$$(\varphi_u \varphi_{v_1})_v - \varphi_v \varphi_{u_1 v_1} = \frac{H_v}{H} \varphi_{u_1} \varphi_{v_1}.$$

Mithin wird:

$$a_1 \varphi_{v_1} + a_2 \varphi_{u_1} = -\frac{(u+v) \varphi_v}{2(u_1+v_1)} \left(\varphi_{u_1 v_1} - \frac{\varphi_{u_1} + \varphi_{v_1}}{u_1 + v_1} \right).$$

Zur weiteren Vereinfachung dient die Beziehung:

$$\varphi_{u_1 v_1} - \frac{\varphi_{u_1} + \varphi_{v_1}}{u_1 + v_1} = \frac{2(u+v)}{u_1 + v_1},$$

die man aus (72) mit Benutzung der Werte (73) von \mathfrak{B} und \mathfrak{C} herleitet. Man hat hiernach unter Hinzufügung einer zweiten analog gebildeten Formel:

$$(85) \quad \begin{cases} a_1 \varphi_{v_1} + a_2 \varphi_{u_1} = -\frac{(u+v)^2}{(u_1+v_1)^2} \varphi_v, \\ b_1 \varphi_{v_1} + b_2 \varphi_{u_1} = -\frac{(u+v)^2}{(u_1+v_1)^2} \varphi_u. \end{cases}$$

Setzt man diese Ausdrücke in (84) ein, so ergibt sich die beachtenswerte Relation:

$$\varphi_{v_1}(x_1)_{u_1} + \varphi_{u_1}(x_1)_{v_1} = -\frac{(u+v)^2}{(u_1+v_1)^2} (\varphi_v x_u + \varphi_u x_v),$$

durch die man aus der Formelgruppe (64) unmittelbar ihr zu der inversen Transformation gehöriges Gegenstück:

$$(86) \quad x = x_1 - \frac{u_1 + v_1}{2(u+v)} [\varphi_{v_1}(x_1)_{u_1} + \varphi_{u_1}(x_1)_{v_1}]$$

gewinnt. Man erkennt daraus zunächst, daß die Tangenten von S , auf denen die korrespondierenden Punkte von S_1 liegen, gleichzeitig Tangenten an S_1 sind.

Nimmt man jetzt das auf u, v bezogene konfokale Hyperboloid $\mathfrak{S}'(\xi', \eta', \zeta')$ hinzu, so besteht neben (86), der Beziehung (63) entsprechend, die folgende:

$$(87) \quad \xi' = \xi_1 - \frac{u_1 + v_1}{2(u + v)} [\varphi_{v_1}(\xi_1)_{u_1} + \varphi_{u_1}(\xi_1)_{v_1}].$$

Man braucht sich daher nur die Entwicklung in § 4, 3 zu vergegenwärtigen, um zu übersehen, daß sich für die inverse Transformation Schritt für Schritt die gleichen Formeln lediglich mit veränderter Stellung des Index 1 ergeben. Ein prinzipieller, aber für die Rechnung belangloser Unterschied liegt nur darin, daß die durch die Gleichung

$$\sum dx^2 = \sum d\xi^2$$

ausgedrückte Isometrie von $S(x, y, z)$ und $\mathfrak{S}(\xi, \eta, \zeta)$ hier bereits feststeht und nicht erst gefordert wird. Angemerkt sei als Gegenstück zu (74) die Differentialrelation:

$$(88) \quad \varphi_u du - \varphi_v dv = -\frac{\bar{\varepsilon}}{\mathfrak{M}_1} [(\varphi_{v_1} L_1 + \varphi_{u_1} M_1) du_1 + (\varphi_{v_1} M_1 + \varphi_{u_1} N_1) dv_1] \\ (\bar{\varepsilon}^2 = 1),$$

in der L_1, M_1, N_1 und \mathfrak{M}_1 die zweiten Fundamentalgrößen der auf u_1, v_1 bezogenen Biegungsfläche S_1 und des Hyperboloids \mathfrak{S}_1 bedeuten. Die im folgenden noch auftretenden Größen $H_1, X_1, \dots, \mathfrak{X}_1, \dots$ bedürfen hiernach keiner besonderen Erklärung.

Wichtig ist es noch nachzuweisen, daß in (88) $\bar{\varepsilon} = \varepsilon$ ist, daß also der Satz gilt: Die inverse Transformation B_k gehört der gleichen Klasse an.

Wir gehen dazu von der Relation

$$\sum (x_1)_{u_1} X = -\frac{\varepsilon \varphi_{u_1}}{(u_1 + v_1) \sqrt{H}}$$

aus, die man leicht aus (80) und (82) gewinnt. Ihr entspricht bei der inversen Transformation B_k jedenfalls die folgende Gleichung:

$$(89) \quad \sum x_u X_1 = -\frac{\varepsilon \varphi_u}{(u + v) \sqrt{H_1}}.$$

Wir gelangen zu dem gewünschten Ergebnis, indem wir die linke Seite derselben für sich berechnen. Bilden wir nämlich X_1, \dots nach dem Muster der Formel (65), ersetzen dann $(x_1)_{u_1}, \dots, (x_1)_{v_1}, \dots$ durch die Ausdrücke (80) und beachten die Multiplikationsregel für Determinanten sowie die Gleichung (66), so wird der Reihe nach:

$$\begin{aligned} \sum x_u X_1 &= \frac{(u_1 + v_1)^3}{2abc\sqrt{H_1}} \sum x_u [(y_1)_{u_1} (z_1)_{v_1} - (z_1)_{u_1} (y_1)_{v_1}] \\ &= \frac{(u_1 + v_1)^3}{2abc\sqrt{H_1}} \left| \begin{array}{ccc} x_u, & & \dots \\ a_1 x_u + b_1 x_v + c_1 X, & \dots & \\ a_2 x_u + b_2 x_v + c_2 X, & \dots & \end{array} \right| \\ &= \frac{(u_1 + v_1)^3 \sqrt{H}}{(u+v)^3 \sqrt{H_1}} (b_1 c_2 - b_2 c_1). \end{aligned}$$

Aus (82) und (85) ergibt sich aber:

$$b_1 c_2 - b_2 c_1 = \frac{\varepsilon (b_1 \varphi_{v_1} + b_2 \varphi_{u_1})}{(u_1 + v_1) \sqrt{H}} = - \frac{\varepsilon (u+v)^2 \varphi_u}{(u_1 + v_1)^3 \sqrt{H}};$$

mithin wird:

$$(90) \quad \sum x_u X_1 = - \frac{\varepsilon \varphi_u}{(u-v) \sqrt{H_1}}.$$

Nach (89) und (90) ist in der Tat $\bar{\varepsilon} = \varepsilon$.

3. Eine für die geometrische Natur der Transformation B_k wesentliche Eigenschaft muß noch hergeleitet werden: *Zwischen der Biegungsfläche S und ihrer Bianchischen Transformierten S_1 besteht Zuordnung ihrer Asymptotenlinien.* Der hier zunächst gegebene neue Beweis benutzt die Wechselseitigkeit der Beziehungen, von der wir uns im vorigen Artikel überzeugt haben.

Wir quadrieren die Gleichung (74), addieren die folgende aus (70) gewonnene:

$$4 \varphi_u \varphi_v du dv - (\varphi_u du_1 + \varphi_v dv_1)^2 = - (\varphi_u du - \varphi_v dv)^2$$

und berücksichtigen, daß $\mathfrak{M}^2 = M^2 - LN$ ist. Nach einer leicht zu übersehenden Umformung ergibt sich:

$$\begin{aligned} & \varphi_u \varphi_v du dv - \varphi_{u_1} \varphi_{v_1} du_1 dv_1 \\ &= \frac{1}{4\mathfrak{M}^2} [L(\varphi_v)^2 + 2M\varphi_u \varphi_v + N(\varphi_u)^2] (L du^2 + 2M du dv + N dv^2). \end{aligned}$$

An die inverse Operation knüpft sich die analoge Formel:

$$\begin{aligned} & \varphi_{u_1} \varphi_{v_1} du_1 dv_1 - \varphi_u \varphi_v du dv \\ &= \frac{1}{4\mathfrak{M}_1^2} [L_1(\varphi_{v_1})^2 + 2M_1\varphi_{u_1} \varphi_{v_1} + N_1(\varphi_{u_1})^2] (L_1 du_1^2 + 2M_1 du_1 dv_1 + N_1 dv_1^2). \end{aligned}$$

Aus dem gleichzeitigen Bestehen dieser beiden Relationen geht aber hervor, daß sich die zweiten Fundamentalformen von S und S_1 nur um einen Faktor voneinander unterscheiden. Die Korrespondenz der Asymptotenlinien ist eine unmittelbare Folge davon. Da nach Art. 2 bereits feststeht

daß die Verbindungslinien entsprechender Punkte von S und S_1 nicht nur S , sondern auch S_1 berühren, erweist sich die von ihnen gebildete Strahlenkongruenz als ein W -System. Die Transformation B_k gehört damit der allgemeineren Klasse der *asymptotischen Flächentransformationen* an.

Die beim Beweise noch nicht berücksichtigte Annahme, daß für ein *partikuläres* Lösungspaar u_1, v_1 (denn um ein solches könnte es sich nur handeln) die in den eckigen Klammern stehenden Ausdrücke identisch verschwinden, kann auf Grund einer geometrischen Überlegung ausgeschlossen werden. Aus $L(\varphi_v)^2 + 2M\varphi_u\varphi_v + N(\varphi_u)^2 = 0$ würde nämlich $\frac{\varphi_u}{\varphi_v} = p$ bzw. $= q$ folgen, wobei hinsichtlich der Bedeutung von p und q auf § 2, 2 verwiesen sei. Dem Formelsystem (64) wäre dann zu entnehmen, daß die Kongruenz, deren Brennflächenmängel die beiden Flächen S und S_1 sind, aus den Tangenten der einen Asymptotenlinienschar von S besteht. Diese letztere Eigenschaft bedingt aber einem bekannten Satze zufolge das Zusammenfallen der beiden Brennflächenmängel.

4. Die nachfolgende ergänzende Betrachtung, bei der die Parameter α, β der Asymptotenlinien explicite eingeführt werden, enthält einen zweiten Beweis des soeben bestätigten Satzes und liefert eine übersichtliche Darstellung der Größen L_1, M_1, N_1 . Sie läßt überdies erkennen, wie man die in § 2, 2 erwähnte Äquivalenz zwischen der Verbiegung der Flächen zweiten Grades und dem speziellen Tschebyscheffschen Problem benutzen kann, um ohne die grundlegenden geometrischen Gesichtspunkte eine rein analytische (intrinsic) Theorie der Operation B_k zu entwickeln.

Wir gehen von den Beziehungen

$$(91) \quad \begin{cases} \frac{1}{\mathfrak{M}}(Mdu + Ndv) = u_\alpha d\alpha - v_\beta d\beta, \\ \frac{1}{\mathfrak{M}}(Ldu + Mdv) = -v_\alpha d\alpha + u_\beta d\beta \end{cases}$$

aus, die, da nach (44)

$$(92) \quad \frac{v_\beta}{u_\beta} = p, \quad \frac{v_\alpha}{u_\alpha} = q$$

ist, mit den Formeln (36) gleichwertig sind, bemerken aber, daß dabei die Parameter α, β der Asymptotenlinien nicht mit den durch die Quadraturen (40) gegebenen Größen α, β identisch zu sein brauchen, sondern daß diese durch eine willkürlich gewählte Funktion von α bzw. von β ersetzt werden können. Mit Hilfe von (91) bringen wir die im Verein mit der endlichen Relation $\varphi = 0$ geltenden Differentialgleichungen (75) der Transformation B_k auf die folgende einfache Form, wobei wir die beiden durch $\varepsilon = \pm 1$ gekennzeichneten Klassen trennen:

$$(93) \quad \left\{ \begin{array}{l} \text{(a) } \varepsilon = +1 \\ \text{(b) } \varepsilon = -1 \end{array} \right. \left\{ \begin{array}{l} \varphi_{u_1}(u_1)_\alpha + \varphi_u u_\alpha = 0, \quad \varphi_{v_1}(u_1)_\beta + \varphi_v v_\beta = 0, \\ \varphi_{v_1}(v_1)_\alpha + \varphi_v v_\alpha = 0, \quad \varphi_{v_1}(v_1)_\beta + \varphi_u u_\beta = 0, \\ \varphi_{u_1}(u_1)_\alpha + \varphi_v v_\alpha = 0, \quad \varphi_{u_1}(u_1)_\beta + \varphi_u u_\beta = 0, \\ \varphi_{v_1}(v_1)_\alpha + \varphi_u u_\alpha = 0, \quad \varphi_{v_1}(v_1)_\beta + \varphi_v v_\beta = 0. \end{array} \right.$$

Bezeichnen wir nun für einen Augenblick mit α_1, β_1 die Parameter der Asymptotenlinien von S_1 , so besteht infolge der Umkehrbarkeit der Operation das analoge Formelsystem:

$$(94) \quad (\varepsilon = +1) \quad \varphi_{u_1}(u_1)_{\alpha_1} + \varphi_u u_{\alpha_1} = 0 \text{ usw.}$$

Der aufgeschriebenen ersten Gleichung desselben geben wir, indem wir den Fall $\varepsilon = +1$ ins Auge fassen und α und β als Funktionen der beiden Variablen α_1 und β_1 betrachten, die folgende Gestalt:

$$[\varphi_{u_1}(u_1)_\alpha + \varphi_u u_\alpha] \alpha_{\alpha_1} + [\varphi_{u_1}(u_1)_\beta + \varphi_u u_\beta] \beta_{\alpha_1} = 0.$$

Hierin ist die erste Klammer nach (93a) gleich Null, die zweite dagegen ist von Null verschieden, da man andernfalls mit Hilfe der hinzutretenden Relation $\varphi_{u_1}(u_1)_\beta + \varphi_v v_\beta = 0$ zu der speziellen Annahme $\frac{\varphi_u}{\varphi_v} = \frac{v_\beta}{u_\beta} = p$ gelangen würde, die wir am Schluß des vorigen Artikels als unzulässig erkannt haben. Es muß demnach $\beta_{\alpha_1} = 0$, also $\beta = f(\beta_1)$ sein, so daß wir $\beta_1 = \beta$ und als Ergebnis einer entsprechenden Überlegung $\alpha_1 = \alpha$ schreiben können. Die Korrespondenz der Asymptotenlinien von S und S_1 ist damit abermals bestätigt. Das Formelsystem (94) fällt jetzt mit (93) zusammen.

Wir führen noch die zu S_1 gehörigen Größen p_1, q_1 ein. Analog zu (92) wird:

$$(95) \quad \frac{(v_1)_\beta}{(u_1)_\beta} = p_1, \quad \frac{(v_1)_\alpha}{(u_1)_\alpha} = q_1.$$

Aus den Differentialgleichungen (93) gewinnen wir im Verein mit (92) und (95) die bemerkenswerten Beziehungen:

$$\begin{aligned} (\varepsilon = +1) \quad p_1 &= \frac{\varphi_{u_1} \varphi_u}{\varphi_{v_1} \varphi_v} \cdot \frac{1}{p}, & q_1 &= \frac{\varphi_{u_1} \varphi_v}{\varphi_{v_1} \varphi_u} \cdot q, \\ (\varepsilon = -1) \quad p_1 &= \frac{\varphi_{u_1} \varphi_v}{\varphi_{v_1} \varphi_u} \cdot p, & q_1 &= \frac{\varphi_{u_1} \varphi_u}{\varphi_{v_1} \varphi_v} \cdot \frac{1}{q}, \end{aligned}$$

die sich zur Darstellung von L_1, M_1, N_1 benutzen lassen. Man hat dazu die Ausdrücke für p_1 und q_1 in die durchweg mit dem Index 1 versehenen Formeln (37) einzutragen und gleichzeitig die noch auftretende Größe \mathfrak{M}_1 der Formel (31) entsprechend zu bilden.

Es werde schließlich diejenige Auffassung der Operation B_k zur Sprache gebracht, die sich auf die Entwicklung am Ende des Art. 2 von

§ 2 stützt. Man kann die Frage nach einem transformatorischen Zusammenhang zwischen Lösungen des dort betrachteten Tschebyscheffschen Problems an die Spitze der Theorie stellen. Die neue Form (93), die wir den Differentialgleichungen der Transformation B_k gegeben haben, erscheint diesem Standpunkt angepaßt. Ist u, v das der Biegungsfläche S entsprechende Lösungspaar von

$$(96) \quad \varepsilon_1 u_\alpha v_\alpha = \varepsilon_2 u_\beta v_\beta = H \quad (\varepsilon_1 = \pm 1, \varepsilon_2 = \pm 1),$$

so erfolgt mit der Aufstellung des einen oder des anderen Systems (93) eine Zerspaltung der durch $\varphi = 0$ bedingten Relationen

$$\varphi_{u_1}(u_1)_\alpha + \varphi_{v_1}(v_1)_\alpha + \varphi_u u_\alpha + \varphi_v v_\alpha = 0,$$

$$\varphi_{u_1}(u_1)_\beta + \varphi_{v_1}(v_1)_\beta + \varphi_u u_\beta + \varphi_v v_\beta = 0,$$

die sowohl für $\varepsilon = +1$ wie für $\varepsilon = -1$ die Gleichungen

$$\varphi_{u_1} \varphi_{v_1} (u_1)_\alpha (v_1)_\alpha = \varphi_u \varphi_v u_\alpha v_\alpha, \quad \varphi_{u_1} \varphi_{v_1} (u_1)_\beta (v_1)_\beta = \varphi_u \varphi_v u_\beta v_\beta$$

unmittelbar nach sich zieht. Auf Grund von (96) und mit Benutzung der Formeln

$$\varphi_{u_1} \varphi_{v_1} = -kH, \quad \varphi_u \varphi_v = -kH_1$$

folgt man daraus:

$$\varepsilon_1 (u_1)_\alpha (v_1)_\alpha = \varepsilon_2 (u_1)_\beta (v_1)_\beta = H_1,$$

so daß man in u_1, v_1 ein neues Lösungspaar von (96) besitzt. Die transformierte Biegungsfläche S_1 ist damit allerdings nur *intrinsek*, d. h. durch ihre flächentheoretischen Fundamentalgrößen und ohne Rücksicht auf ihre Lage im Raume bestimmt. Wir begnügen uns mit diesen Andeutungen und verzichten insbesondere darauf, die für die betrachtete Form der Funktion φ durch die Rechnung § 3, 4 erledigte Frage nach der unbeschränkten Integrierbarkeit der Systeme (93) unter allgemeineren Voraussetzungen zu diskutieren.

§ 5.

Simultane asymptotische Transformationen der Achsenspurflächen.

1. Nachdem wir unter Vereinfachung des analytischen Apparats die Theorie der Bianchischen Transformation in ihren Grundzügen entwickelt und damit die Aufdeckung einer neuen geometrischen Eigenschaft vorbereitet haben, wenden wir uns jetzt wieder zu den in § 1, 5 eingeführten und in § 2, 4 dargestellten *Achsenspurflächen*, um ihr Verhalten der Operation B_k gegenüber zu untersuchen. Bezeichnen wir mit $S_1^{(1)}(x_1^{(1)}, \dots)$, $S_1^{(2)}(x_1^{(2)}, \dots)$, $S_1^{(3)}(x_1^{(3)}, \dots)$ die drei Achsenspurflächen, die sich zu der

transformierten Biegungsfläche S_1 konstruieren lassen, so gelten die von (52), (54) nur durch den Index 1 unterschiedenen Formeln:

$$(97) \quad x_1^{(1)} = x_1 - \frac{(u_1 + v_1)^2}{2\zeta_1} [(\zeta_1)_{v_1} (x_1)_{u_1} + (\zeta_1)_{u_1} (x_1)_{v_1}] \text{ usw.};$$

bei $S_1^{(2)}$ und $S_1^{(3)}$ tritt η_1 bzw. ζ_1 an die Stelle von ζ_1 . Es soll nun der folgende seinem geometrischen Inhalt nach außerordentlich einfache Satz bewiesen werden:

Die drei mit einer Biegungsfläche S des einschaligen Hyperboloids verbundenen Achsenspurflächen $S^{(1)}$, $S^{(2)}$, $S^{(3)}$ gehen bei Anwendung der Transformation B_k durch simultane asymptotische Transformationen in die drei entsprechenden zu der transformierten Biegungsfläche S_1 gehörigen Achsenspurflächen $S_1^{(1)}$, $S_1^{(2)}$, $S_1^{(3)}$ über.

Zunächst leuchtet unmittelbar ein, daß eine für die asymptotische Transformation wesentliche Bedingung erfüllt ist: auf den Flächen eines jeden der drei Paare entsprechen sich die Asymptotenlinien. In der Tat besteht eine derartige Verwandtschaft einmal nach § 1, 5 zwischen der Biegungsfläche und den ihr zugeordneten Achsenspurflächen, nach § 4, 3 aber außerdem zwischen den durch die Transformation B_k zusammenhängenden Biegungsflächen, so daß sie sich auf die Gesamtheit der acht betrachteten Flächen ausdehnt. Es handelt sich hiernach noch darum, zu bestätigen, daß die Flächen $S^{(1)}$, $S_1^{(1)}$, ebenso $S^{(2)}$, $S_1^{(2)}$ und $S^{(3)}$, $S_1^{(3)}$ paarweise die Brennflächenmäntel je einer Strahlenkongruenz bilden. Eine gewisse Schwierigkeit bietet dabei die ganz natürliche Forderung, die für $S^{(1)}$, $S_1^{(1)}$ durchzuführende Rechnung so zu gestalten, daß sie sich ohne weiteres auf die beiden anderen Paare überträgt. Wir sind dadurch genötigt, die Darstellung (52), (97) der Achsenspurflächen den durchsichtigeren Formeln (55) vorzuziehen. Eine durchaus zulässige Vereinfachung ergibt sich andererseits aus dem Umstande, daß wir uns in Anbetracht der in § 4, 2 erkannten Reziprozität aller bei der Transformation B_k gültigen Beziehungen darauf beschränken können, den Nachweis zu erbringen, daß die Verbindungslinien korrespondierender Punkte von $S^{(1)}$ und $S_1^{(1)}$ die eine der beiden Flächen, z. B. $S^{(1)}$ berühren. Als Bedingung dafür erhalten wir mit Berücksichtigung der Formeln (53), durch die für $S^{(1)}$ die Normalenrichtung gegeben ist, die folgende Gleichung, deren linke Seite wir abkürzend mit $\Theta^{(1)}$ bezeichnen:

$$(98) \quad \Theta^{(1)} = \sum (x_1^{(1)} - x^{(1)}) (\zeta_v x_u - \zeta_u x_v) = 0.$$

Dabei ist die dreigliedrige Summe unter Festhaltung des ζ über die lateinischen Buchstaben x, y, z zu erstrecken. In den analog gebildeten Summen $\Theta^{(2)}$, $\Theta^{(3)}$ steht η bzw. ζ an Stelle von ζ . Wir zeigen jetzt, daß (98) tatsächlich erfüllt ist.

2. Aus (97) folgt zunächst mit Hilfe von (52) und (64):

$$x_1^{(1)} - x^{(1)} = -\frac{(u_1 + v_1)^2}{2\xi_1} [(\xi_1)_{v_1}(x_1)_{u_1} + (\xi_1)_{u_1}(x_1)_{v_1}] + \frac{(u+v)^2}{2\xi} (\xi_v x_u + \xi_u x_v) - \frac{u+v}{2(u_1 + v_1)} (\varphi_v x_u + \varphi_u x_v).$$

Hierin ersetzen wir $(x_1)_{u_1}$ und $(x_1)_{v_1}$ durch die Ausdrücke (80), summieren in der durch (98) vorgeschriebenen Weise, wobei wir berücksichtigen, daß $\sum (\xi_u)^2 = E$ usw. ist, und erhalten für $\Theta^{(1)}$ die Formel:

$$(99) \quad \Theta^{(1)} = A^{(1)}(E \xi_v - F \xi_u) - B^{(1)}(G \xi_u - F \xi_v),$$

in der

$$(100) \quad \begin{cases} A^{(1)} = -\frac{(u_1 + v_1)^2}{2\xi_1} [a_2(\xi_1)_{u_1} + a_1(\xi_1)_{v_1}] + \frac{(u+v)^2 \xi_v}{2\xi} - \frac{(u+v) \varphi_v}{2(u_1 + v_1)}, \\ B^{(1)} = -\frac{(u_1 + v_1)^2}{2\xi_1} [b_2(\xi_1)_{u_1} + b_1(\xi_1)_{v_1}] + \frac{(u+v)^2 \xi_u}{2\xi} - \frac{(u+v) \varphi_u}{2(u_1 + v_1)} \end{cases}$$

ist. Wir drücken a_1, b_1, a_2, b_2 durch (83) aus und benutzen zu einer ersten Reduktion die nach (87) bestehende Beziehung:

$$(101) \quad \varphi_{v_1}(\xi_1)_{u_1} + \varphi_{u_1}(\xi_1)_{v_1} = -\frac{2(u+v)}{u_1 + v_1}(\xi' - \xi_1).$$

Es ergibt sich so, indem wir uns vorläufig auf $A^{(1)}$ beschränken:

$$(102) \quad A^{(1)} = \frac{(u+v)(u_1 + v_1)}{4\xi_1} \Phi^{(1)} + \frac{(u+v)^2}{4\xi_1} \frac{H_v}{H} (\xi' - \xi_1) + \frac{(u+v)^2}{2} \frac{\xi_v}{\xi},$$

wobei

$$(103) \quad \Phi^{(1)} = \varphi_{v_1 v}(\xi_1)_{u_1} + \varphi_{u_1 v}(\xi_1)_{v_1} - \frac{\varphi_v}{u_1 + v_1} \left[(\xi_1)_{u_1} + (\xi_1)_{v_1} + \frac{2\xi_1}{u_1 + v_1} \right]$$

ist.

Um nun $\Phi^{(1)}$ umzuformen, stellen wir uns diejenige *Identität* her, die für $\varphi = 0$ mit der Relation (101) zusammenfällt. Sie bietet wieder den Vorteil, daß sie sich ohne Rücksicht auf die Abhängigkeit der vier Argumente u, v, u_1, v_1 partiell differenzieren läßt. Wir gehen dazu von dem in ξ, η, ζ geltenden Ansatz

$$(104) \quad \xi' = \lambda^*(\xi_1)_{u_1} + \mu^*(\xi_1)_{v_1} + \nu^* \xi_1$$

aus. Zwecks Bestimmung der Koeffizienten λ^*, μ^*, ν^* werden diese drei Gleichungen einmal mit $\frac{\xi_1}{a^2}, \frac{\eta_1}{b^2}, -\frac{\zeta_1}{c^2}$, ein zweites Mal mit $\frac{(\xi_1)_{u_1}}{a^2}, \frac{(\eta_1)_{u_1}}{b^2}, -\frac{(\zeta_1)_{u_1}}{c^2}$ und ein drittes Mal mit $\frac{(\xi_1)_{v_1}}{a^2}, \frac{(\eta_1)_{v_1}}{b^2}, -\frac{(\zeta_1)_{v_1}}{c^2}$ multipliziert und danach addiert. Auf der rechten Seite sind die Gleichung des Hyperboloids und die daraus durch Differentiation hervorgehenden Beziehungen, insbesondere die für den Index 1 geschriebenen Formeln (51) anzuwenden,

links dagegen die aus (59) mit Berücksichtigung von (58) zu entnehmende Identität:

$$\frac{\xi' \xi_1}{a^2} + \frac{\eta' \eta_1}{b^2} - \frac{\delta' \delta_1}{c^2} = 1 + \frac{\varphi}{(u+v)(u_1+v_1)},$$

sowie die beiden weiteren, die aus ihr durch partielle Differentiation nach u_1 und v_1 folgen. Die auf Grund des Ansatzes (104) gewonnene Identität schreiben wir ähnlich wie (101) in der Form:

$$\varphi_{v_1}(\xi_1)_{u_1} + \varphi_{u_1}(\xi_1)_{v_1} - \frac{\varphi}{u_1+v_1} [(\xi_1)_{u_1} + (\xi_1)_{v_1} + \frac{2\xi_1}{u_1+v_1}] = -\frac{2(u+v)}{u_1+v_1}(\xi' - \xi_1). \quad (30)$$

Wir differenzieren dieselbe nach v , setzen nachträglich $\varphi = 0$, wodurch sich die Formel nicht weiter ändert, und erhalten, da links der Ausdruck (103) entstanden ist:

$$\Phi^{(1)} = -\frac{2}{u_1+v_1}[\xi' - \xi_1 - (u+v)\xi'_v].$$

Beim Eintragen dieses Wertes in (102) ersetzen wir noch ξ'_v durch $\frac{\xi'}{\xi}\xi_v$ und finden so:

$$A^{(1)} = -\frac{(u+v)^2(\xi' - \xi_1)}{2\xi_1} \left(\frac{\xi_v}{\xi} - \frac{1}{2} \frac{H_v}{H} + \frac{1}{u+v} \right).$$

Hier kann an die Stelle des letzten Faktors $\frac{\xi_v}{\xi}$ treten, wovon man sich überzeugt, indem man die logarithmische Abbildung von (29) bildet. Es wird demnach unter Hinzufügung des analogen Ausdrucks für $B^{(1)}$:

$$(105) \quad A^{(1)} = -\frac{(u+v)^2(\xi' - \xi_1)\xi_v}{2\xi_1\xi}, \quad B^{(1)} = -\frac{(u+v)^2(\xi' - \xi_1)\xi_u}{2\xi_1\xi}.$$

Man braucht nur die sog. Weingartenschen Gleichungen:

$$\xi_u = -\frac{\mathfrak{M}}{EG - F^2}(E\xi_v - F\xi_u), \quad \xi_v = -\frac{\mathfrak{M}}{EG - F^2}(G\xi_u - F\xi_v)$$

heranzuziehen, um durch Einsetzen der Werte (105) in (99) jetzt unmittelbar zu bestätigen, daß $\Theta^{(1)} = 0$ wird.

Aus der vorstehenden Rechnung, in der ein leicht zu übersehender Wechsel der Buchstaben einzutreten hätte, schließt man ohne weiteres auch auf das Bestehen der beiden anderen Gleichungen $\Theta^{(2)} = 0$ und $\Theta^{(3)} = 0$. Der am Anfang des Paragraphen ausgesprochene Satz ist damit bewiesen.

³⁰⁾ Der Ausdruck zwischen den eckigen Klammern hat den Wert a , wird in δ geschrieben $-c$ und in η gleich Null. In η ist also bereits die Gleichung (101) eine Identität.

On Families of Surfaces.

Von

C. E. Weatherburn in Christchurch (Neu-Seeland).

1. Introduction. Curvature.

This paper is concerned with various properties of a singly infinite family of surfaces in Euclidean space. Such a family may be defined by an equation of the form

$$(1) \quad \varphi(x, y, z) = \text{const.}$$

It is also determined if the unit vector \mathbf{n} normal to a surface is given as a point-function, satisfying the condition of normality

$$(2) \quad \mathbf{n} \cdot \text{rot } \mathbf{n} = 0.$$

For this vector is the unit tangent to a normal congruence of curves, which are the orthogonal trajectories of the family of surfaces. Conversely, if the equation of the family is given in the form (1), the unit normal may be taken as defined by

$$\mathbf{n} = \frac{\nabla \varphi}{|\nabla \varphi|}$$

where $\nabla \varphi$ denotes the gradient of the function φ .

Let s be the arc-length of an orthogonal trajectory of the family of surfaces, measured from a fixed point, and let ds be the intercept between consecutive surfaces φ and $\varphi + d\varphi$. Then if

$$ds = \psi d\varphi$$

we may call ψ the distance function for the family of surfaces. On any one surface the curves $\psi = \text{const.}$ are the *lines of equidistance*. Since the gradient of the function φ is defined by

$$\nabla \varphi = \mathbf{n} \frac{d\varphi}{ds}$$

it follows that

$$(3) \quad \mathbf{n} = \psi \nabla \varphi$$

and therefore

$$(4) \quad \psi^2 = \frac{1}{(\nabla \varphi)^2}.$$

The *first curvature* (or mean curvature), J , of a surface of the family is then given by¹⁾

$$\begin{aligned} J &= -\operatorname{div} \mathbf{n} = -\operatorname{div}(\psi \nabla \varphi) \\ &= -(\psi \nabla^2 \varphi + \nabla \varphi \cdot \nabla \psi) \end{aligned}$$

so that

$$(5) \quad J = -(\psi \nabla^2 \varphi + \mathbf{n} \cdot \nabla \log \psi).$$

As for the *second curvature* (or Gaussian curvature), K , we have shown elsewhere²⁾ that this is given by

$$(6) \quad 2K = \operatorname{div}(\mathbf{n} \operatorname{div} \mathbf{n} + \mathbf{n} \times \operatorname{rot} \mathbf{n}).$$

Now

$$(7) \quad \operatorname{rot} \mathbf{n} = \operatorname{rot} \psi \nabla \varphi = \nabla \psi \times \nabla \varphi = -\mathbf{n} \times \nabla \log \psi$$

and therefore

$$(8) \quad \mathbf{n} \times \operatorname{rot} \mathbf{n} = \nabla \log \psi - (\mathbf{n} \cdot \nabla \log \psi) \mathbf{n}.$$

Consequently, in virtue of (5), we have

$$(9) \quad \begin{aligned} \mathbf{n} \operatorname{div} \mathbf{n} + \mathbf{n} \times \operatorname{rot} \mathbf{n} &= (\psi^2 \nabla^2 \varphi) \nabla \varphi + \nabla \log \psi \\ &= \theta \nabla \varphi + \nabla \log \psi \end{aligned}$$

where we have written

$$\theta = \psi^2 \nabla^2 \varphi = \nabla^2 \varphi / (\nabla \varphi)^2.$$

Thus the second curvature, K , of a surface of the family is given by

$$(10) \quad 2K = \operatorname{div}(\theta \nabla \varphi + \nabla \log \psi).$$

2. Orthogonal trajectories. Lines of Equidistance.

Again it follows from (7) that the direction of $\operatorname{rot} \mathbf{n}$ is perpendicular to \mathbf{n} and also to $\nabla \psi$. It is thus tangential to the surface $\varphi = \text{const.}$, and also to the surface $\psi = \text{const.}$, and is therefore the direction of the line of equidistance through the point considered. Hence the theorem:

For a singly infinite family of surfaces, the direction of the line of equidistance through any point is that of the vector $\operatorname{rot} \mathbf{n}$.

¹⁾ See the author's *Differential Geometry*, p. 226 (Cambridge Press).

²⁾ *Loc. Cit.*, p. 261; also a paper "On Families of Curves and Surfaces", recently communicated to the *Quarterly Journal of Mathematics* 50, p. 352.

A family of *parallel surfaces* is characterised by the property³⁾ that $\text{rot } \mathbf{n}$ vanishes identically, and all curves on the surfaces are lines of equidistance.

Consider now the congruence of curves which are the *orthogonal trajectories* of the family of surfaces $\varphi = \text{const.}$ The unit tangent \mathbf{t} to the curve at any point is the unit normal \mathbf{n} to the surface through the point. Let κ be the curvature of the curve, s its arc-length, \mathbf{p} and \mathbf{b} the unit principal normal and binormal respectively. Since \mathbf{n} is a vector of constant length it follows that⁴⁾

$$(11) \quad -\mathbf{n} \times \text{rot } \mathbf{n} = \mathbf{n} \cdot \nabla \mathbf{n} = \frac{d\mathbf{t}}{ds} = \kappa \mathbf{p}$$

and therefore, in virtue of (2),

$$(12) \quad \text{rot } \mathbf{n} = \mathbf{n} \times (\kappa \mathbf{p}) = \kappa \mathbf{t} \times \mathbf{p} = \kappa \mathbf{b}$$

and we have the theorem:

The magnitude of $\text{rot } \mathbf{n}$ is the curvature of the orthogonal trajectory, and its direction is that of the binormal of this curve.

From the equation (8) it follows that $\mathbf{n} \times \text{rot } \mathbf{n}$ is the *two-parametric gradient*⁵⁾ of $\log \psi$ on the surface $\varphi = \text{const.}$ This may be denoted by $\nabla_2 \log \psi$. We have also just seen that the vector has the value $-\kappa \mathbf{p}$. Hence the formula

$$(13) \quad -\kappa \mathbf{p} = \nabla_2 \log \psi$$

which may be expressed:

The two-parametric gradient of the function $\log \psi$ on any surface of the family is the negative of the vector curvature⁶⁾ of the orthogonal trajectory through that point.

Moreover, the curves of equidistance on any surface will be *parallel curves* provided the magnitude of $\nabla_2 \psi$ is constant along each curve⁷⁾, that is to say, provided κ is constant along each curve. Hence the theorem:

A necessary and sufficient condition that the lines of equidistance on each surface may be parallels, and the lines of slope of the distance function therefore geodesics⁸⁾, is that the curvature of the orthogonal trajectories of the family of surfaces be constant along each line of equidistance.

³⁾ *Differential Geometry*, p. 261.

⁴⁾ See the author's *Advanced Vector Analysis*, Art. 8 (19).

⁵⁾ *Differential Geometry*, Art. 114; also a paper by the author "On Differential Invariants in Geometry of Surfaces etc.". *Quart. Journ. of Math.* 50 (1925), pp. 230 to 269.

⁶⁾ *Differential Geometry*, Art. 52.

⁷⁾ *Ibid.*, p. 224.

⁸⁾ *Ibid.*, p. 259; also "On Families of Curves and Surfaces", Art. 7.

3. Isometric System of Surfaces.

The condition that a family of surfaces $\varphi = \text{const.}$ may constitute an isometric system is usually stated in the form that $\nabla^2 \varphi / (\nabla \varphi)^2$ must be a function of φ only. It may, however, be very neatly expressed by a differential equation of the second order, to be satisfied by the unit normal \mathbf{n} , and *there is no need to assume that the family of surfaces is a Lamé family*, that is to say, forms part of a triply orthogonal system⁹). For, on taking the rotation of both members of (9), remembering that the rotation of the gradient vanishes identically, we have

$$\text{rot}(\mathbf{n} \text{ div } \mathbf{n} + \mathbf{n} \times \text{rot } \mathbf{n}) = \nabla \theta \times \nabla \varphi.$$

This vector will vanish if $\nabla \theta$ is parallel to $\nabla \varphi$, that is to say, if θ is a function of φ only. But this is the condition that the family of surfaces should be isometric. Hence the theorem:

An singly infinite family of surfaces with unit normal \mathbf{n} will constitute an isometric system provided

$$(14) \quad \text{rot}(\mathbf{n} \text{ div } \mathbf{n} + \mathbf{n} \times \text{rot } \mathbf{n}) = 0$$

which may also be expressed

$$(14') \quad \nabla \times (\mathbf{n} \nabla \cdot \mathbf{n} - \mathbf{n} \cdot \nabla \mathbf{n}) = 0.$$

When this relation is satisfied, θ being a function of φ only, we may write

$$F = \int \theta d\varphi.$$

The equation (9) is then equivalent to

$$\mathbf{n} \text{ div } \mathbf{n} + \mathbf{n} \times \text{rot } \mathbf{n} = \nabla (F + \log \psi)$$

so that, in virtue of (6), the Gaussian curvature of a surface $\varphi = \text{const.}$ has the value

$$(15) \quad K = \frac{1}{2} \nabla^2 (F + \log \psi).$$

This formula expresses the second curvature of an isometric family of surfaces as the Laplacian of a scalar function.

It has not been assumed in the above proof that an isometric family of surfaces is necessarily a Lamé family. We have shown elsewhere¹⁰) that *the condition that a family of surfaces may form part of a triply orthogonal system may be expressed*

$$\text{div} \left(\frac{\partial}{\partial \mathbf{n}} \text{rot } \mathbf{n} \right) = 0$$

⁹) See also a paper "On Lamé Families of Surfaces", recently communicated by the author to the *Annals of Mathematics* 28, p. 301.

¹⁰) *Ibid.*, Art. 2.

where $\frac{\partial}{\partial n}$ denotes differentiation in the direction of the normal. This relation is expressible in several other forms, such as

$$\operatorname{div}[(\operatorname{rot} \mathbf{n}) \cdot \nabla \mathbf{n}] = (\operatorname{rot} \mathbf{n}) \cdot \nabla \operatorname{div} \mathbf{n}$$

but the above condition (14) appears to be independent of it.

4. Gaussian Curvature of any System.

In conclusion let us investigate other formulæ for the *second curvature*, K , of a surface belonging to a singly infinite family with unit normal \mathbf{n} . We have shown elsewhere¹¹⁾ that, in terms of three parametric differential invariants for the space occupied by the family, the second curvature is expressible by

$$(16) \quad 2K = \mathbf{n} \cdot \nabla^2 \mathbf{n} + (\operatorname{div} \mathbf{n})^2 + (\operatorname{rot} \mathbf{n})^2.$$

Consider any set of fixed rectangular coordinate axes of x, y, z . Since \mathbf{n}^2 is equal to unity, its Laplacian is zero, so that

$$(17) \quad \mathbf{n} \cdot \nabla^2 \mathbf{n} + \left(\frac{\partial \mathbf{n}}{\partial x}\right)^2 + \left(\frac{\partial \mathbf{n}}{\partial y}\right)^2 + \left(\frac{\partial \mathbf{n}}{\partial z}\right)^2 = 0$$

Further, if $\mathbf{i}, \mathbf{j}, \mathbf{k}$ are unit vectors in the directions of the coordinate axes,

$$\operatorname{rot} \mathbf{n} = \mathbf{i} \times \frac{\partial \mathbf{n}}{\partial x} + \mathbf{j} \times \frac{\partial \mathbf{n}}{\partial y} + \mathbf{k} \times \frac{\partial \mathbf{n}}{\partial z}$$

and therefore, on squaring both members,

$$\begin{aligned} (\operatorname{rot} \mathbf{n})^2 &= \left(\frac{\partial \mathbf{n}}{\partial x}\right)^2 + \left(\frac{\partial \mathbf{n}}{\partial y}\right)^2 + \left(\frac{\partial \mathbf{n}}{\partial z}\right)^2 - (\operatorname{div} \mathbf{n})^2 \\ &\quad + 2 \Sigma \left(\mathbf{j} \cdot \frac{\partial \mathbf{n}}{\partial y} \mathbf{k} \cdot \frac{\partial \mathbf{n}}{\partial z} - \mathbf{j} \cdot \frac{\partial \mathbf{n}}{\partial z} \mathbf{k} \cdot \frac{\partial \mathbf{n}}{\partial y} \right). \end{aligned}$$

Hence, in virtue of (17),

$$2K = \mathbf{n} \cdot \nabla^2 \mathbf{n} + (\operatorname{div} \mathbf{n})^2 + (\operatorname{rot} \mathbf{n})^2 = 2 \Sigma (\mathbf{j} \times \mathbf{k}) \cdot \left(\frac{\partial \mathbf{n}}{\partial y} \times \frac{\partial \mathbf{n}}{\partial z} \right).$$

Thus, for any three mutually perpendicular directions, we have the formula

$$(18) \quad K = \Sigma \mathbf{i} \cdot \frac{\partial \mathbf{n}}{\partial y} \times \frac{\partial \mathbf{n}}{\partial z}.$$

This may be expressed in words as follows:

For any choice of three mutually perpendicular directions, the second curvature of a surface of the family is the sum of the resolved parts, in each direction, of the cross-product of the derivatives of \mathbf{n} in the other two, cyclic order being preserved.

¹¹⁾ "On Families of Curves and Surfaces", Art. 2; also *Differential Geometry*, p. 261.

The above formula for K may again be transformed as follows. In terms of the three-parametric operator ∇ we have¹²⁾

$$\nabla \mathbf{n} = \mathbf{i} \frac{\partial \mathbf{n}}{\partial x} + \mathbf{j} \frac{\partial \mathbf{n}}{\partial y} + \mathbf{k} \frac{\partial \mathbf{n}}{\partial z}$$

and the *second* of this dyadic has the value¹³⁾

$$(\nabla \mathbf{n})_2 = \Sigma \mathbf{j} \times \mathbf{k} \frac{\partial \mathbf{n}}{\partial y} \times \frac{\partial \mathbf{n}}{\partial z} = \Sigma \mathbf{i} \frac{\partial \mathbf{n}}{\partial y} \times \frac{\partial \mathbf{n}}{\partial z}.$$

The *scalar*¹⁴⁾ of this dyadic is the expression found above for K , or

$$(19) \quad K = (\nabla \mathbf{n})_{2s}.$$

Thus *the second curvature of a surface of the family is equal to the scalar of the second of $\nabla \mathbf{n}$.*

We have elsewhere proved this formula for a single surface¹⁵⁾, ∇ being then the two-parametric operator associated with the surface.

Lastly, if \mathbf{a} , \mathbf{b} are unit perpendicular vectors, tangential to orthogonal systems of curves on the surfaces, the formula

$$(20) \quad K = (\mathbf{a} \cdot \nabla \mathbf{a}) \cdot (\mathbf{b} \cdot \nabla \mathbf{b}) - (\mathbf{b} \cdot \nabla \mathbf{a}) \cdot (\mathbf{a} \cdot \nabla \mathbf{b})$$

which we have previously shown to hold for a single surface¹⁶⁾, ∇ being two-parametric, is true also for a family of surfaces, ∇ being three-parametric; for the various derivatives have the same values in both cases.

Canterbury College, Christchurch, New Zealand. November 1926.

¹²⁾ *Advanced Vector Analysis*, Art. 83.

¹³⁾ E. B. Wilson, *Vector Analysis*, Art. 118.

¹⁴⁾ The author's *Advanced Vector Analysis*, Art. 56.

¹⁵⁾ *Quarterly Journal of Mathematics* 50, p. 252.

¹⁶⁾ *Ibid.*, p. 251.

Beitrag zum Schwingungsproblem von Duffing.

Von

Kurt Lachmann in Berlin.

Der mathematische Kern des von Herrn Duffing¹⁾ behandelten Schwingungsproblems lautet in der von Herrn Hamel²⁾ gegebenen Form: Die periodischen Lösungen der Differentialgleichung

$$(1) \quad y'' + \alpha^2 \sin y = \beta \sin x$$

sind zu bestimmen, welche die Form

$$(2) \quad y = A_1 \sin x + A_3 \sin 3x + \dots$$

haben.

Eine strenge Lösung des Problems ist nicht bekannt. Unter den bisher vorliegenden Näherungsmethoden ist die folgende hervorzuheben. Setzt man angenähert:

$$y \approx A \sin x$$

und benutzt in der Fourier-Entwicklung für $\sin y$ nur das erste Glied, so erhält man für A eine transzendente Gleichung³⁾, mit deren Hilfe man A auf zeichnerischem oder rechnerischem Wege bestimmen kann.

Im folgenden soll ein Verfahren entwickelt werden, welches es gestattet, die periodischen Lösungen für gegebene Werte der Konstanten α^2 und β mit beliebiger Genauigkeit zu ermitteln.

Hierzu wird der nachstehende Weg eingeschlagen:

a) Man bestimmt ein Intervall, in dem z liegen muß, wenn die den Anfangsbedingungen

$$(3) \quad x = 0, \quad y = 0, \quad y' = z$$

genügende Lösung von (1) periodisch sein soll.

¹⁾ Georg Duffing, *Erzwungene Schwingungen bei veränderlicher Eigenfrequenz und ihre technische Bedeutung*. Sammlung Vieweg Heft 41/42. Braunschweig: Vieweg & Sohn. 1918.

²⁾ Georg Hamel, Über erzwungene Schwingungen bei endlicher Amplitude. *Math. Annalen* 86 (1922) Heft 1/2.

³⁾ Hamel, loc. cit. § 2.

b) Man sucht eine Entwicklung für die den Anfangsbedingungen (3) genügende Lösung von (1) nebst deren erster Ableitung, welche für die gemäß a) in Betracht kommenden Werte z bis $x = \frac{\pi}{2}$ gilt.

c) Evident findet man diejenigen Werte z , welche periodischen Lösungen entsprechen, aus der Bedingung

$$(4) \quad y' \left(\frac{\pi}{2} \right) = 0.$$

Man hat also in der gemäß b) gefundenen Entwicklung für $y'(x)$ $x = \frac{\pi}{2}$ zu setzen und zu untersuchen, für welche Werte z dieser Ausdruck verschwindet.

d) Setzt man diese Werte z in die nach b) ermittelte Reihe für $y(x)$ ein, so sind die periodischen Lösungen von (1) gefunden. Man wird schließlich die Koeffizienten der Entwicklung (2) bestimmen.

Die Methode kann stets angewendet werden, wo es sich um die Bestimmung der periodischen Lösungen einer zahlenmäßig gegebenen Differentialgleichung handelt.

§ 1.

Grenzen für z .

Setzt man

$$\alpha^2 + |\beta| = M,$$

so liegt y'' zwischen $+M$ und $-M$. Die Ableitungen der Lösungen von (1), welche den Anfangsbedingungen (3) genügen, liegen in dem von den Geraden $y = z + Mx$ und $y = z - Mx$ begrenzten Gebiete. Damit $y' \left(\frac{\pi}{2} \right)$ verschwinden kann, muß

$$|z| < M \cdot \frac{\pi}{2}$$

sein.

§ 2.

Die Potenzreihe.

Wir wollen die Lösung der Differentialgleichung (1), welche den Anfangsbedingungen (3) genügt, in eine Potenzreihe entwickeln.

Die Lösung hat die Form

$$(5) \quad y(x) = \frac{y_0'}{1!} x + \frac{y_0'''}{3!} x^3 + \frac{y_0^{(5)}}{5!} x^5 + \dots,$$

da gemäß (1) und (3) die geraden Ableitungen von y an der Stelle $x = 0$, $y = 0$ verschwinden.

Für die ungeraden Ableitungen wird $y'_0 = z$ und für $n > 0$

$$(6) \quad y_0^{(2n+1)} = (-1)^{n+1} \cdot \beta - \alpha^2 (\sin y_0)^{(2n-1)}.$$

Die in (6) auftretenden ungeraden Ableitungen von $\sin y$ an der Stelle $x = 0, y = 0$ findet man am besten dadurch, daß man mit Hilfe der Reihe (5) den Ausdruck $\sin y = y - \frac{y^3}{3!} + \frac{y^5}{5!} - + \dots$ bildet und nach Potenzen von x ordnet. Der Koeffizient von x^{2n-1} ist dann gleich $\frac{(\sin y_0)^{2n-1}}{(2n-1)!}$. Es wird

$$(7) \quad \begin{cases} y_0''' = \beta - \alpha^2 z; & y_0^{\text{v}} = -\beta - \alpha^2 \beta + \alpha^4 z + \alpha^2 z^3; \\ y_0^{\text{viii}} = (\beta + \alpha^2 \beta + \alpha^4 \beta) - \alpha^6 z + 10 \alpha^2 \beta z^2 - 11 \alpha^4 z^3 - \alpha^2 z^5; \\ y_0^{\text{ix}} = (-\beta - \alpha^2 \beta - \alpha^4 \beta - \alpha^6 \beta) + (\alpha^8 + 70 \alpha^2 \beta^2) z \\ \quad + (-21 \alpha^2 \beta - 171 \alpha^4 \beta) z^2 + 102 \alpha^6 z^3 - 35 \alpha^2 \beta z^4 + 57 \alpha^4 z^5 \\ \quad + \alpha^2 z^7 \text{ usw.} \end{cases}$$

Mit diesen Werten ist

$$(8) \quad y' \left(\frac{\pi}{2} \right) = \frac{y'_0}{1!} + \frac{y_0'''}{2!} \left(\frac{\pi}{2} \right)^2 + \frac{y_0^{\text{v}}}{4!} \left(\frac{\pi}{2} \right)^4 + \dots$$

Ist diese Reihe konvergent, so kann man nach einem der bekannten Näherungsverfahren diejenigen Werte z bestimmen, für welche $y' \left(\frac{\pi}{2} \right)$ verschwindet. Diese Werte wurden in (5) eingesetzt und liefern die Potenzreihenentwicklungen der gesuchten periodischen Lösungen. Wir werden zunächst untersuchen, wann die Reihe (5) und mit ihr die Reihe (8) bis $x = \frac{\pi}{2}$ konvergent ist (§ 3). In § 4 werden wir zeigen, wie man aus einer Potenzreihenentwicklung, welche bis $x = \frac{\pi}{2}$ konvergent ist und eine periodische Lösung darstellt, die Fourier-Koeffizienten der Entwicklung (2) ermittelt. Schließlich soll gezeigt werden, welches Verfahren man einschlagen kann, wenn die Konvergenz der Potenzreihe (5) für die in Betracht kommenden Werte z nicht bis $x = \frac{\pi}{2}$ nachgewiesen werden kann.

§ 3.

Konvergenzbereich der Potenzreihe.

Einen Bereich, in dem die Potenzreihe (5) sicher konvergiert, kann man auf Grund der folgenden Betrachtungen finden.

Setzt man die Lösung der Differentialgleichung

$$y'' = \alpha_0 + \alpha_1 x + \alpha_2 y + \alpha_{11} x^2 + 2 \alpha_{12} x y + \alpha_{22} y^2 + \dots,$$

welche den Anfangsbedingungen

$$x = 0; \quad y = y_0; \quad y' = y'_0$$

genügt, in der Form

$$y = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \dots$$

an, und weiß man, daß die Lösung der Differentialgleichung

$$y'' = \alpha_0^* + \alpha_1^* x + \alpha_2^* y + \alpha_{11}^* x^2 + 2\alpha_{12}^* x y + \alpha_{22}^* y^2 + \dots \alpha_{\nu\mu}^* > |\alpha_{\nu\mu}|,$$

welche den Anfangsbedingungen

$$x = 0; \quad y = y_0^*; \quad y' = y_0^{*'}; \quad y_0^* > |y_0|; \quad y_0^{*'} > |y_0'|$$

genügt, durch die für $|x| < R$ konvergente Potenzreihe

$$y = a_0^* + a_1^* x + a_2^* x^2 + a_3^* x^3 + \dots$$

dargestellt werden kann, so ist $|a_n| < a_n^*$, und die Entwicklung

$$y = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \dots$$

konvergiert mindestens für $|x| < R$. (Prinzip der Majorantenreihe.)

Zur Abschätzung des Konvergenzbereiches der gegebenen Differentialgleichung

$$y'' = \beta \sin x - \alpha^2 \sin y = \beta \left(x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots \right) \\ + \alpha^2 \left(y - \frac{y^3}{3!} + \frac{y^5}{5!} - \frac{y^7}{7!} + \dots \right)$$

kann man die Differentialgleichung

$$y'' = m^2 \operatorname{Sin}(x + y) = m^2 \left[\frac{x+y}{1!} + \frac{(x+y)^3}{3!} + \frac{(x+y)^5}{5!} + \frac{(x+y)^7}{7!} + \dots \right]$$

benutzen, wenn m^2 die größere der beiden Zahlen α^2 und $|\beta|$ ist. Als Anfangsbedingungen kann man $x = 0; \quad y = 0; \quad y' = |z|$ wählen.

Um die Differentialgleichung

$$y'' = m^2 \operatorname{Sin}(x + y)$$

zu lösen, setzt man

$$w = x + y; \quad w' = 1 + y'; \quad w'' = y''$$

und erhält

$$w'' = m^2 \operatorname{Sin} w;$$

die Anfangsbedingungen lauten

$$x = 0; \quad w = 0; \quad w' = 1 + |z|.$$

Um diese Differentialgleichung auf einen bekannten Fall zurückzuführen, sei

$$w = i v; \quad \operatorname{Sin} w = \operatorname{Sin}(i v) = i \sin v; \quad m = i \alpha.$$

Hieraus folgt

$$v'' + \alpha^2 \sin v = 0.$$

Das ist die bekannte Pendelgleichung. Für $x = 0$ ist $v = 0; \quad v' = v_0' = \frac{1 + |z|}{i}$.

Bei der Lösung hat man drei Fälle zu unterscheiden⁴⁾.

1. Ist

$$-1 < k = \frac{v'_0}{2\alpha} < 1,$$

so wird

$$\sin \frac{v}{2} = k \operatorname{sn}(\alpha x),$$

worin

$$\operatorname{sn} u = u - \frac{1+k^2}{6} u^3 + \frac{1+14k^2+k^4}{120} u^5 - + \dots$$

die von Jacobi eingeführte elliptische Funktion ist⁵⁾.

2. Ist

$$\frac{v'_0}{2\alpha} = k = 1,$$

so wird $\operatorname{sn} u = \operatorname{Sg} u$,⁶⁾ mithin

$$\operatorname{sn} \frac{v}{2} = \operatorname{Sg}(\alpha x).$$

3. Ist

$$-1 < k = \frac{2\alpha}{v'_0} < 1,$$

so wird

$$\sin \frac{v}{2} = \operatorname{sn}\left(\frac{v'_0}{2} \cdot x\right).$$

Wir setzen $w = iv$, $m = i\alpha$ und $v'_0 = \frac{1+|z|}{i}$ in die vorstehenden Formeln ein.

$$1. \quad k = \frac{v'_0}{2\alpha} = \frac{1+|z|}{i} \cdot \frac{i}{2m} = \frac{1+|z|}{2m}. \quad \text{Also muß}$$

$$k = \frac{1+|z|}{2m} < +1 \quad (m \text{ positiv})$$

sein. Für diesen Fall ist

$$\sin \frac{v}{2} = \sin \frac{w}{2} = -\sin\left(\frac{w}{2} \cdot i\right) = -i \operatorname{Sin} \frac{w}{2} = k \operatorname{sn}(-imx)$$

oder

$$\operatorname{Sin} \frac{w}{2} = ik \operatorname{sn}(-imx) = -ik \operatorname{sn}(imx).$$

Hieraus folgt

$$\operatorname{Cof} \frac{w}{2} = \sqrt{1 + \operatorname{Sin}^2 \frac{w}{2}} = \sqrt{1 - k^2 \operatorname{sn}^2(imx)} = \operatorname{dn}(imx).$$

⁴⁾ Z. B. Wilhelm Hort, Die Differentialgleichungen des Ingenieurs. 2. Auflage. S. 191. Berlin: Julius Springer 1925.

⁵⁾ Heinrich Burkhardt, Funktionentheoretische Vorlesungen. Zweiter Band. Elliptische Funktionen. S. 141. Berlin und Leipzig: Vereinigung wissenschaftlicher Verleger, Walter de Gruyter u. Co. 1920.

⁶⁾ Dr. Eugen Jahnke und Fritz Emde, Funktionentafeln mit Formeln und Kurven. Sammlung mathematisch-physikalischer Lehrbücher 5, S. 51. Leipzig-Berlin: B. G. Teubner 1923.

Differentiiert man die für $\text{Sin} \frac{w}{2}$ gefundene Gleichung, so wird

$$\text{Cos} \frac{w}{2} \cdot \frac{w'}{2} = \text{dn}(imx) \cdot \frac{w'}{2} = -ik \cdot \frac{d}{dx} \text{sn}(imx) = k \cdot m \cdot \text{cn}(imx) \cdot \text{dn}(imx)$$

oder

$$w' = 2km \text{cn}(imx).$$

Da $\text{cn} u$ nur die Pole $2\lambda K + (2\mu + 1)iK'$ als singuläre Stellen hat, worin λ und μ beliebige ganze Zahlen sind⁷⁾, so ist $\text{cn} u$ regulär, wenn $|u|$ kleiner als K' ist. Folglich konvergiert die Potenzreihenentwicklung für w , wenn $m|x| < K'$ ist. Da nun $y = w - x$ ist, so konvergiert auch die Majorantenreihe für $|x| < \frac{K'}{m}$, d. h. diese Reihe hat den Konvergenzradius $\frac{K'}{m}$. Hieraus folgt:

1. Ist $\frac{1+|z|}{m} < 1$, so konvergiert die Potenzreihe (5) sicher bis $x = \frac{\pi}{2}$, wenn $\frac{\pi}{2} < \frac{K'}{m}$ ist.

2. Ist $1 + |z| = m$, also $k = 1$, so wird $K' = \frac{\pi}{2}$ und daher muß m kleiner als 1 sein, wenn wir die Gewißheit haben wollen, daß unsere Potenzreihe wenigstens bis $x = \frac{\pi}{2}$ konvergent ist.

3. Es sei jetzt $\frac{2\alpha}{v'_0} = \frac{2m}{1+|z|} = k < 1$. Dann ist

$$\sin \frac{v}{2} = -i \text{Sin} \frac{w}{2} = \text{sn} \left(\frac{1+|z|}{2i} x \right),$$

oder

$$\text{Sin} \frac{w}{2} = i \text{sn} \left(\frac{1+|z|}{2i} x \right).$$

Ähnlich wie unter 1. folgt

$$\text{Cos} \frac{w}{2} = \sqrt{1 + \text{Sin}^2 \frac{w}{2}} = \sqrt{1 - \text{sn}^2 \left(\frac{1+|z|}{2i} x \right)} = \text{cn} \left(\frac{1+|z|}{2i} x \right),$$

$$\text{Cos} \frac{w}{2} \cdot \frac{w'}{2} = \text{cn} \left(\frac{1+|z|}{2i} x \right) \cdot \frac{w'}{2} = \frac{1+|z|}{2} \text{cn} \left(\frac{1+|z|}{2i} x \right) \text{dn} \left(\frac{1+|z|}{2i} x \right).$$

Es ist also

$$w' = (1 + |z|) \text{dn} \left(\frac{1+|z|}{2i} x \right).$$

Da nun $\text{dn} u$ dieselben Pole wie $\text{cn} u$ hat, so ist $\text{dn} u$ für $|u| < K'$ regulär und daher muß $\frac{1+|z|}{2} \cdot |x| < K'$ sein, wenn w als Funktion von x regulär sein soll. Hieraus ergibt sich: Ist $m < 1 + |z|$, so ist für $\frac{\pi}{2} < \frac{2K'}{1+|z|}$ der Konvergenzradius der Potenzreihe wenigstens gleich $\frac{\pi}{2}$.

⁷⁾ Burkhardt, loc. cit. S. 144.

Zusammenfassend erhält man:

Bezeichnet man die größere der beiden Zahlen α^2 und $|\beta|$ mit m^2 , so konvergiert die Potenzreihe (5) wenigstens bis $x = \frac{\pi}{2}$, wenn für

$$1 + |z| < 2m \quad \text{mit} \quad k = \frac{1 + |z|}{2m} \quad \frac{\pi}{2} < \frac{K'}{m}$$

$$1 + |z| = 2m \quad m < 1$$

$$1 + |z| > 2m \quad \text{mit} \quad k = \frac{2m}{1 + |z|} \quad \frac{\pi}{2} < \frac{2K'}{1 + |z|}$$

ist.

Um die Werte $|z|$ und m zu ermitteln, welche obigen Bedingungen genügen, wurde zunächst in Fig. 1 K' als Funktion von k dargestellt.

In Fig. 2 wurden als Abszissen die Werte m , als Ordinaten die Werte $|z|$ aufgetragen. Wählt man für k einen beliebigen Wert, der aber zwischen 0 und 1 liegen muß, so kann man die Gerade $k = \frac{1 + |z|}{2m}$ in das Ko-

ordinatensystem eintragen. Zu diesem Werte k entnimmt man der Fig. 1 den zugehörigen Wert K' .

Durch $\frac{\pi}{2} = \frac{K'}{m}$ oder $m = \frac{2K'}{\pi}$ wird auf der Geraden ein Punkt bestimmt, der folgende Eigenschaft hat:

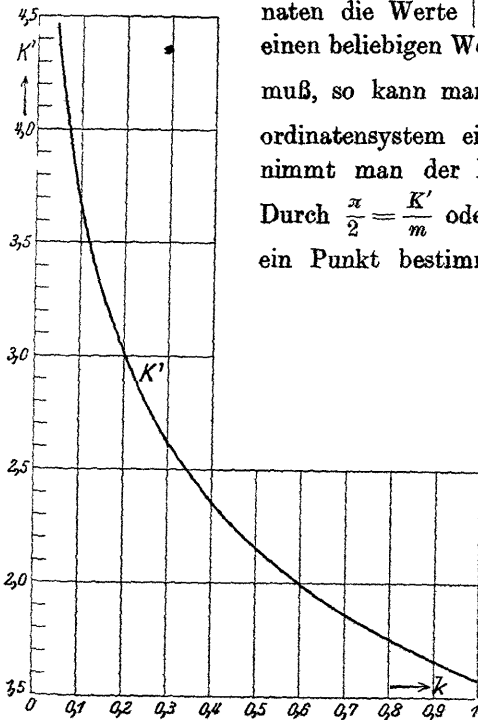


Fig. 1. K' als Funktion von k .

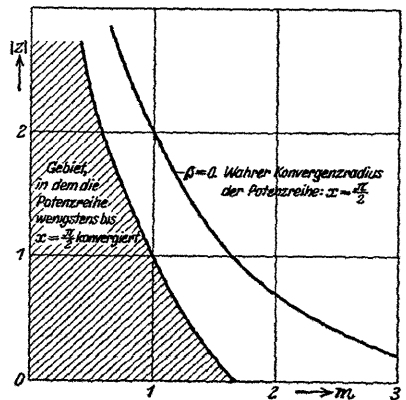


Fig. 2. Konvergenzbereich der Potenzreihe.

links von diesem Punkte erfüllen die den Koordinaten der Geraden entsprechenden Werte m und $|z|$ die Bedingung $\frac{\pi}{2} < \frac{K'}{m}$, rechts von diesem Punkte aber nicht mehr.

Ebenso kann man auch für einen festen Wert k , für den $0 < k < 1$

ist, die Gerade $k = \frac{2m}{1+|z|}$ zeichnen und denjenigen Punkt hervorheben, bis zu dem $\frac{\pi}{2} < \frac{2K'}{1+|z|}$ oder $|z| < \frac{4K'}{\pi} - 1$ ist.

Führt man die Konstruktion für verschiedene Werte k durch, so erhält man das in Fig. 2 schraffiert gezeichnete Gebiet, für welches die Potenzreihe wenigstens bis $x = \frac{\pi}{2}$ konvergiert. Sind also bestimmte Werte α^2 und β gegeben, so hat man nach § 1 den zugehörigen größten Wert $|z|$ zu ermitteln und wählt für m^2 die größere der beiden Zahlen α^2 und β . Liegt dann der Punkt mit den Koordinaten m und $|z|$ in dem schraffierten Gebiet, so sind wir sicher, daß die nach § 2 ermittelte Potenzreihe einen Konvergenzradius hat, der größer als $\frac{\pi}{2}$ ist.

Ist $\beta = 0$, also $y'' + \alpha^2 \sin y = 0$, so läßt sich die den Bedingungen $x = 0$, $y = 0$, $y' = z$ genügende Lösung nach den auf Seite 483 gegebenen Gleichungen in geschlossener Form angeben und daher der wahre Konvergenzradius der Potenzreihe berechnen. Die in Fig. 2 zum Vergleich eingetragene Kurve gibt die wahre Grenze des Gebietes an, für welches der Konvergenzradius für $\beta = 0$ größer oder gleich $\frac{\pi}{2}$ ist.

§ 4.

Bestimmung der Fourier-Koeffizienten.

Wir kommen jetzt zu dem folgenden Problem.

Gegeben ist die Potenzreihenentwicklung einer periodischen Lösung der Form

$$y = \frac{y_0'}{1!} x + \frac{y_0'''}{3!} x^3 + \dots + \frac{y_0^{(2n+1)}}{(2n+1)!} x^{2n+1} + R_{2n+1}(x),$$

wobei

$$|R_{2n+1}(x)| < R$$

ist.

Verlangt wird, die Fourier-Koeffizienten zu bestimmen, d. h. die Konstanten der Entwicklung

$$y = A_1 \sin x + A_3 \sin 3x + \dots$$

Daß diese Entwicklung stets möglich ist, folgt aus bekannten Sätzen; ferner gilt die Beziehung

$$A_k = \frac{1}{\pi} \int_0^{2\pi} y \sin kx \, dx = \frac{4}{\pi} \int_0^{\frac{\pi}{2}} y \sin kx \, dx.$$

Denkt man sich an Stelle von y die Reihe eingesetzt, so erkennt man, daß man die Integrale

$$\int_0^{\frac{\pi}{2}} x^n \sin kx \, dx \quad (k, n \text{ ungerade})$$

zu ermitteln hat.

Aus der Rekursionsformel

$$(-1)^{\frac{k-1}{2}} \int_0^{\frac{\pi}{2}} x^n \sin kx \, dx = \frac{n}{k^2} \left(\frac{\pi}{2}\right)^{n-1} - (-1)^{\frac{k-1}{2}} \frac{n(n-1)}{k^2} \int_0^{\frac{\pi}{2}} x^{n-2} \sin kx \, dx$$

und

$$(-1)^{\frac{k-1}{2}} \int_0^{\frac{\pi}{2}} x \sin kx \, dx = \frac{1}{k^2}$$

folgt

$$\begin{aligned} (-1)^{\frac{k-1}{2}} \int_0^{\frac{\pi}{2}} x^n \sin kx \, dx &= \frac{(-1)^{\frac{n-1}{2}} \cdot n!}{k^{n+1}} \left[1 - \frac{1}{2!} \left(\frac{\pi k}{2}\right)^2 + \frac{1}{4!} \left(\frac{\pi k}{2}\right)^4 \right. \\ &\quad \left. - \frac{1}{6!} \left(\frac{\pi k}{2}\right)^6 \dots - \frac{(-1)^{\frac{n-1}{2}}}{(n-3)!} \left(\frac{\pi k}{2}\right)^{n-3} + \frac{(-1)^{\frac{n-1}{2}}}{(n-1)!} \left(\frac{\pi k}{2}\right)^{n-1} \right]. \end{aligned}$$

Setzt man die eckige Klammer gleich $k^{n+1} \cdot F(k, n)$, also

$$(-1)^{\frac{k-1}{2}} \int_0^{\frac{\pi}{2}} x^n \sin kx \, dx = (-1)^{\frac{n-1}{2}} \cdot n! F(k, n),$$

so wird

$$\begin{aligned} (-1)^{\frac{k-1}{2}} A_k &= (-1)^{\frac{k-1}{2}} \cdot \frac{4}{\pi} \int_0^{\frac{\pi}{2}} \left[\frac{y_0'}{1!} x + \frac{y_0'''}{3!} x^3 + \dots \right] \sin kx \, dx \\ &= \frac{4}{\pi} \left[F(k, 1) y_0' - F(k, 3) y_0''' + F(k, 5) y_0^v + \dots \right. \\ &\quad \left. + (-1)^n F(k, 2n+1) y_0^{(2n+1)} \right] + M_{k, 2n+1}(x), \end{aligned}$$

und ersichtlich ist

$$|M_{k, 2n+1}(x)| = \frac{4}{\pi} \left| \int_0^{\frac{\pi}{2}} R_{2n+1}(x) \sin kx \, dx \right| < \frac{4}{\pi} R \int_0^{\frac{\pi}{2}} \sin kx \, dx = \frac{4R}{k\pi},$$

d. h. der Fehler, den man begeht, wenn man das Restglied M vernachlässigt, ist höchstens gleich dem 1, 3-fachen Betrage des absolut größten Restes der Potenzreihe.

Für $k = 1$ ist

$$F(1, n) = 1 - \frac{1}{2!} \left(\frac{\pi}{2}\right)^2 + \frac{1}{4!} \left(\frac{\pi}{2}\right)^4 \dots + \frac{(-1)^{\frac{n-1}{2}}}{(n-1)!} \left(\frac{\pi}{2}\right)^{n-1},$$

also

$$F(1, 1) = 1, \quad F(1, 3) = -0,233701, \quad F(1, 5) = 0,019969,$$

$$F(1, 7) = -0,000895, \quad F(1, 9) = 0,000025, \quad F(1, 11) = -0,000001.$$

Für $k = 3$ ergibt sich

$$F(3, n) = \frac{1}{3^{n+1}} \left[1 - \frac{1}{2!} \left(\frac{3\pi}{2}\right)^2 + \frac{1}{4!} \left(\frac{3\pi}{2}\right)^4 \dots + \frac{(-1)^{\frac{n-1}{2}}}{(n-1)!} \left(\frac{3\pi}{2}\right)^{n-1} \right],$$

$$F(3, 1) = 0,111111, \quad F(3, 3) = -0,124732, \quad F(3, 5) = 0,014326,$$

$$F(3, 7) = -0,000726, \quad F(3, 9) = 0,000021, \quad F(3, 11) = -0,000000.$$

Mit diesen Werten wird

$$A_1 = \frac{4}{\pi} (y'_0 + 0,233701 y'''_0 + 0,019969 y^V_0 + \dots),$$

$$A_3 = -\frac{4}{\pi} (0,111111 y'_0 + 0,124723 y'''_0 + 0,014326 y^V_0 + \dots).$$

Ebenso können auch die übrigen Koeffizienten ermittelt werden.

§ 5.

Die Konvergenz der Potenzreihe kann nicht bis $x = \frac{\pi}{2}$ nachgewiesen werden.

Sind α^2 oder β so groß, daß die Konvergenz der Potenzreihe (5) für die in Betracht kommenden Werte z nicht bis $x = \frac{\pi}{2}$ nachgewiesen werden kann, so wendet man zweckmäßig das Verfahren der analytischen Fortsetzung an.

Nach den in § 3 gegebenen Formeln sei die Konvergenz die Potenzreihenentwicklung bis $x = x_1$ sichergestellt. Für $x = x_1$ sei $y = y_1$ und $y' = y'_1$, und die weiteren Ableitungen können mit Hilfe von $y''_1 = \beta \sin x_1 - \alpha^2 \sin y_1$; $y'''_1 = \beta_1 \cos x_1 - \alpha^2 \cos y_1 \cdot y'_1$ usw. zahlenmäßig ermittelt werden.

Setzt man nun:

$$x - x_1 = t,$$

so kann man für $y(t)$ und $y'(t)$ die Reihen

$$y(t) = y_1 + \frac{y_1'}{1!}t + \frac{y_1''}{2!}t^2 + \dots,$$

$$y'(t) = y_1' + \frac{y_1''}{1!}t^2 + \dots$$

bilden. Konvergieren diese Reihen für $|t| < x_2$, so rechnet man sich für $x = x_1 + x_2$ die Werte $y = y_2$ und $y' = y_2'$ aus, wendet dasselbe Verfahren an und fährt so fort, bis $x = \frac{\pi}{2}$ erreicht ist.

Daß das Verfahren der analytischen Fortsetzung stets nach einer endlichen Anzahl von Schritten zum Ziel führt, folgt aus dem Umstande, daß auf der reellen x -Achse keine singulären Stellen von y liegen.

Wir wollen eine obere Grenze für diese Anzahl und gleichzeitig eine untere Grenze für den Konvergenzbereich der Potenzreihenentwicklung an der Stelle

$$x = x_n; \quad y = y_n; \quad y' = y_n' \quad \left(0 < x_n < \frac{\pi}{2}\right)$$

angeben. Zu bemerken ist, daß nach § 1 die Werte y_n' zwischen $z + Mx_n$ und $z - Mx_n$ liegen, also beschränkt sind.

Man könnte die gegebene Differentialgleichung, wie in § 3, mit $y'' = m^2 \sin(x + y)$ vergleichen, doch wird die Rechnung bedeutend einfacher, wenn man zum Vergleich die Differentialgleichung

$$y'' = m^2 e^{x+y} = m^2 \left[1 + \frac{x+y}{1!} + \frac{(x+y)^2}{2!} + \dots\right]$$

heranzieht, wobei natürlich die zu ermittelnden unteren Grenzen der Konvergenzradien bedeutend kleiner ausfallen.

Man setzt

$$x - x_n = t; \quad w = x + y = x_n + t + y$$

und erhält die Differentialgleichung

$$w'' = m^2 e^w,$$

welche für die Anfangsbedingungen

$$t = 0; \quad w = x_n + |y_n| = w_0; \quad w' = 1 + |y_n'| = w_0'$$

zu untersuchen ist.

Durch Multiplikation mit $2w'$ wird

$$2w' \cdot w'' = 2m^2 e^w w'.$$

Die Integration dieser Gleichung liefert

$$w'^2 = 2m^2 \cdot e^w - c_1 = 2w'' - c_1.$$

Setzt man

$$w' = v; \quad w'' = v',$$

so folgt

$$v^2 = 2v' - c_1,$$

$$v^2 + c_1 = 2 \cdot \frac{dv}{dt} \quad \text{oder} \quad \frac{t}{2} = \int \frac{dv}{v^2 + c_1}.$$

Folgende Fälle sind zu unterscheiden.

1. $c_1 = C_1^2 > 0$. Es ist

$$\frac{t}{2} = \int \frac{dv}{v^2 + C_1^2} = \frac{1}{C_1} \arctg \frac{v}{C_1} + C_2,$$

$$C_1 \left(\frac{t}{2} - C_2 \right) = \arctg \frac{v}{C_1}, \quad v = C_1 \operatorname{tg} C_1 \left(\frac{t}{2} - C_2 \right).$$

Durch Differentiation erhält man

$$v' = \frac{C_1^2}{2 \cos^2 C_1 \left(\frac{t}{2} - C_2 \right)}.$$

v' und daher auch v bleiben regulär, solange der Nenner nicht gleich Null, d. h.

$$C_1 \left(\frac{t}{2} - C_2 \right) < \frac{\pi}{2}$$

oder

$$|t| < \frac{\pi}{C_1} + 2C_2$$

ist. $\frac{\pi}{C_1} + 2C_2$ ist also eine untere Grenze für den Konvergenzradius der Potenzreihen-Entwicklung unserer Reihe.

Die Integrationskonstanten bestimmen sich aus den Anfangsbedingungen; für $t = 0$ muß $v = w'_0$ und $v' = w'' = m^2 e^{w_0}$ sein. Mithin ist

$$w'_0 = C_1 \operatorname{tg} C_1 C_2,$$

$$m^2 e^{w_0} = \frac{C_1^2}{2 \cos^2(C_1 C_2)} = \frac{C_1^2}{2} (1 + \operatorname{tg}^2(C_1 C_2)) = \frac{C_1^2}{2} \left(1 + \frac{w_0'^2}{C_1^2} \right) = \frac{C_1^2}{2} + \frac{w_0'^2}{2}.$$

Ist also

$$C_1^2 = 2m^2 e^{w_0} - w_0'^2 > 0$$

und bestimmt man C_2 aus

$$w'_0 = C_1 \operatorname{tg} C_1 C_2,$$

so gilt als untere Grenze für den Konvergenzradius

$$\frac{\pi}{C_1} + 2C_2.$$

2. Ist $c_1 = 0$, so wird

$$\frac{t}{2} = \int \frac{dv}{v^2} = -\frac{1}{v} + C_2, \quad \frac{1}{v} = C_2 - \frac{t}{2}$$

oder

$$v = \frac{1}{C_2 - \frac{t}{2}} \quad \text{und daher} \quad v' = \frac{1}{2 \left(C_2 - \frac{t}{2} \right)^2}.$$

Für $t = 0$ wird $v = w_0' = \frac{1}{C_2}$ und $v' = m^2 e^{w_0} = \frac{1}{2 C_2^2}$.

Es muß also

$$\frac{1}{2} w_0'^2 = \frac{1}{2 C_2^2} = m^2 e^{w_0}$$

oder

$$2 m^2 e^{w_0} - w_0'^2 = 0$$

sein; ferner ist $C_2 = \frac{1}{w_0'}$.

Ersichtlich muß hier

$$\left| \frac{t}{2} \right| < C_2 \quad \text{oder} \quad |t| < \frac{2}{w_0'}$$

sein. Als untere Grenze für den Konvergenzradius erhält man, wenn

$$2 m^2 e^{w_0} - w_0'^2 = 0$$

ist,

$$\frac{2}{w_0'}.$$

3. Ist $c_1 = -C_1^2 < 0$, so wird

$$\frac{t}{2} = \int \frac{dv}{v^2 - C_1^2} = \frac{1}{2 C_1} \ln \frac{v - C_1}{v + C_1} + C_2$$

oder

$$\frac{v - C_1}{v + C_1} = e^{C_1 t - 2 C_1 C_2}, \quad v = -C_1 + \frac{2 C_1}{1 - e^{C_1 t - 2 C_1 C_2}}.$$

Aus $v^2 + c_1 = 2 v'$ folgt für $t = 0$ mit $c_1 = -C_1^2$

$$w_0'^2 - C_1^2 = 2 m^2 e^{w_0}$$

oder

$$C_1^2 = w_0'^2 - 2 m^2 e^{w_0} > 0.$$

Ferner ist für $t = 0$

$$C_2 = -\frac{1}{2 C_1} \ln \frac{w_0' - C_1}{w_0' + C_1}.$$

Als Grenze des Konvergenzradius ergibt sich aus

$$|C_1 t - 2 C_1 C_2| < 0,$$

wenn wir C_1 und C_2 positiv wählen,

$$2 C_2.$$

Für gegebene Werte x_n , y_n und y'_n hat man zunächst w_0 und w'_0 zu berechnen und dann mit Hilfe von $2m^2 e^{w_0} - w'_0{}^2 \geq 0$ zu untersuchen, welcher der drei Fälle vorliegt. Schließlich berechnet man C_1 und C_2 und die untere Grenze für den Konvergenzradius.

Führt man dieselben Betrachtungen durch, indem man für y_n und y'_n die absolut größten Werte, welche gemäß § 1 möglich sind, einsetzt, nämlich

$$y'_n = M\pi, \quad y_n = \frac{3}{8} M\pi^2, \quad M = \alpha^2 + |\beta|,$$

und

$$x_n = \frac{\pi}{2}$$

wählt, so erhält man eine untere Grenze für die Konvergenzradien *aller* Entwicklungen und kann hiernach berechnen, wieviel verschiedene Entwicklungen man höchstens nötig hat, um bis $x = \frac{\pi}{2}$ zu gelangen.

Die Bestimmung der Fourier-Koeffizienten hat so zu erfolgen, daß man die Integrale über die einzelnen Teilintervalle erstreckt und die für y geltenden Potenzreihenentwicklungen einsetzt.

Ersichtlich wird die Bestimmung der periodischen Lösungen sehr schwierig, wenn man nicht mit *einer* Potenzreihenentwicklung auskommt. Man wird daher nicht mit dem unbekanntem Werte z arbeiten, sondern die angenäherte Richtung der Anfangstangente der periodischen Lösung entweder mit Hilfe der Näherungslösung und nach einem zeichnerischen Verfahren bestimmen, für z bestimmte Werte annehmen und diese so lange verändern, bis für das letzte Intervall $y' \left(\frac{\pi}{2} \right)$ mit genügender Genauigkeit verschwindet. Die Bestimmung der Fourier-Koeffizienten ist dann verhältnismäßig einfach.

Berichtigung und Bemerkungen zur Arbeit „Über p -Relationen“.

Von

R. Weitzenböck in Amsterdam.

In der genannten Arbeit (Math. Annalen 97 (1927), S. 788—795) glaubte ich für den Satz, daß eine p -Relation $G(p) = 0$ stets auf die Gestalt $G(p) = \sum AII = \sum AII(p, p) = 0$ zu bringen ist, einen von Reihenentwicklungen¹⁾ unabhängigen Beweis gegeben zu haben.

Wie sich nach einem Briefwechsel mit Herrn E. Fischer herausstellte, führen die Überlegungen des § 4 der genannten Arbeit $G(p)$ wohl formal in die gewünschte Gestalt $\sum AII$ über, beweisen aber den Satz nicht, weil bei dieser Umwandlung von $G(p)$ die Koordinatenreihen x, y, \dots, z aus $p_{ik\dots s} = (xy\dots z)_{ik\dots s}$ gebraucht werden.

Gleichfalls von Herrn Fischer wurde ich auf einen Beweis des Satzes durch F. Mertens²⁾ hingewiesen. Dieser gedanklich sehr einfache, von Reihenentwicklungen unabhängige Beweis besitzt nur den Übelstand, daß er eine zwölf Seiten lange Indizesrechnung benötigt. Ich halte es deshalb nicht für überflüssig, diesen Mertensschen Beweis in der viel kürzeren Form wiederzugeben, die die symbolische Methode der Invariantentheorie an die Hand gibt, wobei ich mich möglichst den Mertensschen Bezeichnungen anschließe.

¹⁾ Der in meiner „Invariantentheorie“ (1923), S. 115, gegebene Beweis des Theorems stützt sich auf den zweiten Fundamentalsatz der symbolischen Methode, der seinerseits wieder durch Reihenentwicklungen gewonnen wird. Vgl. hierüber auch B. van der Waerden, Math. Annalen 95 (1926), S. 706—735, wo das Theorem als Spezialfall des Satzes II von S. 725 auftritt.

²⁾ F. Mertens, Über ganze Funktionen von m Systemen und je n Unbestimmten, Monatsh. f. Math. u. Physik 4 (1893), S. 193—228 und S. 297—329. In Betracht kommen S. 316—329.

Es sei $G = G(p) = G(p_{ik\dots s})$ eine Form l -ten Grades der G_d -Koordinaten $p_{ik\dots s} = (xy\dots z)_{ik\dots s}$, und $G'(t)$ bezeichne dieselbe Form, gebildet mit unabhängigen, zu den $p_{ik\dots s}$ kogredienten Veränderlichen $t_{ik\dots s}$.

Wir berechnen den Ausdruck $T(G(p))$, wo T den Prozeß

$$T = \sum_{(ik\dots s)} t_{ik\dots s} \frac{\partial^m}{\partial x_i \partial y_k \dots \partial z_s} = \frac{1}{d!} \sum_{(ik\dots s)} t_{ik\dots s} (\Omega_{xy\dots z})_{ik\dots s}$$

bedeutet. Man erhält bei $l > 1$ die Formel ($c =$ rationaler Zahlenkoeffizient):

$$(1) \quad TG(p) = c_0 \left(t \frac{\partial}{p} \right) G(p) + \sum c' \Pi(t, p) G'.$$

Hier ist der erste Term rechts die Polare $\sum t_{ik\dots s} \frac{\partial G}{\partial p_{ik\dots s}}$ und in \sum hat jedes Glied einen Faktor

$$\Pi(t, p) = t_{ik\dots s} p_{\lambda\mu\dots\nu} - t_{\lambda k\dots s} p_{i\mu\dots\nu} + \dots + (-1)^d t_{i\lambda k\dots s} p_{s\mu\dots\nu},$$

[$\Pi(p, p) = 0$ ist eine quadratische p -Relation]. G' ist eine Form $(l-2)$ -ten Grades in den $p_{ik\dots s}$. Bei $l=1$ steht in (1) rechts nur die Polare allein.

Bevor wir (1) beweisen, sei der Beweis des Theorems über p -Relationen nach Mertens zu Ende geführt. $TG(p)$ ist wieder eine Form der p , wir können also zur Berechnung von $T(TG(p)) = T^2 G$ wieder (1) anwenden; analog bei $T^3 G$ usw. Man erhält nach l Schritten:

$$(2) \quad T^l G(p) = c_0' \left(t \frac{\partial}{p} \right)^l G(p) + \sum c'' \Pi(t, t) \cdot G''.$$

Hier ist der erste Term rechts $= cG(t)$ mit $c \neq 0$, und die G'' sind Formen der $t_{ik\dots s}$. Wegen $G(p) = G((xy\dots z)_{ik\dots s}) = 0$ haben wir also:

$$(3) \quad G(t) = -\frac{1}{c} \sum c'' \Pi(t, t) \cdot G''(t), \quad \text{w. z. b. w.}$$

Zum Beweise von (1) setzen wir

$$(4) \quad G(p) = (A' p)^d (B' q)^d \dots (L' r)^d \\ = (d!)^l (A' x)(A' y) \dots (A' z) \cdot (B' x)(B' y) \dots (B' z) \cdot \dots \cdot (L' x)(L' y) \dots (L' z).$$

Hier sind die p d -fältige Komplexsymbole; q, r, \dots sind mit p äquivalent. Ebenso sind die A' d -fältige Komplexsymbole und B', \dots, L' sind mit A' äquivalent.

Die Wirkung von T auf G ergibt sich dann aus der Formel

$$T((a' x)(\beta' y) \dots (\gamma' z)) = \sum_{(ik\dots s)} t_{ik\dots s} \alpha'_i \beta'_k \dots \gamma'_s \\ = \frac{1}{d!} \sum t_{ik\dots s} (\alpha' \beta' \dots \gamma')_{ik\dots s} = \frac{1}{d!} (\alpha' t)(\beta' t) \dots (\gamma' t),$$

d. h. man erhält $T((a' x)(\beta' y) \dots (\gamma' z))$ bis auf einen Zahlenfaktor, wenn man x, y, \dots, z durch t ersetzt.

Demgemäß wird $T(G)$ nach (4):

$$(5) \quad T(G(p)) = \sum c(A't)^\alpha (B't)^\beta \dots (D't)^\delta (A'\xi)(A'\eta) \dots \\ \dots (B'\varrho)(B'\sigma) \dots (D'\tau)(E'p)^d \dots (L'r)^d \\ (\alpha + \beta + \dots + \delta = d).$$

Die Reihen $\xi, \eta, \dots, \varrho, \sigma, \dots, \tau$ sind hier: $d - 1$ Reihen x , $d - 1$ Reihen y, \dots , $d - 1$ Reihen z . Die Summe ist zu erstrecken über alle Ausdrücke, die sich ergeben, wenn man in (4) auf alle möglichen Weisen je eine Reihe x, y, \dots, z durch t ersetzt. Die Zahlen $\alpha, \beta, \dots, \delta$ haben die Werte $0, 1, 2, \dots, d$.

In (5) lassen sich die Reihen ξ, η, \dots wieder zu $p_{ik\dots s} = (xy\dots z)_{ik\dots s}$ zusammenfassen. Man erhält³⁾, abgesehen von Faktoren $(E'p)^d \dots (L'r)^d$:

$$TG(p) = \sum_{\alpha, \beta, \dots, \delta} c \{ (A't)^\alpha (B't)^\beta \dots (D't)^\delta \sum_{\lambda, \mu, \dots, \varrho} c' (A'^{\lambda_1} B'^{\mu_1} \dots D'^{e_1})_{(xy\dots z)} \dots \\ (A'^{\lambda_{k-1}} B'^{\mu_{k-1}} \dots D'^{e_{k-1}})_{(xy\dots z)} \},$$

$$(6) \quad \begin{cases} \lambda_1 + \lambda_2 + \dots + \lambda_{k-1} = d - \alpha & 0 \leq \lambda_i \leq d - \alpha \\ \mu_1 + \mu_2 + \dots + \mu_{k-1} = d - \beta & 0 \leq \mu_i \leq d - \beta & \lambda_i + \mu_i + \dots + \varrho_i = d \\ \dots & \dots & \alpha + \beta + \dots + \delta = d. \\ \varrho_1 + \varrho_2 + \dots + \varrho_{k-1} = d - \delta & 0 \leq \varrho_i \leq d - \delta \end{cases}$$

Hier kann $(A't)^\alpha (B't)^\beta \dots (D't)^\delta$ durch den Klammerfaktor $(A'^\alpha B'^\beta \dots D'^\delta t^{n-d})$ und jedes $(A'^{\lambda_i} B'^{\mu_i} \dots D'^{e_i})_{(xy\dots z)}$ durch $(A'p)^{\lambda_i} \times (B'p)^{\mu_i} \dots (D'p)^{e_i}$ ersetzt werden. $T(G)$ geht also über in eine Summe von Termen

$$(7) \quad (A'^\alpha B'^\beta \dots D'^\delta t^{n-d}) (A'p) \dots (B'q) \dots (E'r)^d \dots (L's)^d.$$

Hier bringen wir alle $d - \alpha$ Reihen A' von $(A'p) \dots$ in den Klammerfaktor hinein⁴⁾. Dadurch entstehen Ausdrücke W_2 mit symbolischen Faktoren $(t'p)$ und, wenn keine Reihe t' aus dem Klammerfaktor herausgetreten ist, Ausdrücke W_1 mit $(A'^d t'^{n-d})$, d. h. mit $(A't)^d$:

$$W_1 = (A't)^d \cdot (B'p) \dots (B'q) \dots (C'p) \dots (E'r)^d \dots (L's)^d.$$

³⁾ Die Formel (6) ergibt sich entweder direkt aus (5) durch passende Zusammenfassung der Faktoren $(A'\xi)(A'\eta) \dots$ oder — was freilich einfacher ist — durch Reihenentwicklung. Da nämlich $(\eta \frac{\partial}{\partial \xi})(TG) = 0$ ist für $\xi \neq \eta$, so treten die Reihen x, y, \dots, z nur in den Verbindungen $p_{ik\dots s}$ auf. Es kommt übrigens für das Folgende nur darauf an, daß man in (5) rechts die Reihen x, y, \dots durch Komplexsymbole p, q, \dots ersetzen darf.

⁴⁾ Vgl. „Inv. Theorie“, S. 79.

⁵⁾ Vgl. „Inv. Theorie“, S. 116 oder ausführlicher: Proceedings der Akad. van Wetenschappen in Amsterdam 30 (Januar 1927), S. 222.

Hier formen wir $(B^{n-d} p q \dots)(p C') \dots$ um, indem wir alle Reihen p im Klammerfaktor vereinigen. Wegen $(p^d q \dots) = 0$ bleibt nur $(B^{n-d} p^d) = (n-d)!(B' p)^d$ über usf., so daß schließlich $(A' t)^d (B' p)^d (C' q)^d \dots$ herauskommt, was aber bis auf einen Zahlenfaktor mit der Polare $(t \frac{\partial}{\partial p}) G$ von (1) übereinstimmt.

Bei den Gliedern W_2 ist wenigstens eine Reihe t' aus dem Klammerfaktor (7) nach außen getreten, W_2 hat daher die Gestalt

$$(8) \quad W_2 = (A'^d B'^{\lambda} \dots D'^{\mu} t'^j) (B' p) \dots (D' q) \dots (t' p) \dots \\ = (-l)^{d(n-d)} d! (A B')^{\lambda} \dots (A D')^{\mu} (A t')^j \dots (t' p) \dots$$

und führt also wegen⁵⁾

$$(9) \quad (t' p) t'_{i_2} \dots i_{n-d} p_{k_2} \dots k_d = \sum_{i=1}^{i=n} t'_{i i_2} \dots i_{n-d} p_{i k_2} \dots k_d = \Pi(t, p)$$

auf die Ausdrücke in Σ von (1). Hiermit ist (1) bewiesen. Überdies zeigt (8), daß die G' in Σ von (1) einen Faktor $(A B')$ enthalten und also durch Faltung aus der Form $(l-2)$ -ten Grades

$$(10) \quad \frac{\partial^2 G}{\partial p_{i k} \dots \partial p_{\lambda \mu} \dots} - \frac{\partial^2 G}{\partial p_{\lambda k} \dots \partial p_{i \mu} \dots} + \dots + (-l)^d \frac{\partial^2 G}{\partial p_{\lambda i k} \dots \partial p_{s \mu} \dots} \\ = \Omega_{i k \dots s, \lambda \mu \dots \nu}$$

hervorgehen⁶⁾. Nach (4) haben wir nämlich:

$$\frac{\partial^2 G}{\partial p_{i k} \dots \partial p_{\lambda \mu} \dots} = l(l-1)(d!)^2 A'_{i k} \dots_s B'_{\lambda \mu} \dots_{\nu} (C' p)^d \dots (L' r)^d,$$

und analog zu (9) führt $(A' B)$ auf (B', A') , also auf (10).

Hieraus folgen⁶⁾ dann die Bemerkungen betreffend die Eindeutigkeit der Clebschschen Normalform von $G(t)$, die allen Differentialgleichungen $\Omega_{i k \dots s, \lambda \mu \dots \nu}(G) = 0$ genügt.

⁵⁾ Siehe F. Mertens, a. a. O.

Eine Verallgemeinerung des Bézoutschen Theorems.

Von

Bartel L. van der Waerden in Göttingen.

§ 1.

Einleitung.

Um die Lösungsklassen (Klassen proportionaler Lösungen) eines Systems von n homogenen Gleichungen $f_1 = 0, \dots, f_n = 0$ der Gradzahlen m_1, \dots, m_n mit $n + 1$ Unbekannten x_0, \dots, x_n zu finden, bildet man die „ u -Resultante“ des Gleichungssystems, d. h. die Resultante der Formen f_1, \dots, f_n und einer allgemeinen Linearform $\sum u_k x_k$. Die u -Resultante hat als Funktion der u_i den Grad $\prod_1^n m_i$ und zerfällt (in einem geeigneten Erweiterungskörper des Rationalitätsbereichs der Koeffizienten) in Linearfaktoren $\sum u_k \xi_k^{(\alpha)}$. Im allgemeinen, das heißt wenn die Formenkoeffizienten Unbestimmte sind, sind diese Linearfaktoren alle verschieden; für spezielle Werte können aber vielfache Faktoren vorkommen; oder es kann vorkommen, daß die u -Resultante identisch verschwindet, nämlich dann, wenn unendlich viele Lösungsklassen vorhanden sind. Sehen wir von dem letzten Fall ab, so repräsentieren die Koeffizienten $\xi_0^{(\alpha)}, \dots, \xi_n^{(\alpha)}$ eines jeden Linearfaktors eine Lösungsklasse. Ein μ -facher Linearfaktor gibt eine „Lösung der Multiplizität μ “. Die Summe der Multiplizitäten der Lösungsklassen¹⁾ ist somit $\prod_1^n m_i$; dies ist der Satz von Bézout in moderner Gestalt²⁾. Wie man sieht, gibt hier die u -Resultante das Mittel, die Multiplizitäten so zu definieren, daß die „Erhaltung der Anzahl“ gilt: die

¹⁾ Man sagt oft ungenau „Anzahl“ statt „Summe der Multiplizitäten“.

²⁾ Vgl. etwa F. S. Macaulay, Algebraic Theory of Modular Systems, Cambridge Tract 19 (1916), p. 15—17. Im folgenden mit „Macaulay“ zitiert.

Anzahl der Lösungen im „allgemeinen“ Fall ist gleich der Summe der Multiplizitäten im Spezialfall.

Geometrisch heißt dieser Satz, daß die Summe der Multiplizitäten der Schnittpunkte von n algebraischen Hyperflächen im projektiven n -Raum gleich dem Produkt der Gradzahlen ist, falls die Anzahl der Schnittpunkte endlich bleibt. Die Multiplizitäten sind dabei wie vorhin definiert durch die Exponenten der Linearfaktoren der u -Resultante.

Man hat oft in der Geometrie³⁾ unter dem Namen „Bézoutschen Satz“ einen viel weitergehenden Satz angeführt und meist ohne Beweis angewandt, der besagt, daß die Summe der Multiplizitäten einer r -dimensionalen und einer $(n - r)$ -dimensionalen Mannigfaltigkeit im projektiven n -Raum (oder speziell einer Kurve und einer Fläche im 3 Raum) gleich dem Produkt der Gradzahlen dieser Mannigfaltigkeiten ist, falls die Anzahl der Schnittpunkte endlich bleibt. Dabei versteht man unter dem Grad einer Mannigfaltigkeit von k komplexen Dimensionen die Anzahl der Schnittpunkte mit einem allgemeinen linearen $n-k$ -Raum. Was aber unter der Multiplizität eines Schnittpunktes zu verstehen ist, hat man anzugeben versäumt.

Erst Hensel und Landsberg haben mit funktionentheoretischen Hilfsmitteln für den Fall $r = 1$ eine Definition der Multiplizität und einen Beweis des Satzes gegeben. Die Multiplizität eines Schnittpunktes wird da definiert als Ordnung des Verschwindens einer gewissen Funktion (genauer: eines gewissen Divisors) in einem (oder mehreren) Punkten der zur gegebenen eindimensionalen Mannigfaltigkeit gehörigen Riemannschen Fläche. Der Beweis läßt sich aber nicht auf beliebige Werte von r übertragen.

Dasselbe gilt von einem Beweisansatz von Lasker, der die Multiplizität durch idealtheoretische Begriffe (Idealmultiplizität oder „Länge“) präzisiert⁴⁾.

³⁾ Siehe z. B.: G. Salmon, *Geometry of three dimensions* (1874), S. 303 (Übers. von Fiedler, II. Teil, S. 86); H. Schubert, *Kalkül der abzählenden Geometrie* (1879), S. 47; E. Study, *Über die Geometrie der Kegelschnitte* (1885), S. 36, und *Math. Annalen* **40** (1892), S. 555. Der angebliche Beweis bei Halphen (*Mémoire sur les courbes algébriques*, *Journal Ec. Pol.* **52** (1882), p. 19) für den dreidimensionalen Fall kann nicht anerkannt werden, weil bei Halphen eine Multiplizitätsdefinition fehlt, also der Satz keinen klaren Sinn hat. Derselbe und noch andere Einwände lassen sich geltend machen auch gegen einen allgemeinen Beweis von M. Noether, *Zur Eliminationstheorie*, *Math. Annalen* **11** (1877), S. 571.

⁴⁾ E. Lasker, *Zur Theorie der Moduln und Ideale*, *Math. Annalen* **60** (1905), S. 20; vgl. auch B. L. v. d. Waerden, *On Hilbert's function etc.*, *Proc. Kon. Ac. Amsterdam* Sept. 1927, wo die Laskerschen Gedanken genauer ausgeführt sind.

Von geometrischer Seite her ist nur ein Beweis von Zeuthen zu verzeichnen, ebenfalls nur für den Fall $r = 1$ gültig⁵⁾.

In dieser Arbeit wird das Problem für beliebige Werte von n und r in Angriff genommen. Ausgangspunkt ist die Forderung, daß die Multiplizitäten wie vorhin so definiert werden sollen, daß die Summe der Multiplizitäten der Schnittpunkte im Spezialfall gleich der Anzahl der Schnittpunkte im allgemeinen Fall ist. Was dabei als „allgemeiner Fall“ anzusehen ist, ist nicht ohne weiteres klar⁶⁾; es erweist sich als zweckmäßig, ihn folgendermaßen zu definieren. Man bringe die beiden gegebenen Mannigfaltigkeiten, deren Schnittpunkte gesucht werden, in eine allgemeine Lage zueinander, indem man die eine mit unbestimmten Transformationsparametern linear transformiert, und bringe sodann die Mannigfaltigkeiten zum Schnitt. Aus diesem allgemeineren Schnittproblem gewinnt man das ursprüngliche zurück, indem man die Transformationsmatrix U als Einheitsmatrix E spezialisiert.

Es handelt sich nun darum, eine Multiplizitätsdefinition zu finden, die die obige Forderung der „Erhaltung der Anzahl“ erfüllt. Da hier die Bézoutsche Resultante ebenso wie die Riemannsche Fläche und der idealtheoretische Längenbegriff versagt (vgl. § 11, Schluß), so muß ein neuer Begriff an ihre Stelle treten. Dieser ist der von mir in der schon zitierten Arbeit W_2 eingeführte Begriff der „relationstreuen Spezialisierung“. Es gibt, wie ich dort zeigte, zu dem vollständigen Schnittpunktssystem $X^{(1)}, \dots, X^{(a)}$ der in allgemeine Lage gebrachten Mannigfaltigkeiten ein und im wesentlichen nur ein Schnittpunktssystem $Y^{(1)}, \dots, Y^{(a)}$ der ursprünglichen Mannigfaltigkeiten, derart, daß alle zwischen den Koordinaten der Punkte $X^{(a)}$ und den Transformationsparametern bestehenden algebraischen Relationen bei der Spezialisierung $X^{(a)} \rightarrow Y^{(a)}, U \rightarrow E$ bestehen bleiben. Die *Multiplizität* ist die Zahl, die angibt, wie oft ein Schnittpunkt unter den Punkten $Y^{(a)}$ vorkommt.

Auf Grund dieser Multiplizitätsdefinition, die selbstverständlich dem Prinzip der „Erhaltung der Anzahl“ genügt, hat die zu beweisende Behauptung einen präzisen Sinn, und eben wegen dieser „Erhaltung der Anzahl“ ist sie äquivalent mit der folgenden:

Sind M und M' algebraische Mannigfaltigkeiten von r und $n - r$ Dimensionen im projektiven n -Raum, und unterwirft man eine der beiden, etwa M , einer linearen Transformation U mit unbestimmten

⁵⁾ H. G. Zeuthen, *Abzählende Methoden* (1914), S. 30. Für beliebige r hat Zeuthen einige Andeutungen für einen Beweis gegeben, aber keinen vollständigen Beweis und keine Multiplizitätsdefinition.

⁶⁾ Vgl. die Ausführung dazu in § 6 meiner Arbeit „Der Multiplizitätsbegriff der algebraischen Geometrie“, *Math. Annalen* 97 (1927), S. 756 (im folgenden mit W_2 zitiert)

Koeffizienten, so ist die Anzahl der Schnittpunkte von M' mit der transformierten Mannigfaltigkeit MU gleich dem Produkt der Gradzahlen von M und M' ⁷⁾ (Hauptsatz).

Zum Beweis dieses Satzes wird eine Methode angewandt, deren Grundgedanke sich schon in der Literatur der abzählenden Geometrie vorfindet, und die systematisch (wenn auch nicht exakt) vor allem von G. Schaake ⁸⁾ angewandt worden ist, nämlich die folgende. Die Mannigfaltigkeit M wird einer ausgearteten linearen Transformation (vermittelt durch eine möglichst allgemeine Matrix vom Rang $n - r + 1$) unterworfen. Sie geht dabei in eine Mannigfaltigkeit MW über, die in so viele lineare Räume zerfällt, wie der Grad von M beträgt. Die Anzahl der Schnittpunkte dieser zerfallenden Mannigfaltigkeit MW mit M' ist gleich dem Produkt der Gradzahlen von M und M' . Dieses Schnittpunktssystem entsteht nun wie vorhin durch relationstreue Spezialisierungen aus dem Schnittpunktssystem von MU und M' , und zwar entsteht, wie gezeigt wird, jeder Schnittpunkt genau einmal (m. a. W. alle Multiplizitäten sind eins). Also ist die Anzahl der Schnittpunkte von MU und M' auch gleich dem Produkt der Gradzahlen von M und M' .

Die nähere Ausführung des eben skizzierten Beweises geschieht mit den Hilfsmitteln der Idealtheorie ⁹⁾, die zu diesem Zwecke vorher in einigen Punkten ausgebaut werden muß, was in den einleitenden §§ 2—6 geschehen wird. In § 2 wird ein allgemeiner Prozeß behandelt, der aus Idealen wieder Ideale, und zwar isolierte Komponenten der ersteren erzeugt, und der eine Erweiterung der Hentzelt-Noetherschen „Grundidealbildung“ darstellt. Dieser Prozeß wird gebraucht für die Untersuchung des Verhaltens eines Polynomideals in einem gegebenen Punkt, und wird in § 3 noch einmal verwendet, wo es sich darum handelt, den Übergang vom projektiven zum Cartesischen Raum (oder von homogenen zu inhomogenen Koordinaten) idealtheoretisch zu erfassen.

⁷⁾ Es ist natürlich für die Anzahl der Schnittpunkte gleichgültig, ob man M allein, oder M' allein, oder beide unabhängig voneinander einer allgemeinen linearen Transformation unterwirft, da man immer durch eine simultane Transformation von M und M' die Transformation von M' rückgängig machen kann.

⁸⁾ G. Schaake, Afbeelding van stelsels figuren op de punten eener lineaire ruimte, Diss. Amsterdam 1922, S. 147—188.

⁹⁾ Für die Grundlagen der Idealtheorie in Polynombereichen und ihrer geometrischen Anwendungen, sowie für weitere Literatur darüber verweise ich auf meine zusammenfassende Arbeit „Zur Nullstellentheorie der Polynomideale“, Math. Annalen 96 (1926), S. 183. Sie wird im folgenden mit W_1 zitiert. Für die körpertheoretischen Sätze, die zur Verwendung kommen werden, verweise ich auf die grundlegende Steinitzsche Arbeit „Algebraische Theorie der Körper“, Journ. f. Math. 137 (1910). S. 167—309. Im folgenden mit „Steinitz“ zitiert.

§ 4 ist eine Zusammenfassung der Ergebnisse von Hilbert und Lasker über die „charakteristische Funktion“ eines von Formen erzeugten Ideals. Diese charakteristische Funktion eignet sich in einigen einfachen Fällen zur Bestimmung der Schnittpunktmultiplizität, eignet sich aber nicht zu einer allgemeinen Definition derselben, weil sie nicht der „Erhaltung der Anzahl“ genügt (vgl. § 11). Sie gibt aber immer eine obere Grenze für die Multiplizität (§ 10). Mit Hilfe der Sätze von § 5 über das Verhalten der Polynomideale bei Erweiterung des Grundkörpers (z. B. bei Adjunktion von Parametern) werden in § 6 die Schnittpunkte einer irreduziblen Mannigfaltigkeit mit allgemeinen linearen Räumen bestimmt.

§ 7 enthält den ersten Schritt zum Beweis des Hauptsatzes: den Nachweis, daß eine irreduzible Mannigfaltigkeit durch eine passend gewählte „ausgeartete“ lineare Transformation in so viele lineare Räume zerfällt wie ihr Grad beträgt. Die definitive Formulierung des Hauptsatzes erfolgt in § 8; der Beweis wird in §§ 9—11 zu Ende geführt. § 9 bringt außerdem den Nachweis, daß die Schnittpunktmultiplizitäten niemals Null (also immer positiv) sind.

In § 12 werden schließlich für den Fall $r = 1$, also wenn M eine Kurve ist, die Multiplizitäten durch charakteristische Funktionen von gewissen Idealen ausgedrückt, und damit der Anschluß an die Laskersche Multiplizitätsdefinition hergestellt. Der Nachweis, daß die gefundene Formel mit der Henselschen Multiplizitätsdefinition übereinstimmt, soll einer späteren Arbeit vorbehalten bleiben.

§ 2.

Die Komponenten der Ideale in einem Punkt.

Es seien einige allgemeine Betrachtungen vorausgeschickt, die auch später noch zur Anwendung kommen werden.

In einem Ring \mathfrak{R} sei eine nichtleere Menge G gegeben, die zu je zwei Elementen a, b auch das Produkt ab enthält. Ist nun \mathfrak{m} ein Ideal in \mathfrak{R} , so verstehe ich unter der *isolierten Komponente*¹⁰⁾ \mathfrak{m}_G von \mathfrak{m} in bezug

¹⁰⁾ Diese Bezeichnung rechtfertigt sich durch Satz 2. E. Noether versteht nämlich unter einem *isolierten Komponentenideal* von \mathfrak{m} einen Durchschnitt von irgend r Primärideal $\mathfrak{q}_{i_1}, \dots, \mathfrak{q}_{i_r}$, die in einer unverkürzbaren Darstellung von \mathfrak{m} durch „größte Primär ideale“

$$\mathfrak{m} = [\mathfrak{q}_1, \dots, \mathfrak{q}_r]$$

auftreten, vorausgesetzt, daß, wenn \mathfrak{p}_{i_2} ein zu einem \mathfrak{q}_{i_2} gehöriges Primideal ist, und wenn irgendein zu einem \mathfrak{q}_j gehöriges Primideal \mathfrak{p}_j ein Vielfaches von \mathfrak{p}_{i_2} ist, dann auch \mathfrak{q}_j unter den Idealen $\mathfrak{q}_{i_1}, \dots, \mathfrak{q}_{i_r}$ vorkommt. Diese Voraussetzungen sind nach Satz 2 für \mathfrak{m}_G alle erfüllt.

auf G die Menge aller Elemente f von \mathfrak{R} , zu denen ein Element g aus G existiert, so daß

$$fg \equiv 0(m).$$

Offenbar liegt m in m_G . Weiter: wenn m ganz in n liegt, so liegt m_G ganz in n_G . Außerdem gelten die folgenden Sätze:

Satz 1. m_G ist ein Ideal.

Beweis. Wenn $fg \equiv 0(m)$, so folgt daraus $rfg \equiv 0(m)$, d. h. wenn f zu m_G gehört, so gehört auch rf dazu. Und wenn $f_1g_1 \equiv 0(m)$ und $f_2g_2 \equiv 0(m)$, so folgt daraus

$$(f_1 - f_2)g_1g_2 = (f_1g_1)g_2 - (f_2g_2)g_1 \equiv 0(m),$$

d. h. wenn f_1 und f_2 zu m_G gehören, so gehört auch $f_1 - f_2$ dazu.

Satz 2. Wenn m als Durchschnitt von Primäridealen¹¹⁾ dargestellt ist:

$$m = [q_1, \dots, q_r],$$

und wenn von den zugehörigen Primäridealen¹²⁾ p_1, \dots, p_r nur p_1, \dots, p_s ($s \leq r$) ein Element von G enthalten, so ist

$$m_G = [q_{s+1}, \dots, q_r]$$

(bzw. $m_G = \mathfrak{R}$ für den Fall $s = r$).

Beweis. Erstens: Aus

$$f \equiv 0(m_G)$$

folgt

$$fg \equiv 0(m),$$

$$fg \equiv 0(q_i) \quad (i = 1, \dots, r).$$

Nun ist aber für $i = s + 1, \dots, r$

$$g \not\equiv 0(p_i),$$

also folgt¹³⁾

$$f \equiv 0(q_i) \quad (i = s + 1, \dots, r),$$

$$f \equiv 0[q_{s+1}, \dots, q_r].$$

In der Hentzelt-Noetherschen Eliminationstheorie (Math. Annalen 88 (1923), S. 53) wird der Prozeß, der uns von m auf m_G führte, angegeben für den Fall, daß G besteht aus allen von x_1, \dots, x_{i-1} freien Polynomen. Das Ideal m_G heißt dort „Grundideal ($i-1$)-ter Stufe“. Auch bei Macaulay kommt es vor (Macaulay, S. 45, Nr. 43).

Es sei noch bemerkt, daß alle Sätze und Beweise ihre Geltung beibehalten, wenn man für G nicht eine Menge von Ringelementen, sondern eine Menge von Idealen nimmt, die zu a und b auch das Produkt $a \cdot b$ enthält. m_G wird dann die Menge aller Elemente f von \mathfrak{R} , zu denen ein g aus G existiert, so daß

$$gf \equiv 0(m).$$

¹¹⁾ W_1 , § 1, 13.

¹²⁾ W_1 , § 1, 13.

¹³⁾ W_1 , § 1, 13.

Zweitens: Aus

$$f \equiv 0 [q_{s+1}, \dots, q_r]$$

folgt

$$(1) \quad fg \equiv 0 [q_{s+1}, \dots, q_r]$$

für beliebige g . Nun liegen in p_1, \dots, p_s Elemente g_1, \dots, g_s von G . Wenn g_i in p_i liegt, so muß eine Potenz $g_i^{e_i}$ in q_i liegen. Setzen wir also

$$g = \prod_1^s g_i^{e_i},$$

so folgt

$$g \equiv 0 [q_1, \dots, q_s],$$

$$(2) \quad fg \equiv 0 [q_1, \dots, q_s].$$

Aus (1) und (2):

$$fg \equiv 0(m).$$

Satz 3. *Es ist $m_{GG} = m_G$, oder: aus $fg \equiv 0(m_G)$ folgt $f \equiv 0(m_G)$.*

Beweis. Aus $fg \equiv 0(m_G)$ folgt $fgg' \equiv 0(m)$, also $f \equiv 0(m_G)$.

Satz 4. *Sind m und n Ideale in \mathfrak{R} , so gelten die Relationen ¹⁴⁾*

$$(3) \quad (m, n)_G = (m_G, n)_G,$$

$$(4) \quad (m \cdot n)_G = (m_G \cdot n)_G.$$

Beweis von (3). Erstens ist

$$(m, n) \equiv 0(m_G, n),$$

also

$$(m, n)_G \equiv 0(m_G, n)_G.$$

Zweitens: Aus

$$f \equiv 0(m_G, n)_G$$

folgt

$$fg \equiv 0(m_G, n),$$

$$fg = a + b, \quad ag' \equiv 0(m), \quad b \equiv 0(n),$$

$$fgg' = ag' + bg' \equiv 0(m, n),$$

$$f \equiv 0(m, n)_G.$$

Genau ebenso beweist man (4).

Es sei nun \mathfrak{R} insbesondere ein Polynombereich $P[x_1, \dots, x_n]$, und G die Menge aller Polynome, die in einem festen Punkt $A = \{\alpha_1, \dots, \alpha_n\}$ des Raumes $C_n(P)$ nicht verschwinden. Nach Satz 1 ist jedem Ideal m ein Ideal m_G zugeordnet, das in diesem Fall mit m_A bezeichnet werden soll. Ist $m = [q_1, \dots, q_r]$, so hat man nach Satz 2, um m_A zu finden, von den q_i nur diejenigen beizubehalten, deren zugehörige Primideale den

¹⁴⁾ Für die Bedeutung von (m, n) und mn siehe W_1 , § 1, 19.

Punkt A als Nullstelle¹⁵⁾ haben. Dementsprechend soll \mathfrak{m}_A die A -Komponente von \mathfrak{m} heißen.

Aus Satz 4 folgen weiter die Gleichungen

$$(5) \quad (\mathfrak{m}, \mathfrak{n})_A = (\mathfrak{m}_A, \mathfrak{n})_A,$$

$$(6) \quad (\mathfrak{m}\mathfrak{n})_A = (\mathfrak{m}_A\mathfrak{n})_A.$$

Das zum Punkt A gehörige Primideal

$$(x_1 - \alpha_1, \dots, x_n - \alpha_n)$$

werde immer mit \mathfrak{a} bezeichnet. Ein zu \mathfrak{a} gehöriges Primärideal, oder ein Ideal mit nur einer Nullstelle A , heißt nach Macaulay¹⁶⁾ *einfach*.

Satz 5. Ist \mathfrak{m}_A einfach, und $\rho \geq$ dem Exponenten¹⁷⁾ von \mathfrak{m}_A , so ist

$$\mathfrak{m}_A = (\mathfrak{m}, \mathfrak{a}^\rho).$$

Beweis. Erstens:

$$\begin{cases} \mathfrak{m} \equiv 0(\mathfrak{m}_A) \\ \mathfrak{a}^\rho \equiv 0(\mathfrak{m}_A), \end{cases}$$

also

$$(7) \quad (\mathfrak{m}, \mathfrak{a}^\rho) \equiv 0(\mathfrak{m}_A).$$

Zweitens: Ist

$$f \equiv 0(\mathfrak{m}_A),$$

so heißt das

$$fg \equiv 0(\mathfrak{m}), \quad g(A) \neq 0.$$

Daraus

$$(8) \quad fg \equiv 0(\mathfrak{m}, \mathfrak{a}^\rho).$$

Das Ideal $(\mathfrak{m}, \mathfrak{a}^\rho)$ hat nur die Nullstelle A , also enthält g keine Nullstelle des Ideals. Daraus und aus (8) folgt (vgl. W_1 , § 5, 8).

$$f \equiv 0(\mathfrak{m}, \mathfrak{a}^\rho),$$

also

$$(9) \quad \mathfrak{m}_A \equiv 0(\mathfrak{m}, \mathfrak{a}^\rho).$$

Aus (7) und (9) folgt die Behauptung.

Man kann diese Sätze noch verallgemeinern, indem man für G nimmt die Menge aller Polynome, die in endlichvielen Punkten A, B, \dots, D nicht verschwinden. Um jetzt \mathfrak{m}_G zu erhalten, hat man von den Primärkomponenten von \mathfrak{m} diejenigen beizubehalten, die in A oder $B \dots$ oder D eine Nullstelle haben. Daraus folgt:

$$\mathfrak{m}_G = [\mathfrak{m}_A, \mathfrak{m}_B, \dots, \mathfrak{m}_D].$$

¹⁵⁾ W_1 , § 3, 4.

¹⁶⁾ Macaulay, S. 36.

¹⁷⁾ W_1 , § 1, 13.

Genau so wie Satz 5 beweist man jetzt:

Satz 6. Sind A, B, \dots, D isolierte Nullstellen eines Ideals \mathfrak{m} , d. h. sind $\mathfrak{m}_A, \mathfrak{m}_B, \dots, \mathfrak{m}_D$ einfache Komponenten von \mathfrak{m} , und ist $\varrho \geq$ allen Exponenten dieser Komponenten, so ist

$$[\mathfrak{m}_A, \mathfrak{m}_B, \dots, \mathfrak{m}_D] = (\mathfrak{m}, a^e b^e \dots d^e).$$

§ 3.

Homogene Ideale und projektive Räume.

Sei P ein algebraisch-abgeschlossener Körper¹⁸⁾, x_0, \dots, x_n Unbestimmte und $\mathfrak{R} = P[x_0, \dots, x_n]$.

Unter einem Punkt des Cartesischen Raumes $C_n(P)$ ¹⁹⁾ wird verstanden ein Wertsystem $\{\xi_1, \dots, \xi_n\}$ aus P . Unter einem Punkt $X = \{\xi_0, \dots, \xi_n\}$ des projektiven Raumes $P_n(P)$ ²⁰⁾ wird verstanden die Gesamtheit der Wertsysteme $\{\lambda \xi_0, \dots, \lambda \xi_n\}$ aus P , wo ξ_0, \dots, ξ_n feste Elemente sind, nicht alle $= 0$, und wo λ alle Werte aus P durchläuft. Ein Punkt von $P_n(P)$ ist also nichts anderes als eine Gerade durch den Koordinatenursprung im $C_{n+1}(P)$. Die $n+1$ Elemente ξ_0, \dots, ξ_n heißen die Koordinaten des Punktes. Die Elemente $\lambda \xi_0, \dots, \lambda \xi_n$, wo $\lambda \neq 0$, können als Koordinaten des nämlichen Punktes verwertet werden.

Ist $\xi_0 \neq 0$, so kann man dem Punkt $X = \{\xi_0, \dots, \xi_n\}$ des $P_n(P)$ den Punkt $\bar{X} = \left\{ \frac{\xi_1}{\xi_0}, \dots, \frac{\xi_n}{\xi_0} \right\}$ des $C_n(P)$ zuordnen. Die Punkte des $P_n(P)$ für die $\xi_0 \neq 0$, und die Punkte des $P_n(P)$ entsprechen sich so umkehrbar eindeutig. Die Elemente $\frac{\xi_i}{\xi_0}$ heißen die inhomogenen Koordinaten des Punktes X . Sind eine endliche Anzahl Punkte gegeben, so kann man durch eine lineare Koordinatentransformation immer erreichen, daß für alle diese Punkte zugleich $\xi_0 \neq 0$ ist, sodaß man für alle diese Punkte zugleich inhomogene Koordinaten einführen kann.

Ein *homogenes Ideal* oder *H-Ideal*²¹⁾ in $\mathfrak{R} = P[x_0, \dots, x_n]$ ist ein solches Ideal, das mit einem Polynom f zugleich alle homogenen Bestandteile von f enthält. Ein homogenes Ideal hat offenbar eine Basis aus homogenen Polynomen oder *Formen*, denn die Polynome einer beliebigen Basis können in homogene Bestandteile aufgelöst werden.

Umgekehrt ist jedes Ideal, dessen Basis aus Formen besteht, homogen.

Sei \mathfrak{m} ein *H-Ideal* in \mathfrak{R} . Vorausgesetzt werde, daß $\{0, \dots, 0\}$ nicht die einzige Nullstelle des Ideals ist. Ist dann $\{\xi_0, \dots, \xi_n\}$ eine andere

¹⁸⁾ Steinitz S. 260.

¹⁹⁾ W_1 § 4, 1.

²⁰⁾ W_2 § 2, Schluß.

²¹⁾ Macaulay, S. 36.

Nullstelle von m im $C_{n+1}(P)$, so sind auch alle Punkte $\{\lambda \xi_0, \dots, \lambda \xi_n\}$ Nullstellen. Also kann man die Nullstellen zusammenfassen zu *Nullstellenklassen* $\{\lambda \xi, \dots, \lambda \xi_n\}$, die *Punkte von* $P_n(P)$ sind. M. a. W. man kann die Mannigfaltigkeit von m in $C_{n+1}(P)$ auch auffassen als Mannigfaltigkeit in $P_n(P)$.

Es soll gezeigt werden, daß zwischen den H -Idealen in $\mathfrak{R} = P[x_0, \dots, x_n]$ und den Idealen in $\overline{\mathfrak{R}} = P[x_1, \dots, x_n]$ ein Entsprechen stattfindet, ähnlich dem zwischen den Punkten X des $P_n(P)$ und den Punkten \overline{X} des $C_n(P)$ (siehe oben).

Aus einem Polynom $f(x_0, \dots, x_n)$ von \mathfrak{R} entsteht, wenn man $x_0 = 1$ setzt, ein Polynom $\bar{f}(x_1, \dots, x_n)$ von $\overline{\mathfrak{R}}$. Jedes Polynom \bar{f} von $\overline{\mathfrak{R}}$ entsteht dabei mindestens einmal. Summen $f_1 + f_2$ und Produkte $f_1 f_2$ gehen dabei wieder in Summen $\bar{f}_1 + \bar{f}_2$ und Produkte $\bar{f}_1 \bar{f}_2$ über („Meromorphismus“). Daraus folgt, daß aus einem Ideal m in \mathfrak{R} wieder ein Ideal \bar{m} in $\overline{\mathfrak{R}}$ entsteht.

Wählt man für m insbesondere ein H -Ideal, so braucht man für die Berechnung der Polynome \bar{f} von \bar{m} nicht alle Polynome f von m zu benutzen, sondern man kann sich auf die *Formen* von m beschränken. Man kann nämlich jedes Polynom f von m in homogene Bestandteile spalten, die ebenfalls zu m gehören, und diese Bestandteile mit solchen Potenzen x_0^e multiplizieren, daß sie den gleichen Grad bekommen. Die Summe ist dann eine Form von m und ergibt für $x_0 = 1$ dasselbe Resultat wie das ursprüngliche Polynom.

Ist umgekehrt ein Ideal \bar{m} in $\overline{\mathfrak{R}}$ gegeben, so erzeugen alle Formen in \mathfrak{R} , die durch die Substitution $x_0 = 1$ in Polynome von \bar{m} übergehen, ein Ideal m_0 , das (weil es von Formen erzeugt wird) H -Ideal ist, und das bei der Substitution $x_0 = 1$ offenbar wieder das Ideal \bar{m} ergibt.

Das so konstruierte H -Ideal soll *das zu* m *äquivalente* H -Ideal heißen²³⁾.

Satz 7. *Das zu* \bar{m} *äquivalente* H -Ideal m_0 *ist das umfassendste* H -Ideal, *das bei der Substitution* $x_0 = 1$ *wieder* \bar{m} *ergibt.*

Das heißt: Jedes H -Ideal m , das bei der Substitution $x_0 = 1$ wieder \bar{m} ergibt, ist Untermenge (Vielfaches) von m_0 .

Beweis. m wird erzeugt von Formen, die bei der Substitution $x_0 = 1$ in Polynome von \bar{m} übergehen; diese Formen liegen in m_0 , also ist m Untermenge von m_0 , *q. e. d.*

Ist m ein H -Ideal und \bar{m} wie oben definiert, so soll das zu \bar{m} äquivalente H -Ideal immer mit m_0 bezeichnet werden. Wir studieren die Beziehung zwischen m und m_0 .

²³⁾ Macaulay, S. 39.

Aus Satz 7 folgt zunächst

$$m \equiv 0(m_0).$$

Weiter gilt

Satz 8. m_0 ist die Menge aller Polynome f , für die eine Potenz x_0^e existiert so daß

$$x_0^e f \equiv 0(m).$$

Beweis. Erstens: Aus $x_0^e f \equiv 0(m)$ folgt, indem man f in homogene Bestandteile f_i spaltet, $x_0^e f_i \equiv 0(m)$; daraus weiter, indem man $x_0 = 1$ setzt, $\bar{f}_i \equiv 0(\bar{m})$, und daraus $f_i \equiv 0(m_0)$, mithin $f \equiv 0(\bar{m}_0)$.

Ist umgekehrt f ein Polynom von m_0 , so ist f eine Summe von Formen f_i , die bei der Substitution $x_0 = 1$ in Polynome \bar{f}_i von \bar{m} übergehen. Diese \bar{f}_i entstehen ihrerseits durch die Substitution $x_0 = 1$ aus Formen g_i von m . Wenn aber zwei Formen f_i, g_i bei der Substitution $x_0 = 1$ dasselbe Polynom \bar{f}_i ergeben, so unterscheiden sie sich bloß um Faktoren x_0 , d. h. es ist

$$\begin{aligned} x_0^{e_i} f_i &= x_0^{e_i} g_i \\ &\equiv 0(m). \end{aligned}$$

Ist nun $\rho = \max(\rho_i)$, so folgt

$$x_0^\rho f_i \equiv 0(m),$$

mithin wegen $f = \sum f_i$:

$$x_0^\rho f \equiv 0(m),$$

womit der Satz bewiesen ist.

Aus diesem Satz folgt, daß man die Beziehung zwischen m und m_0 der in § 2 gegebenen allgemeinen Theorie unterordnen kann, indem man unter G versteht die Menge aller Potenzen von x_0 . Insbesondere gilt Satz 2:

Wenn m als Durchschnitt von Primäridealen dargestellt ist:

$$m = [q_1, \dots, q_r],$$

und wenn von den zugehörigen Primidealen p_1, \dots, p_r nur p_1, \dots, p_s ($s \leq r$) eine Potenz von x_0 enthalten, so ist

$$m_0 = [q_{s+1}, \dots, q_r].$$

Soll ein Primideal eine Potenz von x_0 , also x_0 selbst enthalten, so muß in der allgemeinen Nullstelle $\{\xi_0, \dots, \xi_n\}$ des Primideals $\xi_0 = 0$ sein.

Wir heben nun aus den q_i besonders diejenigen hervor, die als einzige Nullstelle den Ursprung $\{0, \dots, 0\}$ haben, d. h. deren zugehöriges Primideal ist $u = (x_0, \dots, x_n)$, falls solche vorhanden sind. In den allgemeinen Nullstellen aller übrigen zugehörigen Primidealen p_i ist mindestens eine Koordinate $\xi_i \neq 0$. Also läßt es sich durch eine lineare Koordinaten-

transformation erreichen, daß in allen diesen allgemeinen Nullstellen $\xi_0 \neq 0$ ist. Wir wollen uns diese Koordinatentransformation immer vorher ausgeführt denken. Unter dieser Voraussetzung vereinfacht sich der eben formulierte Satz zu:

Satz 9. *Aus einer Darstellung von m als Durchschnitt von Primäridealien erhält man eine ebensolche Darstellung für m_0 , indem man die Primäridealien ausläßt, die nur die eine Nullstelle $\{0, \dots, 0\}$ haben.*

Es sollen jetzt noch einige formale Relationen zwischen den Idealen m, \bar{m}, m_0 abgeleitet werden.

Satz 10. *Ist $m = (f_1, \dots, f_r)$, so ist*

$$\bar{m} = (\bar{f}_1, \dots, \bar{f}_r).$$

Beweis: Klar.

Folgen. 1. $(\bar{m}, \bar{n}) = (\bar{m}, \bar{n}).$

2. $\overline{m n} = \bar{m} \cdot \bar{n}.$

Satz 11. *Sind m, n zwei H -Idealien derart, daß nach vorheriger Koordinatentransformation (siehe oben)*

$$\bar{m} = \bar{n},$$

so bestehen Relationen der Gestalt

$$\begin{cases} u^e m \equiv 0(n) \\ u^e n \equiv 0(m) \end{cases}^{23},$$

wo $u = (x_0, \dots, x_n)$ gesetzt ist.

Beweis. Sei, wie in Satz 9,

$$m = [q_1, \dots, q_r]$$

und

$$m_0 = [q_{s+1}, \dots, q_r].$$

Da m_0 das zu \bar{m} äquivalente und n_0 das zu \bar{n} äquivalente H -Ideal ist, und da $\bar{m} = \bar{n}$, so muß $m_0 = n_0$. Ist nun ρ der größte der Exponenten der (zum Primideal u gehörigen) Primäridealien q_1, \dots, q_s , so ist

$$(1) \quad u^\rho \equiv 0([q_1, \dots, q_s]).$$

Auch ist

$$n \equiv 0(n_0)$$

oder

$$(2) \quad n \equiv 0([q_{s+1}, \dots, q_r]).$$

²³) Diese Relationen bilden den bequemsten algebraischen Ausdruck für die Tatsache, daß, sobald $\bar{m} = \bar{n}$, die Idealien m, n sich bloß unterscheiden können um Primärkomponenten, die zum Primideal u gehören, also die einzige Nullstelle $\{0, \dots, 0\}$ haben.

Aus (1) und (2) folgt

$$u^e n \equiv 0 [q_1, \dots, q_{s+1}, \dots, q_r],$$

oder

$$u^e n \equiv 0 (m).$$

Genau so beweist man

$$u^s m! \equiv 0 (n); \quad \text{q. e. d.}$$

Ist $X = \{\lambda \xi_0, \dots, \lambda \xi_n\}$ eine Nullstellenklasse von m im $P_n(P)$, und ist $\xi_0 \neq 0$, so kann man (durch Multiplikation mit einem passenden Proportionalitätsfaktor) erreichen, daß $\xi_0 = 1$ ist. Dann ist $\bar{X} = \{\xi_1, \dots, \xi_n\}$ Nullstelle von \bar{m} .

Ist umgekehrt \bar{X} Nullstelle von \bar{m} , so ist X Nullstellenklasse von m . Die Mannigfaltigkeit von \bar{m} im $C_n(P)$ wird also aus der von m im $P_n(P)$ dadurch konstruiert, daß man zu allen Punkten X der letzteren, soweit sie nicht in der Hyperebene $\xi_0 = 0$ liegen, die entsprechenden Punkte \bar{X} des $C_n(P)$ aufsucht.

Ist insbesondere \mathfrak{p} ein Primideal, und ist $\{0, \dots, 0\}$ nicht die einzige Nullstelle von \mathfrak{p} , so ist nach Satz 9 $\mathfrak{p}_0 = \mathfrak{p}$. Es gilt weiter:

Satz 12. Ist \mathfrak{p} prim, und $\{0, \dots, 0\}$ nicht seine einzige Nullstelle, so ist auch $\bar{\mathfrak{p}}$ prim, und jede allgemeine Nullstelle²⁴⁾ $\{\xi_1, \dots, \xi_n\}$ von $\bar{\mathfrak{p}}$ ergibt, wenn λ eine neu adjungierte Unbestimmte ist, eine allgemeine Nullstelle $\{\lambda, \lambda \xi_1, \dots, \lambda \xi_n\}$ von \mathfrak{p} .

Beweis. 1. Ist $\bar{f}\bar{g} \equiv 0 (\bar{\mathfrak{p}})$ und sind f, g die durch Homogenisieren aus \bar{f}, \bar{g} gebildeten Formen, so folgt $fg \equiv 0 (\mathfrak{p}_0)$, also entweder $f \equiv 0 (\mathfrak{p}_0)$ oder $g \equiv 0 (\mathfrak{p}_0)$, also entweder $\bar{f} \equiv 0 (\bar{\mathfrak{p}})$ oder $\bar{g} \equiv 0 (\bar{\mathfrak{p}})$. Also ist $\bar{\mathfrak{p}}$ prim.

2. Daß $\{\lambda, \lambda \xi_1, \dots, \lambda \xi_n\}$ Nullstelle von \mathfrak{p} ist, wurde oben schon bemerkt. Ist nun f ein Polynom in $P[x_0, \dots, x_n]$, und ist $f(\lambda, \lambda \xi_1, \dots, \lambda \xi_n) = 0$, so kann man zunächst f in homogene Bestandteile zerlegen:

$$f = \sum_1^s f_i.$$

Ist γ_i der Grad von f_i , so folgt

$$f(\lambda, \lambda \xi_1, \dots, \lambda \xi_n) = \sum \lambda^{\gamma_i} f_i(1, \xi_1, \dots, \xi_n) = 0,$$

also

$$f_i(1, \xi_1, \dots, \xi_n) = 0,$$

also

$$f_i(1, x_1, \dots, x_n) \equiv 0 (\bar{\mathfrak{p}}),$$

$$f_i(x_0, x_1, \dots, x_n) \equiv 0 (\mathfrak{p}_0) \equiv 0 (\mathfrak{p}),$$

$$f = \sum_1^s f_i \equiv 0 (\mathfrak{p}).$$

Mithin ist $\{\lambda, \lambda \xi_1, \dots, \lambda \xi_n\}$ allgemeine Nullstelle von \mathfrak{p} .

²⁴⁾ W_1 § 3, 4.

Ist r die Dimensionszahl von \bar{p} ²⁵), so ist offenbar $r + 1$ die Dimensionszahl von p , weil durch die Adjunktion von λ der Transzendenzgrad der allgemeinen Nullstelle um eins vermehrt ist. Man ist gewohnt, bei der Berechnung der Dimensionszahl der Mannigfaltigkeit M von p im $P_n(P)$ den Parameter λ als unwesentlich zu betrachten, also unter dieser Dimensionszahl die Zahl r zu verstehen. Die Dimensionszahl einer irreduziblen Mannigfaltigkeit im projektiven Raum unterscheidet sich demnach um eins von der Dimensionszahl des zugehörigen H -Ideals. Die Zahl r , die für H -Ideale im allgemeinen eine größere Rolle spielt als die wirkliche Dimensionszahl $r + 1$, wollen wir die *reduzierte Dimension* des H -Ideals p nennen. Die reduzierte Dimension eines Primärideals wird definiert als die des zugehörigen Primideals, und die eines beliebigen Ideals als die der höchst-dimensionalen Primärkomponente.

Auch für H -Ideale kann man die Komponente in einem Punkt bilden (vgl. § 2). Ist $Y = \{\eta_0, \dots, \eta_n\}$ ein Punkt im $P_n(P)$, m ein H -Ideal, so wird die Y -Komponente m_Y von m definiert als Gesamtheit derjenigen Polynome f , zu denen eine Form g existiert, so daß

$$\begin{cases} fg \equiv 0(m), \\ g(Y) \neq 0. \end{cases}$$

Offenbar ist m_Y wieder ein H -Ideal. Der Übergang $m \rightarrow m_Y$ läßt sich ohne weiteres in die allgemeine Theorie des § 2 einordnen, indem man unter G versteht die Menge aller Formen, die im Punkt Y nicht verschwinden. Demnach gelten auch hier die Gleichungen (5), (6) § 2, und kann man m_0 dadurch erhalten, daß man von den Primärkomponenten von m nur diejenigen beibehält, deren zugehörige Primideale in Y eine Nullstelle haben.

Der Übergang zu inhomogenen Koordinaten wird ermöglicht durch

Satz 13. *Ist die Numerierung der Koordinaten so gewählt, daß $\eta_0 \neq 0$ ist, so ist*

$$\overline{m}_Y = \overline{m}_{\bar{Y}}.$$

Beweis. Sei \bar{f} Element von $\overline{m}_{\bar{Y}}$. Dann entsteht \bar{f} aus einer Form f von m_Y durch die Substitution $x_0 = 1$. Also ist

$$fg \equiv 0(m), \quad g(Y) \neq 0.$$

Daraus

$$\bar{f}\bar{g} \equiv 0(\overline{m}), \quad \bar{g}(\bar{Y}) \neq 0,$$

also gehört \bar{f} zu $\overline{m}_{\bar{Y}}$. Gehört umgekehrt \bar{f} zu $\overline{m}_{\bar{Y}}$, so folgt

²⁵) W_1 § 3, 5.

$$\begin{aligned}\bar{f}\bar{g} &\equiv 0(\bar{m}), & \bar{g}(\bar{Y}) &\neq 0, \\ fgx_0^e &\equiv 0(m), & g(Y) &\neq 0, & \eta_0^e &\neq 0, \\ f &\equiv 0(m_Y), \\ \bar{f} &\equiv 0(\bar{m}_Y).\end{aligned}$$

§ 4.

Die Hilbertsche charakteristische Funktion.

Hilbert und Lasker haben einige Sätze entwickelt, die sich auf die Anzahl der modulo einem H -Ideal linear-unabhängigen Formen gegebener Ordnung beziehen. Sätze und Beweise sollen hier kurz rekapituliert werden, weil man die Sache noch etwas einfacher darstellen und allgemeiner fassen kann, als es in der (gegenüber Hilbert durch Vermeidung der Syzygienketten schon sehr vereinfachten) Laskerschen Darstellung geschehen ist.

Definitionen. Es sei P ein Körper und $P[x_0, \dots, x_n]$ der aus ihm abgeleitete Polynombereich. Mit $\varphi(\varrho)$ bezeichnen wir die Anzahl der Potenzprodukte der x_0, \dots, x_n vom Grad ϱ . Mit $\varphi(\varrho; a)$ bezeichnen wir die Anzahl der linear-unabhängigen Formen vom Grad ϱ in einem H -Ideal a . Mit $\chi(\varrho; a)$ schließlich bezeichnen wir die Anzahl der modulo dem Ideal a linear-unabhängigen Formen vom Grad ϱ , oder die Anzahl der unabhängigen linearen Gleichungen, die die Koeffizienten einer Form ϱ -ten Grades zu erfüllen haben, damit die Form im Ideal liegt. Auf die χ -Funktion kommt es im folgenden hauptsächlich an.

Offenbar ist:

$$\begin{aligned}(1) \quad & \varphi(\varrho) = \varphi(\varrho; (1)), \\ (2) \quad & \chi(\varrho; a) = \varphi(\varrho) - \varphi(\varrho, a).\end{aligned}$$

Satz 14. Sind a, b zwei H -Ideale, (a, b) ihre Summe, $[a, b]$ ihr Durchschnitt, so ist

$$\begin{aligned}(3) \quad & \varphi(\varrho; (a, b)) = \varphi(\varrho; a) + \varphi(\varrho; b) - \varphi(\varrho; [a, b]), \\ (4) \quad & \chi(\varrho; (a, b)) = \chi(\varrho; a) + \chi(\varrho; b) - \chi(\varrho; [a, b]).\end{aligned}$$

Beweis. (3) ist unmittelbar klar. (4) folgt aus (2) und (3).

Satz 15. Ist f eine Form vom Grad γ , und relativ prim zu a (das letztere heißt, daß aus $fg \equiv 0(a)$ notwendig $g \equiv 0(a)$ folgt), so ist

$$(5) \quad \chi(\varrho; (a, f)) = \chi(\varrho; a) - \chi(\varrho - \gamma; a).$$

Beweis. Sei \mathfrak{f} das von f erzeugte Ideal. Wir bestimmen zunächst die Funktionen $\varphi(\varrho; \mathfrak{f})$ und $\varphi(\varrho, [a, \mathfrak{f}])$. Alle Formen vom Grad ϱ in \mathfrak{f} haben die Gestalt $f \cdot g$, wo g den Grad $\varrho - \gamma$ hat. Also folgt

$$(6) \quad \varphi(\varrho; \mathfrak{f}) = \varphi(\varrho - \gamma).$$

Die Formen vom Grad ϱ in $[a, f]$ müssen sowohl in f liegen, also die Gestalt $f \cdot g$ haben, als auch in a liegen:

$$f \cdot g \equiv 0(a).$$

Aus dieser Kongruenz folgt aber $g \equiv 0(a)$, und umgekehrt. Also ist die Anzahl der linear-unabhängigen Formen $f \cdot g$ vom Grad ϱ in $[a, f]$ gleich der Anzahl der linear-unabhängigen Formen g vom Grad $\varrho - \gamma$ in a . Also:

$$(7) \quad \varphi(\varrho; [a, f]) = \varphi(\varrho - \gamma; a).$$

Nunmehr hat man nach (4):

$$\chi(\varrho; (a, f)) = \chi(\varrho; a) + \chi(\varrho; f) - \chi(\varrho; [a, f])$$

nach (2):

$$= \chi(\varrho; a) - \varphi(\varrho; f) - \varphi(\varrho; [a, f])$$

nach (6), (7):

$$= \chi(\varrho; a) - \varphi(\varrho - \gamma) + \varphi(\varrho - \gamma; a),$$

nach (2):

$$= \chi(\varrho; a) - \chi(\varrho - \gamma; a).$$

Satz 16. Ist $\{0, \dots, 0\}$ die einzige Nullstelle des Ideals a , so ist für hinreichend hohe ϱ :

$$\chi(\varrho; a) = 0.$$

Beweis. Nach dem Hilbertschen Nullstellensatz²⁶⁾ liegt jedes Potenzprodukt von hinreichend hohem Grad im Ideal.

Satz 17. Ist a ein H -Ideal und d seine reduzierte Dimension, so wird die Funktion $\chi(\varrho; a)$ für hinreichend hohe ϱ dargestellt durch ein Polynom in ϱ von der Gestalt

$$(8) \quad \chi(\varrho; a) = a_0 \binom{\varrho}{d} + a_1 \binom{\varrho}{d-1} + \dots + a_d \quad (a_0 \geq 0)$$

mit ganzzahligen Koeffizienten a_r . Man nennt dieses Polynom die charakteristische Funktion des Ideals.

Beweis. Ist die reduzierte Dimension $d = -1$, also die gewöhnliche Dimension Null, so hat das Ideal a nur die Nullstelle $\{0, \dots, 0\}$ und nach Satz 16 ist $\chi(\varrho; a) = 0$. Wenn man also die Null als Funktion (-1) -ten Grades von ϱ ansieht, so stimmt der Satz für Ideale der reduzierten Dimension -1 . Es sei also $d \geq 0$, und der Satz sei für alle reduzierten Dimensionen kleiner als d bewiesen. Das gegebene Ideal a läßt sich darstellen als Durchschnitt von Primäridealen:

$$a = [q_1, \dots, q_r],$$

wo alle q_i höchstens die reduzierte Dimension d haben, und wo mindestens ein q_i wirklich diese Dimension erreicht. Setzt man den Satz für Primär-

²⁶⁾ W₁, § 5, 9.

ideale, sowie (falls $r > 1$) für Durchschnitte von weniger als r Primär-idealen als bewiesen voraus, so folgt aus (4)

$$\begin{aligned}\chi(\varrho; \alpha) &= \chi(\varrho; [q_1, \dots, q_{r-1}, q_r]) \\ &= \chi(\varrho; [q_1, \dots, q_{r-1}]) + \chi(\varrho; q_r) - \chi(\varrho; ([q_1, \dots, q_{r-1}], q_r)).\end{aligned}$$

Das letzte Ideal $([q_1, \dots, q_{r-1}], q_r)$ hat eine reduzierte Dimension $< d$, also hat seine χ -Funktion die gesuchte Gestalt mit $a_0 = 0$. Die ersten beiden Glieder in der vorigen Gleichung haben diese Gestalt nach Annahme auch (mit $a_0 \geq 0$); also muß auch die linke Seite diese Gestalt haben. Somit bleibt nur, den Satz für Primär Ideale zu beweisen. Sei α primär und \mathfrak{p} das zugehörige Primideal. Wir wählen eine Linearform l , die nicht in \mathfrak{p} liegt. Dann hat (α, l) höchstens die reduzierte Dimension d , und l ist zu α relativ prim. Also gilt (5):

$$(9) \quad \chi(\varrho; (\alpha, l)) = \chi(\varrho; \alpha) - \chi(\varrho - 1; \alpha).$$

Nach Induktionsvoraussetzung hat die linke Seite für $\varrho \geq \varrho_0$ die Gestalt

$$a_0 \binom{\varrho}{d-1} + a_1 \binom{\varrho}{d-2} + \dots + a_{d-1} \quad (a_0 \geq 0).$$

Aus der Rekursionsformel (9) findet man nun:

$$\begin{aligned}\chi(\varrho; \alpha) - \chi(\varrho_0; \alpha) &= a_0 \left\{ \binom{\varrho+1}{d} - \binom{\varrho_0+1}{d} \right\} \\ &\quad + a_1 \left\{ \binom{\varrho+1}{d-1} - \binom{\varrho_0+1}{d-1} \right\} + \dots + a_{d-1} (\varrho - \varrho_0).\end{aligned}$$

Wenn man alle Konstanten zusammensucht und a_d nennt, und beachtet, daß $\binom{\varrho+1}{k} = \binom{\varrho}{k} + \binom{\varrho}{k-1}$ ist, so steht die Behauptung (8) da.

Den (nichtnegativen) Koeffizienten a_0 von $\binom{\varrho}{d}$ kann man den Grad des Ideals α nennen, weil er auf das engste mit der Anzahl der Schnittpunkte der zugehörigen Mannigfaltigkeit mit einem allgemeinen linearen Raum von $n - d$ Dimensionen zusammenhängt²⁷⁾. Aus dem eben gegebenen Beweis folgt noch:

Satz 18. *Der Grad eines H-Ideals ist die Summe der Grade der Primärkomponenten gleicher Dimension.*

Für $d = 0$ besteht die charakteristische Funktion nur aus einem Glied a_0 . Das Ideal hat Nullstellen, die von $\{0, \dots, 0\}$ verschieden sind, und es gibt von beliebig hohem Grad Formen, die nicht in allen diesen

²⁷⁾ Für Primideale ist nämlich Grad des Ideals = Grad der Mannigfaltigkeit. Für Primär Ideale läßt sich zeigen: der Grad ist gleich l mal dem Grad des zugehörigen Primideals, wo l die „Länge“ des Primär Ideals ist, d. h. die Länge einer maximalen Kette von Primär Idealen $\mathfrak{q} = \mathfrak{q}_1, \mathfrak{q}_2, \dots, \mathfrak{q}_l = \mathfrak{p}$ (\mathfrak{q}_{i+1} echter Teiler von \mathfrak{q}_i), die zum selben Primideal gehören. Vgl. meine unter 4) zitierte Arbeit.

Nullstellen verschwinden, also nicht dem Ideal angehören; demnach ist $a_0 > 0$. Dasselbe gilt nun, wie wir zeigen werden, für jede beliebige Dimension:

Satz 19. *Der Grad eines H -Ideals α der reduzierten Dimension $d \geq 0$ ist eine positive ganze Zahl.*

Es genügt, den Satz für Primärideale zu beweisen; wir können außerdem annehmen, es sei $d > 0$, und der Satz sei für Ideale niedrigerer Dimension schon bewiesen. Ist α primär, \mathfrak{p} das zugehörige Primideal, und $f \not\equiv 0(\mathfrak{p})$, so folgt aus (5) durch Vergleichung der Koeffizienten von ϱ^{d-1} auf beiden Seiten:

$$(10) \quad \text{Grad}(\alpha, f) = \gamma \cdot \text{Grad } \alpha,$$

falls (α, f) genau die reduzierte Dimension $d - 1$ hat. Hat aber (α, f) eine niedrigere Dimension, so hat man in (10) links *Null* zu schreiben. Die Behauptung wird nun bewiesen sein, sobald man eine Form f angeben kann, derart, daß (α, f) genau die reduzierte Dimension $d - 1$ hat. Geht man zu inhomogenen Idealen über, so ist also zu zeigen, daß es ein Polynom \bar{f} gibt, derart, daß $(\bar{\alpha}, \bar{f})$ die Dimension $d - 1$ hat. Nun sei $\bar{\mathfrak{p}}$ das zu $\bar{\alpha}$ gehörige Primideal, $\bar{\mathfrak{p}}_1$ ein Primteiler der Dimension $d - 1$ von $\bar{\mathfrak{p}}$ ²⁸⁾, und \bar{f} so gewählt, daß $\bar{f} \equiv 0(\bar{\mathfrak{p}}_1)$, aber $\bar{f} \not\equiv 0(\bar{\mathfrak{p}})$. Dann genügt \bar{f} unseren Anforderungen und der Satz ist bewiesen. Nachträglich folgt nunmehr, daß in (10) rechts niemals *Null* stehen kann, wie auch $f \not\equiv 0(\mathfrak{p})$ gewählt werde, also daß (α, f) immer genau die reduzierte Dimension $d - 1$ hat.

Ist \mathfrak{m} ein H -Ideal und definiert man \mathfrak{m}_0 wie im § 3, unter Berücksichtigung der für Satz 9 nötigen Koordinatentransformation, so haben \mathfrak{m} und \mathfrak{m}_0 dieselben Primärkomponenten bis auf solche Komponenten, die nur eine Nullstelle haben, also dieselbe charakteristische Funktion. Nach § 3 entspricht dem Ideal \mathfrak{m} bei Übergang zu inhomogenen Koordinaten ein Ideal $\bar{\mathfrak{m}}$, derart, daß die Formen vom Grad ϱ in \mathfrak{m}_0 eineindeutig den Polynomen vom Grad $\leq \varrho$ in $\bar{\mathfrak{m}}$ entsprechen, in die sie durch $x_0 = 1$ übergehen. Also ist die Anzahl der linear-unabhängigen Polynome vom Grad $\leq \varrho$ in $\bar{\mathfrak{m}}$ gleich der Anzahl der linear-unabhängigen Formen vom Grad ϱ in \mathfrak{m}_0 , und ebenso die Anzahl der modulo $\bar{\mathfrak{m}}$ linear-unabhängigen Polynome vom Grad $\leq \varrho$ gleich der Anzahl der modulo \mathfrak{m}_0 linear-unabhängigen Formen vom Grad ϱ . Daraus und aus der Tatsache, daß \mathfrak{m} und \mathfrak{m}_0 dieselbe charakteristische Funktion haben, folgt:

Satz 20. *Ist $\bar{\mathfrak{m}}$ ein d -dimensionales Ideal, so wird die Anzahl der modulo $\bar{\mathfrak{m}}$ linear-unabhängigen Polynome vom Grad $\leq \varrho$ für große ϱ*

²⁸⁾ Einen solchen gibt es nach W_1 § 3, 8.

durch die charakteristische Funktion des H -Ideals \mathfrak{m} , also durch ein Polynom d -ten Grades der Gestalt (8) dargestellt.

Ist insbesondere $\overline{\mathfrak{m}}$ nulldimensional ($d = 0$), so ist die charakteristische Funktion von \mathfrak{m} eine Konstante a_0 , Summe der ebenfalls konstanten charakteristischen Funktionen der Primärkomponenten gleicher Dimension. Also ist auch die Anzahl der modulo \mathfrak{m} linear-unabhängigen Polynome beliebig hohen Grades eine Konstante und gleich der Summe der entsprechend gebildeten Konstanten der Primärkomponenten. Man kann dieses auch so ausdrücken:

Satz 21. *Der Restklassenbereich $\overline{\mathfrak{R}}/\overline{\mathfrak{m}}$ ist, wenn $\overline{\mathfrak{m}}$ nulldimensional ist, ein P -Modul von endlichem Rang. Dieser Rang ist gleich der Summe der Rangzahlen der Restklassenbereiche nach den Primärkomponenten und gleich der charakteristischen Funktion des H -Ideals \mathfrak{m} .*

Für nulldimensionale Primideale $\overline{\mathfrak{p}}$ ist der Rang von $\overline{\mathfrak{R}}/\overline{\mathfrak{p}}$ wegen der Isomorphie $\overline{\mathfrak{R}}/\overline{\mathfrak{p}} \cong P(\xi_1, \dots, \xi_n)$ (W_1 § 3, 2) gleich dem Grad des Nullstellenkörpers $P(\xi_1, \dots, \xi_n)$ in bezug auf P . Ist insbesondere P algebraisch abgeschlossen, so ist dieser Grad 1. Also ist in diesem Fall auch die charakteristische Funktion von $\overline{\mathfrak{p}}$ gleich 1.

§ 5.

Erweiterung des Grundkörpers.

Es sei Ω ein Erweiterungskörper von P ; es sei

$$\mathfrak{R} = P[x_1, \dots, x_n]$$

und

$$\mathfrak{R}_\Omega = \Omega[x_1, \dots, x_n].$$

Ein Ideal \mathfrak{m} in \mathfrak{R} erzeugt ein Ideal \mathfrak{m}_Ω in \mathfrak{R}_Ω . Wir untersuchen die Beziehungen zwischen \mathfrak{m} und \mathfrak{m}_Ω .

Satz 22. *Der Durchschnitt von \mathfrak{m}_Ω mit \mathfrak{R} ist wieder \mathfrak{m} .*

Beweis²⁹⁾. Es sei f ein gemeinsames Element von \mathfrak{m}_Ω und \mathfrak{R} . Zu beweisen ist:

$$f \in \mathfrak{m} \text{ } ^{30)}.$$

Nach der Definition von \mathfrak{m}_Ω ist

$$(1) \quad f = \sum_1^r \alpha_i f_i, \quad f_i \in \mathfrak{m}, \quad \alpha_i \in \Omega.$$

Drückt man die Größen α_i durch endlichviele linear-unabhängige Elemente $1, \omega_1, \omega_2, \dots$ von Ω mit Koeffizienten aus P aus, so kommt:

$$f = g_0 + \omega_1 g_1 + \dots, \quad \text{wo } g_i \in \mathfrak{m},$$

²⁹⁾ Vgl. W_1 § 6, 2.

³⁰⁾ ε heißt: ist Element von.

mithin, da die ω_i linear-unabhängig in bezug auf P sind:

$$f = g_0, \quad 0 = g_1, \dots$$

also

$$f \in \mathfrak{M}.$$

Satz 23. Die charakteristische Funktion eines H -Ideals ändert sich bei Erweiterung des Grundkörpers nicht.

Beweis. Wir haben zu zeigen, daß

$$\chi(\varrho; \mathfrak{a}) = \chi(\varrho; \mathfrak{a}_\Omega)$$

oder, was nach (2) § 5 auf dasselbe hinauskommt, daß

$$\varphi(\varrho; \mathfrak{a}) = \varphi(\varrho; \mathfrak{a}_\Omega)$$

Sei nun f_1, \dots, f_φ ein Maximalsystem von linear-unabhängigen Formen vom Grad ϱ in \mathfrak{a} , so daß sich alle Formen desselben Grades in \mathfrak{a} linear durch diese ausdrücken. Dann drücken sich auch alle Formen von \mathfrak{a}_Ω linear durch die f_i aus, und die lineare Unabhängigkeit bleibt bestehen, denn würde eine lineare Abhängigkeit im Körper Ω bestehen, so könnte man damit ebenso verfahren wie oben mit der Gleichung (1), und käme zu einer linearen Abhängigkeit in P . Damit ist $\varphi(\varrho; \mathfrak{a}) = \varphi(\varrho; \mathfrak{a}_\Omega)$ bewiesen.

Satz 24. Voraussetzung. Ω sei Galoisch in bezug auf P . \mathfrak{A} sei ein Ideal in \mathfrak{K}_Ω , daß von lauter solchen Polynomen erzeugt wird, deren Koeffizienten von erster Art^{30a)} in bezug auf P sind. \mathfrak{A} sei mit allen seinen konjugierten Idealen in bezug auf P identisch.

Behauptung. \mathfrak{A} wird von einem Ideal \mathfrak{a} aus \mathfrak{K} erzeugt.

Beweis. Sei \mathfrak{a} der Durchschnitt von \mathfrak{K} mit \mathfrak{A} , und \mathfrak{a}_Ω wie vorhin das von \mathfrak{a} in Ω erzeugte Ideal. \mathfrak{a}_Ω ist sicher Untermenge von \mathfrak{A} ; wenn wir zeigen können, daß alle erzeugenden Elemente von \mathfrak{A} zu \mathfrak{a}_Ω gehören, so sind wir fertig.

Sei f eine solche Erzeugende. Der Unterkörper Γ von Ω , der durch Adjunktion der Koeffizienten von f und aller dazu konjugierten Größen entsteht, ist Galoisch, von erster Art und von endlichem Grad in bezug auf P . Eine Basis von Γ sei $(\omega_1, \dots, \omega_m)$; die konjugierten Basen seien $(\omega_1^{(k)}, \dots, \omega_m^{(k)})$ ($k = 1, \dots, m$). Das Polynom f drückt sich durch $\omega_1, \dots, \omega_m$ linear aus:

$$(2) \quad f = \sum f_i \omega_i, \quad f_i \in \mathfrak{K}.$$

Durch Anwendung desjenigen Automorphismus von Γ , der $(\omega_1, \dots, \omega_m)$ in $(\omega_1^{(k)}, \dots, \omega_m^{(k)})$ überführt, gehe f in $f^{(k)}$ über. Es folgt

$$f^{(k)} = \sum f_i \omega_i^{(k)}.$$

^{30a)} Steinitz § 13.

Aus diesem Gleichungssystem lassen sich wegen $|\omega_i^{(k)}| \neq 0$ die f_i ausrechnen:

$$f_i = \sum \alpha_i^{(k)} f^{(k)}.$$

Da die $f^{(k)}$, als Konjugierte von f , in \mathfrak{A} liegen, liegen auch die f_i in \mathfrak{A} , also in \mathfrak{a} . Aus (2) folgt jetzt, daß f in \mathfrak{a}_Ω liegt, q. e. d.

Ist \mathfrak{p} ein nulldimensionales Primideal in $P[x_1, \dots, x_n]$, so ist die allgemeine Nullstelle $\{\xi_1, \dots, \xi_n\}$ algebraisch in bezug auf P . Man nennt \mathfrak{p} von erster Art, wenn alle ξ_i von erster Art in bezug auf P sind; sonst von zweiter Art.

Eine wichtige Eigenschaft der Primideale erster Art ist:

Satz 25. *Es sei \mathfrak{p} ein nulldimensionales Primideal von erster Art in \mathfrak{K} . Es sei Ω ein Galoischer Erweiterungskörper von P , der den Nullstellenkörper $P(\xi_1, \dots, \xi_n)$ umfaßt. Es seien $\{\xi_1^{(\alpha)}, \dots, \xi_n^{(\alpha)}\}$ ($\alpha = 1, \dots, l$) die zu ξ_1, \dots, ξ_n konjugierten Wertsysteme in Ω (oder die Nullstellen von \mathfrak{p} in Ω). Dann zerfällt das von \mathfrak{p} erzeugte Ideal \mathfrak{p}_Ω in \mathfrak{K}_Ω in lauter verschiedene Primideale:*

$$\mathfrak{p}_\Omega = [\mathfrak{p}^{(1)}, \dots, \mathfrak{p}^{(l)}] = \mathfrak{p}^{(1)} \cdot \mathfrak{p}^{(2)} \dots \mathfrak{p}^{(l)},$$

wo

$$\mathfrak{p}^{(\alpha)} = (x_1 - \xi_1^{(\alpha)}, \dots, x_n - \xi_n^{(\alpha)}).$$

Beweis. Da $\mathfrak{p}^{(1)}, \dots, \mathfrak{p}^{(l)}$ paarweise teilerfremd sind, ist Durchschnitt = Produkt:

$$[\mathfrak{p}^{(1)}, \dots, \mathfrak{p}^{(l)}] = \mathfrak{p}^{(1)} \cdot \mathfrak{p}^{(2)} \dots \mathfrak{p}^{(l)}.$$

Das Produkt $\mathfrak{p}^{(1)} \dots \mathfrak{p}^{(l)}$ ist mit allen seinen Konjugierten identisch, und wird erzeugt von Polynomen, deren Koeffizienten von erster Art in bezug auf P sind (sie hängen nämlich nur von den $\xi_i^{(\alpha)}$ ab). Also wird es von P -Polynomen erzeugt (Satz 24). Ebenso wird \mathfrak{p}_Ω von P -Polynomen, nämlich von den Polynomen von \mathfrak{p} erzeugt. Um also den Satz zu beweisen, genügt es, zu zeigen, daß jedes P -Polynom in $[\mathfrak{p}^{(1)}, \dots, \mathfrak{p}^{(l)}]$ liegt sobald es in \mathfrak{p} liegt, und umgekehrt.

Ein P -Polynom f gehört dann nur zu \mathfrak{p} , wenn

$$f(\xi_1, \dots, \xi_n) = 0,$$

da ja $\{\xi_1, \dots, \xi_n\}$ eine allgemeine Nullstelle von \mathfrak{p} ist. Ist $f(\xi_1, \dots, \xi_n) = 0$, so ist allgemeiner für jedes α

$$f(\xi_1^{(\alpha)}, \dots, \xi_n^{(\alpha)}) = 0,$$

also

$$f \equiv 0 \pmod{\mathfrak{p}^{(\alpha)}}$$

für alle α , also

$$f \equiv 0 \pmod{[\mathfrak{p}^{(1)}, \dots, \mathfrak{p}^{(l)}]},$$

und derselbe Schluß gilt auch in umgekehrter Richtung. Damit ist die Behauptung bewiesen.

Bei Erweiterung des Grundkörpers P ändern sich nicht nur die Ideale, sondern auch ihre Mannigfaltigkeiten.

Sei m ein Ideal und M seine Mannigfaltigkeit im $C_n(P)$, d. h. die Gesamtheit einer Nullstelle im Körper P . Sei wieder m_Ω das von m in \mathfrak{R}_Ω erzeugte Ideal. Wir bezeichnen mit M_Ω seine Mannigfaltigkeit. Sie besteht aus den Nullstellen des Ideals m_Ω im Körper Ω , oder, was auf dasselbe hinauskommt, aus den Nullstellen der erzeugenden Polynome des Ideals, d. h. aus den Nullstellen von m im Körper Ω .

Ist insbesondere P algebraisch-abgeschlossen, und ist \mathfrak{n} das zu M gehörige Ideal (die Gesamtheit aller P -Polynome, die auf M verschwinden), so ist nach dem Hilbertschen Nullstellensatz

$$\mathfrak{n}^e \equiv 0(m), \quad m \equiv 0(\mathfrak{n})$$

also auch

$$(\mathfrak{n}_\Omega)^e \equiv 0(m_\Omega), \quad m_\Omega \equiv 0(\mathfrak{n}_\Omega),$$

mithin stimmen nicht nur die Mannigfaltigkeiten von m und \mathfrak{n} , sondern auch die von m_Ω und \mathfrak{n}_Ω überein. Da \mathfrak{n} unabhängig von m durch M allein bestimmt wird, so werden auch \mathfrak{n}_Ω und M_Ω durch M allein bestimmt: M_Ω hängt nur von M , nicht von m ab.

§ 6.

Schnittpunkte einer irreduziblen Mannigfaltigkeit mit allgemeinen linearen Räumen.

Allen folgenden Untersuchungen werde ein fester algebraisch-abgeschlossener Körper P , für den man sich etwa den Körper der komplexen Zahlen denken kann, zugrunde gelegt.

Dem Körper P sollen abzählbar-unendlich-viele Unbestimmte adjungiert werden; zum so entstehenden rationalen Funktionenkörper sei Ω ein algebraisch-abgeschlossener Erweiterungskörper³¹⁾. Ω hat dann die Eigenschaft, daß es immer, wenn im Laufe der Untersuchung endlichviele Größen aus Ω verwendet worden sind, noch beliebig viele neue, von diesen Größen unabhängige Unbestimmte in Ω gibt. Die Zugrundelegung des ein für allemal konstruierten Körpers Ω erspart uns also die immer erneute Adjunktion von Unbestimmten und alle Konstruktionen von algebraischen Erweiterungskörpern. Wenn im folgenden an irgendeiner Stelle „Unbestimmte aus Ω “ eingeführt werden, so sind damit immer gemeint solche Unbestimmte von Ω , die von allen bis dahin verwendeten Größen aus Ω algebraisch-unabhängig sind.

³¹⁾ Steinitz, § 21, Satz 9 (S. 287).

Außerdem hat Ω die folgende Eigenschaft: wenn irgendein Erweiterungskörper von P von endlichem Transzendenzgrad³²⁾ gegeben ist, so enthält Ω einen damit äquivalenten Unterkörper³³⁾. Insbesondere enthält Ω zu jedem Primideal aus einem Polynombereich $P[x_0, \dots, x_n]$ einen Nullstellenkörper³⁴⁾.

Unter einem *allgemeinen linearen $n-r$ Raum* im $P_n(\Omega)$ werde verstanden eine solche $(n-r)$ -dimensionale Mannigfaltigkeit in $P_n(\Omega)$, die aus den Nullstellen von r Linearformen

$$l_i = \sum u_{ik} x_k$$

mit *unbestimmten* Koeffizienten aus Ω besteht.

Es sei ein für allemal M eine irreduzible r -dimensionale Mannigfaltigkeit in $P_n(P)$, und \mathfrak{p} das zugehörige Primideal in $P[x_0, \dots, x_n]$. \mathfrak{p}_Ω und M_Ω seien definiert wie im § 4. Wir werden in diesem Paragraphen sehen, daß M mit einem allgemeinen linearen $n-r-1$ Raum, keine Schnittpunkte, mit einem allgemeinen linearen $n-r$ Raum aber endlichviele Schnittpunkte hat. Die Anzahl dieser Schnittpunkte heißt der *Grad* von M .

Wie man diese Schnittpunkte zu finden hat, ist in der Eliminationstheorie schon wiederholt untersucht worden. Man verfährt dabei entweder so (Kronecker, Hentzelt), daß man den Unbestimmten x_1, \dots, x_n zuvor einer allgemeinen linearen Transformation unterwirft und dann diejenigen Punkte von M sucht, in denen etwa x_{n-r+1}, \dots, x_n gegebene (allgemeine) Werte haben; oder man sucht direkt durch Resultantenbildung (Mertens) die gemeinsamen Nullstellen von \mathfrak{p} und $(n-r)$ allgemeinen Linearformen.

Die Sätze über die Schnittpunkte von M mit allgemeinen linearen Räumen, die wir hier brauchen werden, kommen daher in der Eliminationstheorie, vor allem in der Hentzelschen, größtenteils schon vor, jedoch nicht in der idealtheoretischen Fassung, die für unsere Untersuchung nötig ist, und außerdem verquickt mit Eliminantanten, Resolventen und Resultantenformen, die wir nicht brauchen.

Satz 26. *Ein allgemeiner linearer $n-r-1$ Raum in $P_n(\Omega)$ hat mit M_Ω keine Schnittpunkte.*

Beweis. Der $n-r-1$ Raum bestehe aus den Nullstellen der Formen

$$l_i = \sum_0^n u_{ik} x_k \quad (i = 1, \dots, r+1).$$

³²⁾ Steinitz, § 23 (S. 299).

³³⁾ Wenn man nämlich dem gegebenen Erweiterungskörper von endlichem Transzendenzgrad abzählbar unendlich viele Unbestimmte adjungiert und algebraisch abschließt, so entsteht ein zu Ω äquivalenter Erweiterungskörper von P (Steinitz § 22, Satz 5; § 21, Satz 9).

³⁴⁾ W., § 1, 4.

Gesetzt, es wäre $X = \{\xi_0, \dots, \xi_n\}$ ein Schnittpunkt. Wir numerieren die Koordinaten so, daß $\xi_0 \neq 0$, und können dann annehmen $\xi_0 = 1$. Da X Nullstelle von p ist, so ist $\bar{X} = \{\xi_1, \dots, \xi_n\}$ Nullstelle von $\bar{p} = p_{(x_0=1)}$ (§ 3). Da \bar{p} die Dimension r hat, so hat das System ξ_1, \dots, ξ_n höchstens den Transzendenzgrad r in bezug auf P , also um so mehr in bezug auf $P(u_{11}, u_{12}, \dots, u_{r+1, n})$.

Da nun vermöge

$$u_{i0} + \sum_1^n u_{ik} \xi_k = 0$$

die u_{i0} sich durch die ξ_k und die übrigen u_{ik} ausdrücken lassen, so hat auch das System $u_{10}, \dots, u_{r+1, 0}$ höchstens den Transzendenzgrad r in bezug auf $P(u_{11}, u_{12}, \dots, u_{r+1, n})$, entgegen der Voraussetzung, daß die u_{i0} Unbestimmte sind.

Hilfssatz 1. Voraussetzung. In Ω sei ξ_1, \dots, ξ_n ein System vom Transzendenzgrad r in bezug auf P . Es seien u_{ik} ($i, k = 1, \dots, n$) Unbestimmte in Ω , unabhängig von den ξ , und es sei gesetzt

$$(1) \quad \xi_i^* = \sum_1^n u_{ik} \xi_k \quad (i = 1, \dots, n).$$

Behauptung 1. ξ_1^*, \dots, ξ_r^* sind algebraisch-unabhängig in bezug auf $P(u)$ ³⁵.

2. ξ_1^*, \dots, ξ_n^* sind algebraisch und von erster Art³⁶ in bezug auf $P(u, \xi_1^*, \dots, \xi_r^*)$ ³⁷.

Beweis. Zunächst eine Vorbemerkung. Eine beliebige Permutation der ersten Indizes der u_{ik} erzeugt einen Automorphismus des Körpers $P(u, \xi)$. Dabei werden die ξ_i^* ebenso permutiert wie die u_{ik} . Wenn wir also zeigen können, daß es überhaupt r linear-unabhängige $\xi_{\lambda_1}^*, \dots, \xi_{\lambda_r}^*$ gibt, so folgt daraus vermöge dieser Automorphismen, daß auch ξ_1^*, \dots, ξ_r^* algebraisch-unabhängig sind (Behauptung 1); wenn wir weiter zeigen können, daß mindestens ein weiteres $\xi_{\lambda_{r+1}}^*$ algebraisch und von erster Art in bezug auf $\xi_{\lambda_1}^*, \dots, \xi_{\lambda_r}^*$ ist, so folgt ebenso vermöge dieser Automorphismen, daß $\xi_{r+1}^*, \dots, \xi_n^*$ alle algebraisch und von erster Art in bezug auf ξ_1^*, \dots, ξ_r^* sind (Behauptung 2).

Der Körper $P(u, \xi^*) = P(u, \xi)$ hat den Transzendenzgrad r in bezug auf $P(u)$. Also gibt es ein irreduzibles System $\xi_{\lambda_1}^*, \dots, \xi_{\lambda_r}^*$, von dem jedes weitere $\xi_{\lambda_{r+1}}^*$ algebraisch abhängt.

³⁵) Das Symbol $P(u)$ soll eine Abkürzung für $P(u_{11}, u_{12}, \dots, u_{n\pi})$ sein.

³⁶) Steinitz, § 13.

³⁷) Der Satz ist enthalten im Hilfssatz V von E. Noether, Eliminationstheorie und allgemeine Idealtheorie, Math. Annalen 90 (1923), S. 245. Der dort gegebene Beweis ist aber fehlerhaft. Der hier gegebene Beweis gilt auch für jenen Hilfssatz V.

Für Charakteristik Null³⁸⁾ ist jede algebraische Größe zugleich von erster Art, also irgendein $\xi_{\lambda_r+1}^*$ algebraisch und von erster Art in bezug auf $P(u, \xi_{\lambda_1}^*, \dots, \xi_{\lambda_r}^*)$, womit nach der Vorbemerkung alles erledigt ist.

Für Charakteristik p bilde man die Gleichung niedrigsten Grades in $\xi_{\lambda_r+1}^*$, welche $\xi_{\lambda_r+1}^*$ mit $\xi_{\lambda_1}^*, \dots, \xi_{\lambda_r}^*$ und den u_{ik} verknüpft. Auf Grund der Vorbemerkung ist nur zu beweisen, daß diese Gleichung in bezug auf mindestens eins der $\xi_{\lambda_i}^*$ von erster Art ist.

Wäre sie das nicht, d. h. enthielte sie $\xi_{\lambda_1}^*, \dots, \xi_{\lambda_r+1}^*$ nur in der p -ten Potenz, so könnte man die linke Seite schreiben als Summe von Potenzprodukten

$$\prod_u(u) = u_{11}^{\mu_{11}} u_{12}^{\mu_{12}}, \dots, u_{nn}^{\mu_{nn}} \quad (0 \leq \mu_{ik} < p)$$

mit Koeffizienten, die nur von

$$u_{11}^p, u_{12}^p, \dots, u_{nn}^p, \dots, \xi_{\lambda_1}^{*p}, \dots, \xi_{\lambda_r+1}^{*p}$$

abhängen. Die Gleichung würde also lauten:

$$(2) \quad \sum_{\mu} \prod_{\mu}(u) f_{\mu}(u_{11}^p, u_{12}^p, \dots, u_{nn}^p, \xi_{\lambda_1}^{*p}, \dots, \xi_{\lambda_r+1}^{*p}) = 0.$$

Da die $\prod_{\mu}(u)$ in bezug auf den Körper $P(u_{11}^p, u_{12}^p, \dots, u_{nn}^p, \xi_1, \dots, \xi_n)$, in dem auch die ξ_i^{*p} liegen, linear-unabhängig sind, so folgt aus (2):

$$f_{\mu}(u_{11}^p, u_{12}^p, \dots, u_{nn}^p, \xi_{\lambda_1}^{*p}, \dots, \xi_{\lambda_r+1}^{*p}) = 0$$

für alle μ . Mindestens eins der f_{μ} muß $\xi_{\lambda_r+1}^{*p}$ wirklich enthalten. Nun ist, da P algebraisch-abgeschlossen ist, f_{μ} eine p -te Potenz eines Polynoms g_{μ} von niedrigerem Grad; also erhält man die Gleichung

$$g_{\mu}(u_{11}, u_{12}, \dots, u_{nn}, \xi_{\lambda_1}^*, \dots, \xi_{\lambda_r+1}^*) = 0,$$

die in $\xi_{\lambda_r+1}^*$ von niedrigerem Grad ist wie die ursprüngliche, entgegen der Voraussetzung.

Hilfssatz 2. *Es sei \mathfrak{p} ein Primideal in $P[x_1, \dots, x_n]$, und $\{\xi_1, \dots, \xi_n\}$ seine allgemeine Nullstelle. Ist nun $r \leq n$, so ist das Ideal*

$$\mathfrak{p}_r = (\mathfrak{p}, x_1 - \xi_1, \dots, x_r - \xi_r)$$

in $P(\xi_1, \dots, \xi_r)[x_1, \dots, x_n]$ ebenfalls prim, und es hat wiederum die allgemeine Nullstelle $\{\xi_1, \dots, \xi_n\}$.

Beweis. Daß $\{\xi_1, \dots, \xi_n\}$ eine Nullstelle von $(\mathfrak{p}, x_1 - \xi_1, \dots, x_r - \xi_r)$ ist, ist klar, denn dieser Punkt ist eine Nullstelle sowohl von \mathfrak{p} , wie von $x_1 - \xi_1, \dots, x_r - \xi_r$. Bleibt also zu zeigen, daß die Nullstelle allgemein ist, d. h. daß, wenn f ein Polynom aus $P(\xi_1, \dots, \xi_r)[x_1, \dots, x_n]$ ist, aus

$$f(\xi_1, \dots, \xi_n) = 0$$

³⁸⁾ Steinitz, § 4.

folgt

$$f \equiv 0 (\mathfrak{p}_r).$$

f hängt rational von ξ_1, \dots, ξ_r ab. Ohne Beschränkung der Allgemeinheit kann man f ganzrational in ξ_1, \dots, ξ_r annehmen:

$$f = F(\xi_1, \dots, \xi_r, x_1, \dots, x_n).$$

Ersetzt man in F überall ξ_i durch x_i , so entsteht das Polynom $F(x_1, \dots, x_r, x_1, \dots, x_n)$, das noch immer die Nullstelle ξ_1, \dots, ξ_n hat, also, da es nur von den x abhängt, in \mathfrak{p} liegt. Weiter ist

$$F(\xi_1, \dots, \xi_r, x_1, \dots, x_n) \equiv F(x_1, \dots, x_r, x_1, \dots, x_n) \pmod{(x_1 - \xi_1, \dots, x_r - \xi_r)},$$

also

$$f \equiv 0 (\mathfrak{p}, x_1 - \xi_1, \dots, x_r - \xi_r).$$

Also ist die Nullstelle allgemein, und das Ideal \mathfrak{p}_r prim.

Eine kleine Spezialisierung der Voraussetzungen führt unmittelbar zu:

Hilfssatz 3. *Es sei \mathfrak{p} ein r -dimensionales Primideal in $\mathbb{P}[x_1, \dots, x_n]$, und $\{\xi_1, \dots, \xi_n\}$ seine allgemeine Nullstelle. Sind nun ξ_{r+1}, \dots, ξ_n algebraisch und von erster Art in bezug auf $\mathbb{P}(\xi_1, \dots, \xi_r)$, so ist für unbestimmte t_1, \dots, t_r das Ideal*

$$\mathfrak{c} = (\mathfrak{p}_{\mathbb{P}(t)}, x_1 - t_1, \dots, x_r - t_r)$$

in $\mathbb{P}(t)[x_1, \dots, x_n]$ prim, nulldimensional und von erster Art.

Beweis. Zunächst sind ξ_1, \dots, ξ_r algebraisch-unabhängig, also gilt die Isomorphie

$$\mathbb{P}(\xi_1, \dots, \xi_r) \cong \mathbb{P}(t_1, \dots, t_r).$$

Wir können demnach in der Behauptung alle t_i durch ξ_i ersetzen. Dann geht \mathfrak{c} über in das Ideal \mathfrak{p}_r , von dem im Hilfssatz 2 nachgewiesen wurde, daß es prim ist und die allgemeine Nullstelle ξ_1, \dots, ξ_n hat. Diese allgemeine Nullstelle ist nach Voraussetzung algebraisch und von erster Art in bezug auf den Grundkörper $\mathbb{P}(\xi_1, \dots, \xi_r)$; also ist \mathfrak{p}_r nulldimensional und von erster Art. Das gleiche gilt vermöge der Isomorphie für \mathfrak{c} .

Auf Satz 25 und den Hilfssätzen 1 und 3 fußt das Hauptergebnis dieses Paragraphen:

Satz 27. *Voraussetzung. Die r -dimensionale Mannigfaltigkeit M in $P_n(\mathbb{P})$ gehöre zum homogenen Primideal \mathfrak{p} in $\mathbb{R} = \mathbb{P}[x_0, \dots, x_n]$. Es seien u_{ik} ($i, k = 1, \dots, n$) und t_i ($i = 1, \dots, r$) Unbestimmte in Ω . Es sei L der allgemeine lineare $n-r$ -Raum in $P_n(\Omega)$, bestehend aus den Nullstellen der r -Linearformen*

$$l_i = -t_i x_0 + \sum_1^n u_{ik} x_k \quad (i = 1, \dots, r).$$

Behauptung. 1. M_Ω hat mit L nur endlichviele Schnittpunkte $Y^{(1)}, \dots, Y^{(s)}$ ($s > 0$) in $P_n(\Omega)^{39}$. Wie schon bemerkt, nennt man die positive Zahl s den Grad von M .

2. Die Koordinaten dieser Schnittpunkte sind, bei passender Wahl der Proportionalitätsfaktoren, konjugiert und von erster Art in bezug auf $P(u, t) = P(u_{11}, \dots, u_{nn}, t_1, \dots, t_r)$.

3. Das Ideal $\bar{a} = (\bar{p}_\Omega, \bar{l}_1, \dots, \bar{l}_r)^{40}$ in $\Omega[x_1, \dots, x_n]$ ist Produkt von lauter verschiedenen Primidealen, die je zu einem der Punkte $\overline{Y^{(1)}}, \dots, \overline{Y^{(s)}}$ gehören.

Beweis. Wir definieren neue Variable x_i^* durch

$$x_i^* = \sum_1^n u_{ik} x_k \quad (i = 1, \dots, n).$$

Dann ist

$$\begin{aligned} P(u)[x_1, \dots, x_n] &= P(u)[x_1^*, \dots, x_n^*], \\ \Omega[x_1, \dots, x_n] &= \Omega[x_1^*, \dots, x_n^*]. \end{aligned}$$

Ist nun ξ_1, \dots, ξ_n eine allgemeine Nullstelle von $\bar{p}_{P(u)}$, so ist, wenn man x_1^*, \dots, x_n^* als Variable desselben Polynombereichs $P(u)[x_1, \dots, x_n]$ betrachtet, eine allgemeine Nullstelle von $\bar{p}_{P(u)}$ gegeben durch $\{\xi_1^*, \dots, \xi_n^*\}$, wo

$$\xi_i^* = \sum u_{ik} \xi_k.$$

Da nun nach Hilfssatz 1 die Größen $\xi_{r+1}^*, \dots, \xi_n^*$ algebraisch und von erster Art in bezug auf $P(\xi_1^*, \dots, \xi_r^*)$ sind, so läßt sich Hilfssatz 3 anwenden, d. h. das Ideal

$$(\bar{p}_{P(u,t)}, x_1^* - t_1, \dots, x_r^* - t_r)$$

in $P(u, t)[x_1^*, \dots, x_n^*]$ ist prim, nulldimensional und von erster Art. Schreibt man für die x^* ihre Bedeutung, so sieht man, daß dieses Ideal dasselbe ist wie

$$(\bar{p}_{P(u,t)}, \bar{l}_1, \dots, \bar{l}_r).$$

Durch Erweiterung des Körpers $P(u, t)$ zu Ω geht es also über in \bar{a} . Bei dieser Erweiterung bleibt es möglicherweise nicht mehr prim, aber nach Satz 25 zerfällt es in lauter verschiedene Primideale

$$\bar{a} = \prod_1^s \bar{a}^{(\alpha)}; \quad \bar{a}^{(\alpha)} = (x_1 - \eta_1^{(\alpha)}, \dots, x_n - \eta_n^{(\alpha)}),$$

³⁹) Dieser Satz zeigt, daß der Mertenssche Begriff der Stufenzahl eines homogenen Gleichungssystems, nämlich die Anzahl der allgemeinen linearen Gleichungen, die man hinzufügen muß, um eine endliche Anzahl von Lösungsklassen zu erhalten, mit der Höchstdimension der zugehörigen Mannigfaltigkeit übereinstimmt.

⁴⁰) Die Bedeutung der Querstriche ist dieselbe wie in § 3.

wo die Punkte $\overline{Y^{(\alpha)}} = \{\eta_1^{(\alpha)}, \dots, \eta_n^{(\alpha)}\}$ konjugiert in bezug auf $P(u, t)$ sind. Sie sind sicher die einzigen Schnittpunkte von M und L , bei denen die erste Koordinate = 1 gewählt werden kann, weil jedem solchen Schnittpunkt Y ein Punkt \overline{Y} in $C_n(\Omega)$ zugeordnet werden kann, welcher eine gemeinsame Nullstelle von $\overline{p_\Omega}$ und $\overline{l_1}, \dots, \overline{l_r}$, also unter den Punkten $\overline{Y^{(\alpha)}}$ schon vertreten ist. Es fragt sich also nur, ob es nicht Schnittpunkte gibt, deren erste Koordinate verschwindet. Ein solcher Punkt wäre Nullstelle von (p, x_0) . Die Mannigfaltigkeit von (p, x_0) ist aber, als echter Teil einer irreduziblen r -dimensionalen Mannigfaltigkeit, höchstens $(r-1)$ -dimensional, hat also nach Satz 26 mit L keine Punkte gemein. Also kann ein solcher Schnittpunkt nicht existieren.

Korollar. Aus der hier abgeleiteten Gleichung

$$\bar{a} = \prod_1^s \overline{a^{(\alpha)}}$$

kann man nach Satz 11 zwei Relationen herleiten zwischen den H -Idealen, definiert durch

$$(3) \quad \begin{cases} u &= (x_0, \dots, x_n), \\ a &= (p_\Omega, l_1, \dots, l_r), \\ a^{(\alpha)} &= (x_1 - \eta_1^{(\alpha)} x_0, \dots, x_n - \eta_n^{(\alpha)} x_0), \end{cases}$$

nämlich die Relationen:

$$(4) \quad u^e a \equiv 0 \left(\prod_1^s a^{(\alpha)} \right),$$

$$(5) \quad u^\sigma \prod_1^r a^{(\alpha)} \equiv 0(a).$$

Diese Gleichungen bilden den bequemsten algebraischen Ausdruck für die Tatsache, daß die Ideale $a = (p_\Omega, l_1, \dots, l_r)$ und $\prod_1^s a^{(\alpha)}$ in allen ihren Primärkomponenten, bis auf die zum Primideal u gehörigen, übereinstimmen.

§ 7.

Das Verhalten der Ideale bei linearen Transformationen.

Sei wieder $R_\Omega = \Omega[x_0, \dots, x_n]$, und sei $V = (v_{ik})$ ($i, k = 0, \dots, n$) eine Matrix mit Elementen aus Ω . Unter dem *Transformierten* fV eines Polynoms f mit der Matrix V soll verstanden werden das Polynom fV , das aus f entsteht nach Ersetzung von x_i durch $\sum v_{ik} x_k$. Unter dem *Transformierten* mV eines Ideals m mit der Matrix V soll verstanden werden das Ideal, das erzeugt wird von den transformierten Polynomen des Ideals m .

Ist $m = (f_1, \dots, f_r)$, so lassen die Transformaten des Polynoms

$$f = \sum a_i f_i$$

von m sich in die Gestalt

$$fV = \sum (a_i V \cdot f_i V)$$

schreiben, also wird das transformierte Ideal erzeugt von $f_1 V, \dots, f_r V$:

$$mV = (f_1 V, \dots, f_r V),$$

d. h.: *ein Ideal wird transformiert, indem man seine Basis transformiert.*

Daraus folgt direkt

$$(a, b)V = (aV, bV).$$

$$a \cdot b \cdot V = aV \cdot bV.$$

Weiter folgt aus $a \equiv 0(b)$ immer $aV \equiv 0(bV)$.

Ist m ein H -Ideal, also die Basiselemente f_1, \dots, f_r Formen, so sind auch $f_1 V, \dots, f_r V$ Formen, also mV ein H -Ideal.

Ist M eine Mannigfaltigkeit, m das zugehörige Ideal, mV das transformierte Ideal, so bezeichnen wir die Mannigfaltigkeit von mV mit MV und nennen sie: *die Transformierte von M .*

Ist V nichtsingulär ($|v_{ik}| \neq 0$), so kann man aus den transformierten Idealen die ursprünglichen zurückgewinnen, indem man mit der inversen Matrix V^{-1} transformiert.

Wir betrachten aber im folgenden insbesondere singuläre Matrizes (geometrisch: „ausgeartete Transformationen“), und zwar Matrizes vom Rang $n - r + 1$ ($0 \leq r \leq n$).

Jede Matrix V vom Rang $n - r + 1$ läßt sich bekanntlich in der Gestalt

$$V = A^{-1}TB$$

darstellen, wo A und B nichtsinguläre Matrizes sind, und wo T eine Diagonalmatrix

$$\begin{cases} t_{ik} = 1 & \text{für } i = k \leq n - r \\ t_{ik} = 0 & \text{sonst.} \end{cases}$$

ist. Aus diesem Grund verstehen wir unter der *allgemeinen Matrix vom Rang $n - r + 1$* die Matrix

$$W = S^{-1}TU,$$

wo T wie oben definiert ist, und wo S und U Matrizes mit unbestimmten Koeffizienten sind.

Unser Ziel ist, das Verhalten des $(r + 1)$ -dimensionalen H -Ideals p_Ω bei Transformation mit der allgemeinen Matrix vom Rang $(n - r + 1)$ zu untersuchen. Wozu diese Untersuchung nötig ist, wird sich aus dem nächsten Paragraphen ergeben.

Die Elemente der Matrizes S und U seien s_{ik} , u_{ik} , die in der inversen Matrizes s'_{ik} , u'_{ik} . Wir setzen

$$l_i = \sum s_{ik} x_k \quad (i = n - r + 1, \dots, n).$$

Dann ist

$$l_i S^{-1} = \sum \sum s_{ik} s'_{ki} x_l = x_i$$

also, wenn $\alpha = (p_\Omega, l_{n-r+1}, \dots, l_n)$ gesetzt wird,

$$\alpha S^{-1} = (p_\Omega S^{-1}, x_{n-r+1}, \dots, x_n)$$

Bei der Substitution $x_i \rightarrow \sum t_{ik} x_k$ gehen x_{n-r+1}, \dots, x_n in Null über. Also folgt

$$\alpha S^{-1} T = (p_\Omega S^{-1} T, 0, \dots, 0) = p_\Omega S^{-1} T$$

$$\alpha S^{-1} T U = p_\Omega S^{-1} T U$$

$$(1) \quad \alpha W = p_\Omega W$$

Für das Ideal α gelten die Gleichungen (4), (5) von § 6, da es natürlich nicht darauf ankommt ob die zur Definition von α verwandten Linearformen mit l_1, \dots, l_r (vgl. (3) § 6) oder mit l_{n-r+1}, \dots, l_n bezeichnet werden. Man hat also

$$\begin{cases} u^\epsilon \alpha \equiv 0 \left(\prod_1^s \alpha^{(\alpha)} \right), \\ u^\sigma \prod_1^s \alpha^{(\alpha)} \equiv 0(\alpha), \end{cases}$$

wobei s der Grad von M ist, und u , $\alpha^{(\alpha)}$ die Bedeutung (3) § 6 haben.

Durch Transformation mit W folgt daraus, wenn man (1) beachtet:

$$(2) \quad (u W)^\epsilon \cdot p_\Omega W \equiv 0 \left(\prod_1^s (\alpha^{(\alpha)} W) \right),$$

$$(3) \quad (u W)^\sigma \cdot \prod_1^s (\alpha^{(\alpha)} W) \equiv 0(p_\Omega W).$$

Die Ideale $u W$, $\alpha^{(\alpha)} W$ sind leicht explizite hinzuschreiben. Die untransformierten Ideale u , $\alpha^{(\alpha)}$ haben, wie ihre Definitionen (3) § 6 zeigen, die allgemeinen Nullstellen

$$O = \{0, \dots, 0\},$$

bzw.

$$Y^{(\alpha)} = \{\mu, \mu \eta_1^{(\alpha)}, \dots, \mu \eta_n^{(\alpha)}\}.$$

Die letztere kann man noch etwas symmetrischer schreiben, indem man $\eta_0^{(\alpha)} = 1$ setzt. Die transformierten Ideale $u S^{-1}$, $\alpha^{(\alpha)} S^{-1}$ haben demnach allgemeine Nullstellen, die man durch Transformation mit S gewinnt:

$$\begin{cases} O S = \{0, \dots, 0\} \\ Y^{(\alpha)} S = \{\mu \lambda_0^{(\alpha)}, \dots, \mu \lambda_k^{(\alpha)}\}, \end{cases}$$

wo

$$\lambda_i^{(\alpha)} = \sum_0^n \varepsilon_{ik} \eta_k^{(\alpha)}$$

gesetzt ist.

Also hat man für uS^{-1} und $\alpha^{(\alpha)}S^{-1}$ die Basisdarstellungen

$$\begin{cases} uS^{-1} = (x_0, \dots, x_n), \\ \alpha^{(\alpha)}S^{-1} = (\lambda_0^{(\alpha)}x_1 - \lambda_1^{(\alpha)}x_0; \dots; \lambda_0^{(\alpha)}x_n - \lambda_n^{(\alpha)}x_0). \end{cases}$$

Nun waren die Punkte $Y^{(\alpha)}$ ihrer Bedeutung nach (vgl. die Definition der $\eta_k^{(\alpha)}$ in § 6) Nullstellen von $l_i = \sum \varepsilon_{ik} x_k$ ($i > n - r$). Also wird $\lambda_i^{(\alpha)} = 0$ für $i > n - r$. Mithin ist

$$\alpha^{(\alpha)}S^{-1} = (\lambda_0^{(\alpha)}x_1 - \lambda_1^{(\alpha)}x_0; \dots; \lambda_0^{(\alpha)}x_{n-r} - \lambda_{n-r}^{(\alpha)}x_0; x_{n-r+1}; \dots; x_n).$$

Transformiert man jetzt weiter mit T und U , so folgt

$$(4) \quad \begin{cases} uS^{-1}T = (x_0, \dots, x_{n-r}), \\ \alpha^{(\alpha)}S^{-1}T = (\lambda_0^{(\alpha)}x_1 - \lambda_1^{(\alpha)}x_0; \dots; \lambda_0^{(\alpha)}x_{n-r} - \lambda_{n-r}^{(\alpha)}x_0), \\ uW = uS^{-1}TU = \left(\sum_0^n u_{0k}x_k, \dots, \sum_0^n u_{n-r,k}x_k \right). \end{cases}$$

$$(5) \quad \begin{aligned} \alpha^{(\alpha)}W &= \alpha^{(\alpha)}S^{-1}TU \\ &= \left(\sum_0^n (\lambda_0^{(\alpha)}u_{1k} - \lambda_1^{(\alpha)}u_{0k})x_k; \dots; \sum_0^n (\lambda_0^{(\alpha)}u_{n-r,k} - \lambda_{n-r}^{(\alpha)}u_{0k})x_k \right). \end{aligned}$$

Aus (4) sieht man, daß uW gehört zu einem allgemeinen r^{-1} Raum L in $P_n(\Omega)$, und aus (5) folgt, daß die Ideale $\alpha^{(\alpha)}W$ gehören zu linearen Räumen $L^{(\alpha)}$ ($\alpha = 1, \dots, s$), welche L enthalten. Die $L^{(\alpha)}$ sind alle verschieden, denn da die Punkte $Y^{(\alpha)} = \{\eta_0^{(\alpha)}, \dots, \eta_n^{(\alpha)}\}$ verschieden sind, so sind die transformierten Punkte $Y^{(\alpha)}S = \{\lambda_0^{(\alpha)}, \dots, \lambda_{n-r}^{(\alpha)}, 0, \dots, 0\}$ es auch, und sie sind die Schnittpunkte der Räume $L^{(\alpha)}U^{-1}$ mit dem Raum $\xi_{n-r+1} = \dots = \xi_n = 0$.

Die Mannigfaltigkeit von $\prod_1^s (\alpha^{(\alpha)}W)$ ist die Vereinigung der linearen Räume $L^{(1)}, \dots, L^{(s)}$. Sie enthält die Mannigfaltigkeit von uW . Also folgt aus (3), daß sie auch die Mannigfaltigkeit von $p_\Omega W$ enthält.

Umgekehrt enthält die Mannigfaltigkeit von $p_\Omega W$ die Räume $L^{(1)}, \dots, L^{(s)}$, denn aus (2) folgt

$$(uW)^e \cdot (p_\Omega W) \equiv 0 (\alpha^{(\alpha)}W)$$

also, da $\alpha^{(\alpha)}W$ prim ist und $uW \not\equiv 0 (\alpha^{(\alpha)}W)$:

$$(6) \quad p_\Omega W \equiv 0 (\alpha^{(\alpha)}W).$$

Damit ist bewiesen:

Satz 28. Die Mannigfaltigkeit von $p_\Omega W$ ist die Vereinigung der linearen Räume $L^{(1)}, \dots, L^{(s)}$.

Dieser Satz enthält die für das folgende grundlegende, in der Einleitung erwähnte Tatsache, daß die singuläre Transformation W die Mannigfaltigkeit M_Ω in eine in lineare Räume zerfallende überführt.

Eine andere Folge der Gleichungen (2), (3) ist:

Satz 29. Ist $Y = \{\eta_0, \dots, \eta_n\}$ ein Punkt von $L^{(\beta)}$, der nicht in L liegt, so ist

$$(\mathfrak{p}_\Omega W)_Y = \alpha^{(\beta)} W^{41}).$$

Beweis. Aus (6) folgt:

$$(7) \quad (\mathfrak{p}_\Omega W)_Y \equiv 0 (\alpha^{(\beta)} W)_Y$$

und aus (3):

$$(8) \quad \left((u W)^e \prod_1^s (\alpha^{(\alpha)} W) \right)_Y \equiv 0 (\mathfrak{p}_\Omega W)_Y.$$

Da Y nicht Nullstelle von $u W$ ist, und ebensowenig von $\alpha^{(\alpha)} W (\alpha \neq \beta)$, so hat man

$$\begin{aligned} (u W)_Y &= (1), \\ (\alpha^{(\alpha)} W)_Y &= (1) \quad (\alpha \neq \beta). \end{aligned}$$

Durch wiederholte Anwendung von (6) (§ 2) auf die linke Seite von (8) folgt jetzt

$$(9) \quad (\alpha^{(\beta)} W)_Y \equiv 0 (\mathfrak{p}_\Omega W)_Y$$

Aus (7) und (9):

$$(\mathfrak{p}_\Omega W)_Y = (\alpha^{(\beta)} W)_Y = \alpha^\beta W,$$

weil Y Nullstelle des Primideals $\alpha^{(\beta)} W$ ist.

§ 8

Der verallgemeinerte Bézoutsche Satz. Multiplizitäten.

Nach den Voruntersuchungen der vorangehenden Paragraphen sind wir instande, den verallgemeinerten Bézoutschen Satz scharf zu formulieren und seinen Beweis in Angriff zu nehmen.

Hauptsatz. *Es sei P ein algebraisch-abgeschlossener Körper. Es seien M und M' irreduzible algebraische Mannigfaltigkeiten der Dimensionen r und $n-r$ im projektiven Raum $P_n(P)$. Es sei U eine Matrix mit unbestimmten (d. h. in bezug auf P algebraisch-unabhängigen) Koeffizienten u_{ik} im algebraisch-abgeschlossenen Erweiterungskörper Ω . Dann ist die Anzahl der Schnittpunkte von $M_\Omega U$ mit M'_Ω gleich dem Produkt der Gradzahlen von M und M' .*

Die Beschränkung auf irreduzible Mannigfaltigkeiten ist natürlich unwesentlich, da eine reduzible rein r -dimensionale bzw. rein $(n-r)$ -dimen-

⁴¹⁾ Für die Bedeutung des Symbols m_Y siehe § 3, Schluß.

sionale Mannigfaltigkeit sich aus ebensolchen irreduziblen aufbaut. Die irreduziblen Bestandteile von M haben miteinander nur Mannigfaltigkeiten von niedrigeren Dimensionszahlen gemein (W_1 § 4, 15), welche, wie man leicht zeigt, nach Transformation mittels U keine Schnittpunkte mit M' haben (vgl. den analogen Satz 26). Also ist die Anzahl der Schnittpunkte der reduziblen Mannigfaltigkeit $M_\Omega U$ mit M'_Ω gleich der Summe der Anzahlen der Schnittpunkte der irreduziblen Bestandteile von $M_\Omega U$ mit M'_Ω . Ebenso zerlegt man diese Anzahlen weiter nach den irreduziblen Bestandteilen von M' . Daraus folgt, daß der Satz für reduzible Mannigfaltigkeiten gilt sobald er für irreduzible bewiesen ist.

Sind $p = (f_1, \dots, f_r)$ und $p' = (f'_1, \dots, f'_l)$ die zu M und M' gehörigen Ideale, so sind die Schnittpunkte $X = \{\xi_0, \dots, \xi_n\}$ von $M_\Omega U$ und M'_Ω die Lösungsklassen des homogenen Gleichungssystems:

$$(1) \quad \begin{cases} f_1(\sum u_{ik} \xi_k) = 0 \\ \dots \dots \dots \\ f_r(\sum u_{ik} \xi_k) = 0 \\ f'_1(\xi) = 0 \\ \dots \dots \dots \\ f'_l(\xi) = 0 \end{cases}$$

im Körper Ω . Ein geometrisches Problem, dessen Lösung mit der Lösung eines homogenen Gleichungssystems äquivalent ist, heißt *Normalproblem*.

Ist V irgendeine spezielle Matrix in Ω , so werden die Schnittpunkte von $M_\Omega V$ mit M'_Ω gegeben durch die Lösungsklassen des spezialisierten Gleichungssystems

$$(2) \quad \begin{cases} f_1(\sum v_{ik} \xi_k) = 0 \\ \dots \dots \dots \\ f_r(\sum v_{ik} \xi_k) = 0 \\ f'_1(\xi) = 0 \\ \dots \dots \dots \\ f'_l(\xi) = 0 \end{cases}$$

Insbesondere gilt das, wenn V die Einheitsmatrix ist: $V = E$; das Gleichungssystem (2) gilt dann für die Schnittpunkte von M und M' .

Wenn ein Normalproblem in dieser Weise durch eine Spezialisierung von Unbestimmten u_{ik} aus einem allgemeineren Normalproblem abgeleitet ist, so gelten die folgenden Sätze⁴²⁾:

I. Sind $X^{(1)}, \dots, X^{(q)}$ q verschiedene Lösungsklassen des allgemeinen Normalproblems (1), so gibt es q (nicht notwendig verschiedene) Lösungsklassen $Y^{(1)}, \dots, Y^{(q)}$ des spezialisierten Normalproblems (2) derart, daß

⁴²⁾ W_2 §§ 3—5.

alle in den Koordinaten der $X^{(i)}$ homogenen Relationen

$$f(u, X^{(1)}, \dots, X^{(q)}) = 0$$

bei der Spezialisierung $u \rightarrow v$, $X \rightarrow Y$ bestehen bleiben.

Ein System $Y^{(1)}, \dots, Y^{(q)}$ mit dieser Eigenschaft heißt eine *relations-treue Spezialisierung* des Systems $X^{(1)}, \dots, X^{(q)}$.

II. Hat das spezialisierte Normalproblem nur endlichviele Lösungsklassen, so hat das allgemeine auch nur endlichviele verschiedene Lösungsklassen, etwa $X^{(1)}, \dots, X^{(q)}$. Alle relationstreuen Spezialisierungen $Y^{(1)}, \dots, Y^{(q)}$ von $X^{(1)}, \dots, X^{(q)}$ sind in diesem Fall bis auf die Reihenfolge miteinander identisch.

Die Zahl, die angibt wie oft eine gegebene Lösung Y in der in Satz II definierten relationstreuen Spezialisierung vorkommt, heißt die *Multiplizität* der Lösung Y des spezialisierten Normalproblems. Sie kann positiv oder Null sein.

III. Die Summe der Multiplizitäten aller Lösungen des spezialisierten Normalproblems ist gleich der Anzahl der Lösungen des allgemeinen Normalproblems (Prinzip der Erhaltung der Anzahl).

Wir wollen diese Sätze auf die durch (1) und (2) gegebenen Normalprobleme anwenden und wählen dabei insbesondere für V die in § 8 definierte Matrix W , die allgemeine Matrix vom Rang $n - r + 1$. Die Mannigfaltigkeit $M_\Omega W$ zerfällt dann in lineare Räume $L^{(1)}, \dots, L^{(s)}$, wo s der Grad von M ist. Die Schnittpunkte von $M_\Omega W$ und M'_Ω verteilen sich auf $L^{(1)}, \dots, L^{(s)}$. Da die linearen Räume $L^{(1)}, \dots, L^{(s)}$ allgemein sind, d. h. je durch r Linearformen mit algebraisch-unabhängigen Koeffizienten gegeben sind, so haben sie mit M'_Ω je so viele Schnittpunkte wie der Grad von M' beträgt. Ein solcher Schnittpunkt kann nur einem der Räume $L^{(a)}$ zugleich angehören, denn der Raum L , in dem sich je zwei $L^{(a)}, L^{(b)}$ schneiden, ist ein allgemeiner linearer $r-1$ -Raum, hat also mit M'_Ω keine Punkte gemein (Satz 26). Also hat die zerfallende Mannigfaltigkeit $M_\Omega W$ mit M'_Ω so viele Schnittpunkte, wie das Produkt der Gradzahlen von M und M' beträgt. Wenn man noch zeigen kann, daß alle diese Schnittpunkte die Multiplizität *eins* haben, so wird nach dem Prinzip der Erhaltung der Anzahl die Anzahl der Schnittpunkte von $M_\Omega U$ und M'_Ω ebenfalls gleich dem Produkt der Gradzahlen von M und M' sein, also der Hauptsatz bewiesen sein.

Dem Beweis dieser Tatsache sind die beiden folgenden Paragraphen gewidmet: in § 9 wird gezeigt werden, daß die Multiplizitäten der Schnittpunkte von $M_\Omega W$ und $M'_\Omega > 0$ sind, in § 10, daß sie ≤ 1 sind. Es wird dabei zugleich noch mehr bewiesen werden: eine Zusammenfassung der erhaltenen Resultate folgt in § 11.

§ 9.

Die Multiplizitäten sind nicht Null.

Hilfssatz 4. Sind f_1, \dots, f_r Polynome in $\Sigma[y_1, \dots, y_s]$, und algebraisch-abhängig in bezug auf $P(y_1, \dots, y_s)$, wo P ein Unterkörper von Σ ist, so bleiben sie algebraisch-abhängig, wenn man $y_s = 0$ setzt.

Beweis. Es sei

$$(1) \quad F(y_1, \dots, y_s, f_1, \dots, f_r) = 0$$

die Relation mit Koeffizienten aus P , die nach Voraussetzung zwischen f_1, \dots, f_r besteht, wo F nicht identisch verschwindet:

$$(2) \quad F(y_1, \dots, y_s, z_1, \dots, z_r) \neq 0 \quad \text{für unbestimmte } z_i.$$

Man kann annehmen, daß $F(y_1, \dots, y_s, z_1, \dots, z_r)$ nicht durch y_s teilbar ist, da man sonst den Faktor y_s weglassen könnte, ohne daß die Gültigkeit von (1) oder (2) gestört sein würde. In dieser Annahme ist

$$F(y_1, \dots, y_{s-1}, 0, z_1, \dots, z_r) \neq 0.$$

Bei der Substitution $y_s = 0$ mögen f_1, \dots, f_r in f_1^0, \dots, f_r^0 übergehen. Aus (1) folgt dann durch die Substitution $y_s = 0$:

$$F(y_1, \dots, y_{s-1}, 0, f_1^0, \dots, f_r^0) = 0, \quad \text{q. e. d.}$$

Bemerkung. Genau dasselbe gilt natürlich, wenn man, statt $y_s = 0$, $y_s = \alpha$ setzt, wo α irgendein Element von P ist, denn das heißt, daß man $y_s - \alpha$ als neue Unbestimmte auffaßt und durch Null ersetzt.

Wendet man den Hilfssatz mehreremal an, so folgt, daß man auch mehrere oder alle Unbestimmte y_i durch Konstante ersetzen kann, ohne daß die algebraische Abhängigkeit der f_i verloren geht. Diese Tatsache ist oft nützlich, um zu beweisen, daß gewisse Größen algebraisch-unabhängig sind, indem man durch eine passend gewählte Spezialisierung von vorkommenden Parametern die Annahme der algebraischen Abhängigkeit ad absurdum führt. In diesem Sinne wird der eben bewiesene Hilfssatz benutzt werden beim Beweis des folgenden Satzes.

Satz 30. Voraussetzung. Es seien M und M' irreduzible algebraische Mannigfaltigkeiten der Dimensionen r und $n - r$ in $P_n(P)$. Es sei U eine Matrix mit unbestimmten Koeffizienten u_{ik} ($i, k = 0, \dots, n$) in Ω .

Behauptung. 1. M'_Ω hat mit der transformierten $M_\Omega U$ mindestens einen Schnittpunkt.

2. Ist V irgendeine Matrix mit Elementen v_{ik} aus Ω , so daß $M_\Omega V$ und M'_Ω nur endlichviele Schnittpunkte haben, so ist die Multiplizität eines solchen Schnittpunkts Y niemals Null.

Beweis. Der Grundgedanke des Beweises von 1. ist folgender: Wir kehren zunächst das Problem um, indem wir einen Punkt G auf M'_Ω (möglichst allgemein, also als allgemeine Nullstelle von p') als gegeben annehmen, und eine möglichst allgemeine lineare Transformation Q so bestimmen, daß die transformierte Mannigfaltigkeit $M_\Omega Q$ auch den Punkt G enthält. Wenn es sich nun herausstellt, daß die so bestimmten Transformationskoeffizienten algebraisch-unabhängig sind, so folgt daraus, daß sie durch einen Körperisomorphismus in die Unbestimmte u_{ik} übergeführt werden können; der Punkt G geht dabei in einen Schnittpunkt X der Mannigfaltigkeiten $M_\Omega U$, M'_Ω über, womit die Existenz eines solchen dargetan ist.

Dieselbe Problemumkehrung führt nun aber auch zum Beweis der zweiten Behauptung. Es zeigt sich nämlich, daß man die allgemeine Nullstelle G und die weiteren beliebig angenommenen Größen relationstreu so spezialisieren kann, daß der Punkt G in jeden gegebenen Schnittpunkt Y von M'_Ω und $M_\Omega V$, und die Matrix Q in die gegebene Matrix V übergeht. Daraus folgt dann ohne weiteres, daß der gegebene Schnittpunkt Y auch eine relationstreu Spezialisierung von X ist, und daraus nach einem in W_2 bewiesenen Satz⁴³⁾, daß Y nicht die Multiplizität Null haben kann⁴⁴⁾.

Es seien also p und p' die zu M und M' gehörigen Primideale; es seien die Koordinaten so gewählt, daß in ihren allgemeinen Nullstellen die Koordinate ξ_0 nicht verschwindet. Dann kann man nach Satz 17 eine allgemeine Nullstelle von p in der Gestalt

$$F = \{\lambda, \lambda \varphi_1, \dots, \lambda \varphi_n\}$$

annehmen. Der Nullstellenkörper $\Phi = P(\lambda, \varphi_1, \dots, \varphi_n)$ kann als Unterkörper von Ω gedacht werden. Der Transzendenzgrad des Systems $\{\lambda, \varphi_1, \dots, \varphi_n\}$ ist $r + 1$; es seien etwa $\lambda, \varphi_1, \dots, \varphi_r$ algebraisch unabhängig in bezug auf P . Mit Rücksicht auf den zweiten Teil des Beweises ist es zweckmäßig, $\lambda, \varphi_1, \dots, \varphi_r$ in Ω so zu wählen, daß sie auch noch algebraisch unabhängig in bezug auf die Größen v_{ik} und η_i (Koordinaten von Y) sind. Die Nullstelle F bleibt dann eine allgemeine auch nach Adjunktion dieser Größen.

Eine allgemeine Nullstelle von p'_Φ kann ebenso in der Gestalt

$$G = \{\mu, \mu \psi_1, \dots, \mu \psi_n\}$$

(wiederum in Ω) angenommen werden. Das Elementsystem ψ_1, \dots, ψ_n hat in bezug auf Φ den Transzendenzgrad $n - r$. Es seien etwa $\psi_i, \dots, \psi_{i-n-r}$ algebraisch-unabhängig in bezug auf Φ .

⁴³⁾ W_2 Satz 5.

⁴⁴⁾ Ein ähnliches Beweisverfahren habe ich schon früher (W_2 § 8) benutzt, um zu zeigen, daß die Multiplizität einer Geraden einer kubischen Fläche (nicht Regelfläche) niemals Null ist.

Nun seien y_{ik} ($i = 0, \dots, n; k = 1, \dots, n$) Unbestimmte in Ω , und es sei

$$q_{ik} = \mu y_{ik} \quad (i = 0, \dots, n; k = 1, \dots, n),$$

$$q_{00} = \lambda - \sum_1^n y_{0k} \mu \psi_k,$$

$$q_{i0} = \lambda \varphi_i - \sum_1^n y_{ik} \mu \psi_k.$$

Dann wird die Transformation Q :

$$x_i = \sum_0^n q_{ik} x_k^*$$

den Punkt F von $P_n(\Omega)$ in $F^* = \{1, \psi_1, \dots, \psi_n\} = G$ transformieren. Die Mannigfaltigkeiten $M_\Omega Q$ und M'_Ω haben also den Punkt G gemein. Wir werden zeigen, daß die Transformation Q eine allgemeine ist, d. h. daß die q_{ik} algebraisch-unabhängig in bezug auf P sind.

Wären die q_{ik} algebraisch-abhängig, so würde eine Relation

$$H(q) = 0$$

mit Koeffizienten aus P bestehen. In dieser Relation könnten $q_{00}, q_{10}, \dots, q_{n0}$ sicher nicht fehlen, da die übrigen q_{ik} offensichtlich algebraisch-unabhängig sind. Also sind $q_{00}, q_{10}, \dots, q_{n0}$ algebraisch-abhängig in bezug auf $P(\mu, y_{ik})$. Diese Abhängigkeit muß nach Hilfssatz 6 bestehen bleiben, wenn man spezialisiert:

$$\begin{cases} \mu = -1, \\ y_{r+1, i_1} = y_{r+2, i_2} = \dots = y_{n, i_{n-r}} = 1, \\ \text{übrige } y_{ik} = 0. \end{cases}$$

Dabei gehen $q_{00}, q_{10}, \dots, q_{n0}$ über in

$$\lambda; \lambda \varphi_1; \dots; \lambda \varphi_r; \lambda \varphi_{r+1} + \psi_{i_1}; \dots; \lambda \varphi_n + \psi_{i_{n-r}}.$$

Zwischen diesen Größen müßte nun eine algebraische Relation mit Koeffizienten aus P bestehen. Da $\psi_{i_1}, \dots, \psi_{i_{n-r}}$ unabhängig in bezug auf $\Phi = P(\lambda, \varphi_1, \dots, \varphi_r)$ sind, so können $\lambda \varphi_{r+1} + \psi_{i_1}, \dots, \lambda \varphi_n + \psi_{i_{n-r}}$ zunächst in dieser Relation nicht vorkommen. Also bleibt eine Relation zwischen $\lambda, \lambda \varphi_1, \dots, \lambda \varphi_r$. Diese sind aber algebraisch-unabhängig in bezug auf P . Der Widerspruch ist da und die Unabhängigkeit der q_{ik} bewiesen.

Da die Unbestimmte u_{ik} ebenfalls unabhängig sind, so folgt die Isomorphie:

$$P(q_{ik}) \cong P(u_{ik}),$$

die sich zu einer Isomorphie des Erweiterungskörpers $P(q_{ik}, \lambda, \varphi_i)$ von $P(q_{ik})$ mit einem Erweiterungskörper von $P(u_{ik})$ innerhalb Ω erweitern

läßt. Dabei geht der Punkt G , Schnittpunkt von $M_\Omega Q$ und M'_Ω , über in einen Punkt X , Schnittpunkt von $M_\Omega U$ und M'_Ω , womit die Behauptung 1. bewiesen ist.

Um 2. zu beweisen, nehmen wir an, $Y = \{\eta_0, \dots, \eta_n\}$ sei ein Schnittpunkt von $M_\Omega V$ und M'_Ω , und es sei etwa $\eta_0 = 1$. Nach W_2 (Satz 5) wird 2. bewiesen sein, sobald man zeigen kann, daß Y eine relationstreue Spezialisierung des Schnittpunktes X von $M_\Omega U$ und M'_Ω ist. Es sei also

$$(2) \quad f \left(\begin{array}{c} \xi_0, \dots, \xi_n; \\ \dots \dots \dots \\ u_{00}, \dots, u_{0n} \\ \dots \dots \dots \\ u_{n0}, \dots, u_{nn} \end{array} \right) = 0$$

eine in ξ_0, \dots, ξ_n homogene, richtige Gleichung mit Koeffizienten aus P . Wir haben zu zeigen, daß daraus folgt

$$(3) \quad f \left(\begin{array}{c} \eta_0, \dots, \eta_n; \\ \dots \dots \dots \\ v_{00}, \dots, v_{0n} \\ \dots \dots \dots \\ v_{n0}, \dots, v_{nn} \end{array} \right) = 0.$$

Aus (2) folgt vermöge des Isomorphismus, der X in G und U in Q überführt:

$$(4) \quad f \left(\begin{array}{c} \mu, \mu \psi_1, \dots, \mu \psi_n; \\ \lambda \varphi_1 - \sum_1^n y_{0k} \mu \psi_k, \mu y_{01}, \dots, \mu y_{0n} \\ \dots \dots \dots \\ \lambda \varphi_n - \sum_1^n y_{nk} \mu \psi_k, \mu y_{n1}, \dots, \mu y_{nn} \end{array} \right) = 0.$$

Die allgemeine Nullstelle $\{\mu, \mu \psi_1, \dots, \mu \psi_k\}$ von p'_ϕ kann durch irgendeine spezielle Nullstelle von p' ersetzt werden, ohne daß die Gleichung (4) verloren geht; wir ersetzen sie durch $\{\eta_0, \dots, \eta_n\}$, dabei $\eta_0 = 1$ beachtend. Zugleich ersetzen wir die Unbestimmten y_{ik} durch v_{ik} .

$$f \left(\begin{array}{c} \eta_0, \dots, \eta_n; \\ \lambda - \sum_1^n v_{0k} \eta_k, v_{01}, \dots, v_{0n} \\ \lambda \varphi_1 - \sum_1^n v_{1k} \eta_k, v_{11}, \dots, v_{1n} \\ \dots \dots \dots \\ \lambda \varphi_n - \sum_1^n v_{nk} \eta_k, v_{n1}, \dots, v_{nn} \end{array} \right) = 0.$$

Hier kann man weiter die allgemeine Nullstelle $\{\lambda, \lambda \varphi_1, \dots, \lambda \varphi_n\}$ von p_P (v_{ik}, η_{kl}) ersetzen durch irgendeine spezielle Nullstelle von p_Ω . Nun ist $\{\eta_0, \dots, \eta_n\}$ Nullstelle von $p_\Omega V$, also

$$\left\{ \sum_0^n v_{0k} \eta_k, \dots, \sum_0^n v_{nk} \eta_k \right\}$$

Nullstelle von p_Ω . Man kann also $\{\lambda, \lambda \varphi_1, \dots, \lambda \varphi_n\}$ dadurch ersetzen. Das ergibt gerade die zu beweisende Gleichung (3).

Satz 30 gilt insbesondere, wenn man für V die in den §§ 7 und 8 betrachtete Matrix W einsetzt. Wie gezeigt, hat $M_\Omega W$ mit M'_Ω nur endlichviele Schnittpunkte und es folgt, daß die Multiplizitäten dieser Schnittpunkte > 0 sind.

§ 10.

Eine obere Grenze für die Multiplizitäten.

Der folgende Satz gibt für beliebige Normalprobleme eine obere Schranke für die Multiplizitäten der Lösungen.

Satz 31. Voraussetzung. *Ein Normalproblem, gegeben durch die Gleichungen*

$$\begin{cases} f_1(\lambda, \xi) = 0 \\ \dots \dots \dots \\ f_r(\lambda, \xi) = 0 \end{cases}$$

(wo $\lambda_1, \dots, \lambda_k$ Unbestimmte in Ω sind), habe im Körper Ω endlichviele Lösungen $X^{(1)}, \dots, X^{(q)}$. Nach der Spezialisierung $\lambda \rightarrow \mu$ habe das Normalproblem immer noch endlichviele Lösungen. Eine relationstreue Spezialisierung der $X^{(1)}, \dots, X^{(q)}$ für $\lambda \rightarrow \mu$ sei $Y^{(1)}, \dots, Y^{(q)}$. Von diesen Punkten seien $Y^{(1)}, \dots, Y^{(m)}$ gleich Y (m ist demnach die Multiplizität von Y). In $\Omega[x_1, \dots, x_n]$ betrachten wir die Ideale

$$\begin{aligned} \mathfrak{a} &= (f_1(\lambda, x), \dots, f_r(\lambda, x)), \\ \mathfrak{b} &= (f_1(\mu, x), \dots, f_r(\mu, x)), \end{aligned}$$

deren Nullstellen gerade die Lösungen unserer Gleichungen sind. Es sei χ die charakteristische Funktion des Ideals \mathfrak{b}_Y , und χ_i die des Ideals $\mathfrak{a}_{X^{(i)}}$.

Behauptung.

$$\chi \geq \sum_1^m \chi_i.$$

Eine Folge der Behauptung ist (weil notwendig jedes $\chi_i > 0$)

$$\chi \geq m,$$

womit eine obere Schranke für die Multiplizität m gegeben ist.

Beweis. Durch eine Koordinatentransformation sei erreicht, daß x_0 in den Punkten $X^{(i)}, Y^{(i)}$ Y nicht verschwindet.

Die Summe $\sum_1^m \chi_i$ ist die charakteristische Funktion des Ideals $[\alpha_{X^{(1)}}, \dots, \alpha_{X^{(m)}}]$ (Satz 18, § 4). Wir behaupten nun, daß, wenn σ eine hinreichend hohe Zahl ist, die beiden Ideale $[\alpha_{X^{(1)}}, \dots, \alpha_{X^{(m)}}]$ und \mathfrak{b}_Y dieselben charakteristischen Funktionen haben wie die beiden folgenden:

$$\mathfrak{v} = \left(\alpha, \prod_1^m m_i^\sigma \right),$$

$$\mathfrak{w} = (\mathfrak{b}, n^{m\sigma}).$$

wo

$$(1) \quad m_i = (\xi_0^{(i)} x_1 - \xi_1^{(i)} x_0, \xi_0^{(i)} x_2 - \xi_2^{(i)} x_0, \dots, \xi_0^{(i)} x_n - \xi_n^{(i)} x_0),$$

$$(2) \quad n = (\eta_0 x_1 - \eta_1 x_0, \eta_1 x_2 - \eta_2 x_0, \dots, \eta_0 x_n - \eta_n x_0)$$

gesetzt ist. Das wird nach Satz 21 bewiesen sein, sobald gezeigt ist, daß die zugehörigen inhomogenen Ideale übereinstimmen, also daß

$$[\bar{\alpha}_{X^{(1)}}, \dots, \bar{\alpha}_{X^{(m)}}] = (\bar{\alpha}, \prod_1^m \bar{m}_i^\sigma),$$

$$\bar{\mathfrak{b}}_Y = (\bar{\mathfrak{b}}, \bar{n}^{m\sigma}).$$

Diese Gleichungen sind aber nach den Sätzen 5 und 6 (§ 2) richtig. Also sind $\sum_1^m \chi_i$ und χ in der Tat die charakteristischen Funktionen der Ideale $\mathfrak{v}, \mathfrak{w}$.

Sei nun (r_1, \dots, r_r) eine Basis für $\prod_1^m m_i^\sigma$, gebildet durch Multiplikation aus den Basen (1) der einzelnen Ideale m_i . Dieselbe gibt nach der Spezialisierung $\xi_k^{(i)} \rightarrow \eta_k^{(i)} = \eta_k$ eine Basis (s_1, \dots, s_r) für $n^{m\sigma}$. Man hat also:

$$\mathfrak{v} = \left(\alpha, \prod_1^m m_i^\sigma \right) = (f_1(\lambda, x), \dots, f_r(\lambda, x), r_1, \dots, r_r),$$

$$\mathfrak{w} = (\mathfrak{b}, n^{m\sigma}) = (f_1(\mu, x), \dots, f_r(\mu, x), s_1, \dots, s_r).$$

Die Basis von \mathfrak{w} ist demnach eine relationstreue Spezialisierung der Basis von \mathfrak{v} . Zu zeigen ist, daß für die charakteristischen Funktionen der Ideale $\mathfrak{v}, \mathfrak{w}$ die Beziehung besteht:

$$\chi(\varrho; \mathfrak{w}) \geq \chi(\varrho; \mathfrak{v})$$

oder, was nach § 4, Gleichung (2), dasselbe ist,

$$(3) \quad \varphi(\varrho; \mathfrak{w}) \leq \varphi(\varrho; \mathfrak{v}).$$

Nun kann man eine Modulbasis für die Formen vom Grad ϱ von \mathfrak{v} dadurch gewinnen, daß man die Basiselemente von \mathfrak{v} multipliziert mit allen solchen Potenzprodukten der x_i , daß die Produkte jedesmal den Grad σ haben. Bildet man aus den Koeffizienten der so entstehenden Formen die Matrix (die „dialytische Matrix des Ideals \mathfrak{v} für den Grad ϱ “),

so gibt der Rang dieser Matrix den Wert von $\varphi(\varrho; \mathfrak{v})$ an. Ebenso kann man aber für das Ideal \mathfrak{w} eine Matrix bilden, dessen Rang den Wert von $\varphi(\varrho; \mathfrak{w})$ angibt. Der Rang der \mathfrak{w} -Matrix ist (wegen der relationstreuen Spezialisierung) \leq dem Rang der \mathfrak{v} -Matrix. Damit ist (3) bewiesen.

Dieser Satz werde nunmehr angewandt auf den Fall von § 8. Das Normalproblem bestand darin, die gemeinsamen Nullstellenklassen der Ideale $\mathfrak{p}_\Omega U$ und \mathfrak{p}'_Ω zu finden. Die Spezialisierung bestand darin, daß für U die singuläre Matrix W eingesetzt wird. Demnach hat man zu setzen:

$$\mathfrak{a} = (\mathfrak{p}_\Omega U, \mathfrak{p}'_\Omega),$$

$$\mathfrak{b} = (\mathfrak{p}_\Omega W, \mathfrak{p}'_\Omega),$$

m = die Multiplizität einer Lösung Y des spezialisierten Normalproblems,

χ = die charakteristische Funktion von \mathfrak{b}_Y .

Der Satz ergibt

$$(4) \quad \chi \geq m.$$

Um χ zu berechnen, beachte man, daß nach (5) § 2:

$$\mathfrak{b}_Y = (\mathfrak{p}_\Omega W, \mathfrak{p}'_\Omega)_Y = ((\mathfrak{p}_\Omega W)_Y, \mathfrak{p}'_\Omega)_Y.$$

Die Mannigfaltigkeit des Ideals $\mathfrak{p}_\Omega W$ zerfällt nach Satz 28 in lineare Räume $L^{(1)}, \dots, L^{(\beta)}$, die einen linearen $r-1$ Raum L gemein haben. Y ist ein Punkt eines $L^{(\beta)}$, der, wie in § 8 schon bemerkt, nicht in L liegt. Daraus folgt nach Satz 29 (§ 7):

$$(\mathfrak{p}_\Omega W)_Y = \mathfrak{a}^{(\beta)} W,$$

wo $\mathfrak{a}^{(\beta)} W$ das zu $L^{(\beta)}$ gehörige Primideal

$$\mathfrak{a}^{(\beta)} W = \left(\sum_k (\lambda_0^{(\beta)} u_{1k} - \lambda_1^{(\beta)} u_{0k}) x_k, \dots, \sum_k (\lambda_0^{(\beta)} u_{n-r,k} - \lambda_{n-r}^{(\beta)} u_{0k}) x_k \right)$$

ist. Setzt man

$$l_i = \sum_k (\lambda_0^{(\beta)} u_{ik} - \lambda_i^{(\beta)} u_{0k}) x_k \quad (i = 1, \dots, n-r),$$

so wird

$$\mathfrak{a}^{(\beta)} W = (l_1, \dots, l_{n-r}),$$

mithin

$$\mathfrak{b}_Y = (\mathfrak{p}'_\Omega, (\mathfrak{p}_\Omega W)_Y)_Y = (\mathfrak{p}'_\Omega, \mathfrak{a}^{(\beta)} W)_Y = (\mathfrak{p}'_\Omega, l_1, \dots, l_{n-r})_Y.$$

Die l_i sind allgemeine Linearformen. Also können wir auf das Ideal

$$(\overline{\mathfrak{p}'_\Omega}, \overline{l_1}, \dots, \overline{l_{n-r}})$$

den dritten Teil von Satz 27 anwenden, der besagt, daß dieses Ideal ein Durchschnitt ist von lauter verschiedenen Primidealen, die je zu einem der Schnittpunkte von $\overline{M'_\Omega}$ und $\overline{L^{(\beta)}}$ gehören. Da Y einer dieser Schnitt-

punkte ist, so ist die \bar{Y} -Komponente

$$\bar{b}_Y = (\bar{p}'_\Omega, \bar{l}_1, \dots, \bar{l}_{n-r})\bar{Y}$$

das zu \bar{Y} gehörige Primideal. Daraus folgt nach der Schlußbemerkung von § 4, daß die charakteristische Funktion χ von b_Y gleich 1 ist.

Demnach kann man für (4) schreiben $m \leq 1$, womit bewiesen ist:

Satz 32. *Die Schnittpunkte von $M_\Omega W$ und M'_Ω haben sämtlich eine Multiplizität ≤ 1 .*

§ 11.

Zusammenfassung der Ergebnisse. Folgerungen.

In § 9 ist gezeigt, daß die Multiplizität der Schnittpunkte von $M_\Omega W$ und M'_Ω nicht Null sind, in § 10, daß sie ≤ 1 sind, also sind sie $= 1$, womit nach § 8 der Hauptsatz (verallgemeinerte Bézoutsche Satz) *bewiesen ist*.

Aus dem Hauptsatz folgt nun auf Grund des Prinzips der Erhaltung der Anzahl, daß *die Summe der Multiplizitäten der Schnittpunkte von $M_\Omega V$ mit M'_Ω immer gleich dem Produkt der Grundzahlen ist, wie auch die Matrix V beschaffen sein mag, vorausgesetzt, daß die Anzahl der Schnittpunkte endlich ist*.

Insbesondere gilt das, wenn man für V die Einheitsmatrix E einsetzt, also die Schnittpunkte von M_Ω mit M'_Ω betrachtet. Man muß dabei wieder voraussetzen, daß M_Ω und M'_Ω , also insbesondere auch M und M' , nur endlichviele Schnittpunkte haben. Nun gilt:

Satz 33. *Haben M und M' nur endlichviele Schnittpunkte, so haben M_Ω und M'_Ω auch nur endlichviele Schnittpunkte, und zwar genau dieselben wie M und M' .*

Beweis. Hätten M_Ω und M'_Ω einen Schnittpunkt, dessen Koordinatenverhältnisse nicht in P lägen, also transzendent in bezug auf P wären, so könnte man diese Koordinatenverhältnisse als algebraische Funktionen von gewissen Parametern auffassen, und unendlichviele reguläre Argumentwerte⁴⁵⁾ für diese Funktionen finden, die zu unendlichvielen Schnittpunkten von M und M' führen würden, entgegen der Voraussetzung.

Nach diesem Satz braucht man nur über die Schnittpunkte von M und M' , nicht auch über die von M_Ω und M'_Ω zu reden; jeder Schnittpunkt von M und M' hat eine gewisse Multiplizität, und *die Summe dieser Multiplizitäten ist gleich dem Produkt der Gradzahlen*. Dies ist eine zweite Fassung des verallgemeinerten Bézoutschen Theorems.

⁴⁵⁾ W_1 § 2, 6.

Die Sätze der §§ 8, 9 ergeben aber noch mehr. Satz 23 besagt, daß die Multiplizität eines Schnittpunktes von $M_{\Omega}V$ und M'_{Ω} , insbesondere eines Schnittpunktes von M und M' , niemals Null ist. Daraus folgt:

Satz 34. *Die Anzahl der Schnittpunkte von M und M' ist höchstens gleich dem Produkt der Gradzahlen.*

Weiter besagt Satz 31, daß die Multiplizität eines Schnittpunktes Y von $M_{\Omega}V$ und M'_{Ω} höchstens gleich der charakteristischen Funktion des Ideals

$$(\mathfrak{p}_{\Omega}V, \mathfrak{p}'_{\Omega})_{\mathcal{F}}$$

ist. Insbesondere ist die Multiplizität m eines Schnittpunktes von M und M' höchstens gleich der charakteristischen Funktion χ des Ideals

$$(\mathfrak{p}_{\Omega}, \mathfrak{p}'_{\Omega})_{\mathcal{F}},$$

oder, was nach Satz 23 (§ 5) auf dasselbe herauskommt, des Ideals

$$(\mathfrak{p}, \mathfrak{p}')_{\mathcal{F}}.$$

Wir werden in § 12 sehen, daß im Fall $r = 1$, also wenn M ein-dimensional ist, die Ungleichung

$$\chi \leq m$$

zu einer Gleichung

$$\chi = m$$

verschärft werden kann. Daß dies aber im allgemeinen nicht der Fall zu sein braucht, sieht man aus dem folgenden Beispiel, das im wesentlichen von Macaulay herrührt⁴⁶⁾.

Das H -Ideal

$$\mathfrak{p}' = (x_1 x_4 - x_2 x_3, x_2^3 - x_1^2 x_3, x_3^3 - x_2 x_4^2, x_2^2 x_4 - x_1 x_3^2)$$

in $\mathcal{P}[x_0, \dots, x_4]$ ist prim, weil es die allgemeine Nullstelle

$$\{\lambda, \mu^4, \mu^3 \nu, \mu \nu^3, \nu^4\}$$

hat. Die Punkte seiner Mannigfaltigkeit M' werden gefunden, indem man den Unbestimmten λ, μ, ν spezielle Werte λ', μ', ν' gibt.

Mit der linearen Mannigfaltigkeit M , gegeben durch das Primideal

$$\mathfrak{p} = (x_1, x_4)$$

hat M' offensichtlich nur einen Schnittpunkt, nämlich

$$Y = \{\lambda, 0, 0, 0, 0\}.$$

Mit der allgemein-transformierten Mannigfaltigkeit $M_{\Omega}U$, gegeben durch

$$\mathfrak{p}_{\Omega}U = (\sum u_{1k} x_k, \sum u_{4k} x_k),$$

⁴⁶⁾ Macaulay, S. 98.

hat M'_0 aber vier Schnittpunkte, wie man durch Auflösung der Gleichungen

$$\begin{cases} u_{10}\lambda' + u_{11}\mu'^4 + u_{12}\mu'^3\nu' + u_{13}\mu'\nu'^2 + u_{14}\nu'^4 = 0, \\ u_{40}\lambda' + u_{41}\mu'^4 + u_{42}\mu'^3\nu' + u_{43}\mu'\nu'^2 + u_{44}\nu'^4 = 0 \end{cases}$$

sieht. Also hat der eine Schnittpunkt Y von M und M' die Multiplizität 4:

$$m = 4.$$

χ ist die charakteristische Funktion des Ideals

$$(\mathfrak{p}, \mathfrak{p}')_Y$$

in $P[x_1, \dots, x_4]$. Nun ist

$$(\mathfrak{p}, \mathfrak{p}') = (x_1, x_4, x_2x_3, x_2^3, x_3^3)$$

ein Primärideal mit der Nullstelle Y , also folgt:

$$(\mathfrak{p}, \mathfrak{p}')_Y = (x_1, x_4, x_2x_3, x_2^3, x_3^3).$$

Geht man auf Grund von Satz 20 (§ 4) auf inhomogene Ideale über durch die Substitution $x_0 = 1$, so wird ein Maximalsystem von modulo dem Ideal $(x_1, x_4, x_2x_3, x_2^3, x_3^3)$ linear-unabhängigen Polynomen gegeben durch

$$1, x_2, x_3, x_2^2, x_3^2.$$

Also ist $\chi = 5$. Damit ist die Gleichung $\chi = m$ widerlegt.

§ 12.

Der Fall $r = 1$.

Satz 35. *Es seien M und M' irreduzible Mannigfaltigkeiten der Dimensionen 1 und $n - 1$ im $P_n(P)$, die sich in endlichvielen Punkten schneiden. Es seien \mathfrak{p} und \mathfrak{p}' die zugehörigen Primideale, Y ein Schnittpunkt von M und M' , m seine Multiplizität, und χ die charakteristische Funktion des Ideals $(\mathfrak{p}, \mathfrak{p}')_Y$. Dann ist $\chi = m$.*

Beweis. In § 11 wurde schon bemerkt, daß $\chi \leq m$. Wenn wir also zeigen können, daß bei Summation über alle Schnittpunkte

$$\sum \chi \geq \sum m$$

so wird der Satz bewiesen sein.

$\sum m$ ist nach § 11 gleich dem Produkt der Gradzahlen von M und M' :

$$(1) \quad \sum m = s \cdot s'.$$

$\sum \chi$ ist nach § 4, Gleichung (4), gleich der charakteristischen Funktion des Ideals $(\mathfrak{p}, \mathfrak{p}')$. Nach W_1 § 5, 10 ist \mathfrak{p}' ein Hauptideal, das von einer irreduziblen Form f vom Grad γ erzeugt wird:

$$\begin{aligned} \mathfrak{p}' &= (f), \\ (\mathfrak{p}, \mathfrak{p}') &= (\mathfrak{p}, f), \end{aligned}$$

wo f zu p prim ist. Daraus folgt nach § 4, Gleichung (10):

$$\text{Grad } (p, f) = \gamma \cdot \text{Grad } p,$$

oder

$$(2) \quad \sum \chi = \gamma \cdot \text{Grad } p.$$

Dieselben Gleichungen gelten auch dann, wenn die Koeffizienten der auftretenden Polynome einem Erweiterungskörper Ω von P entnommen werden, denn bei dieser Erweiterung bleiben alle charakteristischen Funktionen ungeändert (Satz 23). Wählt man für f eine allgemeine Linearform, so wird die linke Seite von (2) mindestens gleich dem Grad von M , da schon die Anzahl der Summanden gleich diesem Grad ist, und die einzelnen Summanden positiv. Auf der rechten Seite wird $\gamma = 1$. Also kommt:

$$\text{Grad } p \geq \text{Grad } M = s.$$

Die Formel (2) gilt aber auch dann, wenn M ein allgemeiner linearer 1 Raum ist, und M' beliebig. Dann ist aber die linke Seite mindestens gleich dem Grad von M' , da schon die Anzahl der Summanden gleich diesem Grad ist, und die einzelnen Summanden positiv. Auf der rechten Seite ist $\text{Grad } p = 1$. Also kommt

$$\gamma \geq \text{Grad von } M' = s',$$

und (2) ergibt

$$(3) \quad \sum \chi \geq s' \cdot s.$$

Aus (1) und (3) folgt die gesuchte Ungleichung

$$\sum \chi \geq \sum m.$$

Hinterher schließt man, daß in allen Formeln dieses Paragraphen das Gleichheitszeichen gelten muß. Es ist also $s' = \gamma$ und $s = \text{Grad } p$.

(Eingegangen am 3. 6. 1927.)

Über gewisse Extremumprobleme der Funktionentheorie.

Von

E. Egerváry in Budapest.

Einleitung.

1. In vorliegender Arbeit wird die explizite Lösung eines Carathéodory-Fejérschen Extremumproblems in folgendem speziellen Falle gegeben:

Unter allen für $|z| < 1$ regulären Potenzreihen, deren n ersten Koeffizienten gleich 1 sind, d. h. die Form haben:

$$1 + z + z^2 + \dots + z^{n-1} + \gamma_n z^n + \dots$$

ist diejenige zu bestimmen, für welche das Maximum des absoluten Betrages für $|z| < 1$ (das Modulmaximum) den kleinstmöglichen Wert annimmt.

Es ergibt sich, daß die untere Grenze des Modulmaximums gleich

$$\frac{1}{2 \cos \frac{n}{2n+1} \pi} = \frac{1}{2 \sin \frac{\pi}{4n+2}}$$

ist und für die rationale Funktion

$$h^*(z) = \frac{1}{2 \cos \frac{n}{2n+1} \pi} \cdot \frac{\sin \frac{\pi}{2n+1} + \sin \frac{2\pi}{2n+1} z + \dots + \sin \frac{n\pi}{2n+1} z^{n-1}}{\sin \frac{n\pi}{2n+1} + \sin \frac{(n-1)\pi}{2n+1} z + \dots + \sin \frac{\pi}{2n+1} z^{n-1}}$$

erreicht wird.

Abgesehen davon, daß hiermit zum ersten Male die explizite Lösung eines vorgelegten Carathéodory-Fejérschen Extremumproblems für Potenzreihen¹⁾ geboten wird, kommt diesem speziellen Falle eine erhöhte Be-

¹⁾ Für ein anderes verwandtes Carathéodory-Fejérsches Problem, welches sich auf die *harmonische* Entwicklung $a_0 + \sum_{\nu=1}^{\infty} \varrho^{\nu} (a_{\nu} \cos \nu \theta + b_{\nu} \sin \nu \theta)$ bezieht, hat Herr O. Hölder einige spezielle Fälle ausgearbeitet. Vgl. O. Hölder, Über einige Determinanten, Leipz. Ber. 65 (1913), S. 110–120.

deutung durch den Umstand zu, daß durch seine Lösung auch die Frage nach der oberen Grenze für den Absolutwert der Koeffizientensumme für Potenzreihen mit beschränktem Mittelmodul²⁾, also die Frage, welche für Potenzreihen mit beschränktem Modulmaximum²⁾ von Herrn E. Landau aufgeworfen und gelöst wurde³⁾, vollständig erledigt wird und zwar durch folgenden Satz:

Für $|z| \leq 1$ sei $f(z) = a_0 + a_1 z + a_2 z^2 + \dots$ regulär und es gelte

$$\frac{1}{2\pi} \int_0^{2\pi} |f(e^{it})| dt \leq 1.$$

Dann ist für die Gesamtheit aller dieser $f(z)$

$$(1) \quad |s_n(1)| = |a_0 + a_1 + \dots + a_{n-1}| \leq \frac{1}{2 \cos \frac{n}{2n+1} \pi} \quad 4)$$

und die obere Grenze wird nur bei der rationalen ganzen Funktion

$$f^*(z) = \frac{4}{2n+1} \left[\sin \frac{n\pi}{2n+1} + \sin \frac{(n-1)\pi}{2n+1} z + \dots + \sin \frac{\pi}{2n+1} z^{n-1} \right]^2$$

erreicht.

2. Die Betrachtung des obengenannten Extremumproblems in dem etwas allgemeineren Falle, wo die vorgeschriebenen n ersten Koeffizienten

²⁾ Eine für $|z| \leq 1$ konvergente Potenzreihe

$$f(z) = a_0 + a_1 z + a_2 z^2 + \dots$$

werde ich der Kürze halber im folgenden überall „Potenzreihe mit beschränktem Mittelmodul“ nennen, falls die Ungleichung gilt

$$\frac{1}{2\pi} \int_0^{2\pi} |f(e^{it})| dt \leq 1.$$

Dieselbe Potenzreihe nenne ich „Potenzreihe mit beschränktem Modulmaximum“, falls die Ungleichung gilt

$$|f(z)| \leq 1 \quad \text{für } |z| \leq 1.$$

Im folgenden wird überall angenommen, daß $f(z)$ und die anderen betrachteten Funktionen auch für $|z| = 1$ regulär sind, was keine Beschränkung der Allgemeinheit bedeutet. Ist nämlich $f(z)$ nur für $|z| < 1$ regulär, so wendet man die bewiesenen Sätze auf $f(\varrho z)$ an ($0 < \varrho < 1$) und erhält durch Grenzübergang für $\varrho \rightarrow 1$ dieselben Ungleichungen für $f(z)$, welche im spezielleren Falle $f(z)$ regulär für $|z| \leq 1$ bewiesen wurden. Auch die Unizitätssätze lassen sich mit Heranziehung eines Fatouschen Satzes auf den allgemeineren Fall $f(z)$ regulär für $|z| < 1$ übertragen.

³⁾ E. Landau, Abschätzung der Koeffizientensumme einer Potenzreihe, Arch. Math. Phys. (3) 21 (1913), S. 42–50 und S. 250–255.

⁴⁾ Sogar $|a_0| + |a_1| + \dots + |a_{n-1}|$ hat dieselbe obere Grenze, vgl. § 3.

$\varrho^{n-1}, \varrho^{n-2}, \varrho^{n-3}, \dots, \varrho, 1$ ($0 < \varrho \leq 1$) sind, ergab eine Abschätzung für die Majorante $\mathfrak{M}(\varrho) = \sum_{n=0}^{\infty} |a_n| \varrho^n$ von Potenzreihen mit beschränktem Mittelmodul⁵⁾, die ich in folgendem Satze ausspreche:

Für $|z| \leq 1$ sei $f(z) = a_0 + a_1 z + a_2 z^2 + \dots$ regulär und es gelte

$$\frac{1}{2\pi} \int_0^{2\pi} |f(e^{it})| dt \leq 1.$$

Dann ist für die Gesamtheit aller dieser $f(z)$

$$(2) \quad \mathfrak{M}(\varrho) = |a_0| + |a_1| \varrho + |a_2| \varrho^2 + \dots \leq \frac{1}{1-\varrho^2} \quad (0 < \varrho < 1)^6)$$

und es gibt für jedes ϱ ($0 < \varrho < 1$) eine Folge von rationalen ganzen Funktionen $f_n^*(z)$ (mit positiven Koeffizienten) von der Beschaffenheit

$$\frac{1}{2\pi} \int_0^{2\pi} |f_n^*(e^{it})| dt = 1; \quad f_n^*(\varrho) \rightarrow \frac{1}{1-\varrho^2} \quad \text{für } n \rightarrow \infty.$$

3. Die Lösung der obengenannten Extremumprobleme gründet sich hauptsächlich auf folgende bekannte Tatsachen⁷⁾.

Es seien c_0, c_1, \dots, c_{n-1} positive, monoton zunehmende Zahlen

$$0 < c_0 \leq c_1 \leq \dots \leq c_{n-1};$$

dann ist für die Gesamtheit aller in $|z| \leq 1$ konvergenten Potenzreihen

$$(3) \quad f(z) = c_0 + c_1 z + \dots + c_{n-1} z^{n-1} + \dots,$$

⁵⁾ Die Majorante von Potenzreihen mit beschränktem Modulmaximum ist von G. H. Hardy untersucht worden, A theorem concerning Taylor series, Quart. J. 44 (1913), S. 147–160.

⁶⁾ Aus $\frac{1}{2\pi} \int_0^{2\pi} |f(e^{it})| dt \leq 1$ folgt nach Cauchy unmittelbar $|a_n| \leq 1$ ($n=0, 1, 2, \dots$),

also

$$|a_0| + |a_1| \varrho + |a_2| \varrho^2 + \dots \leq \frac{1}{1-\varrho}.$$

Das Gewicht im Satze (2) liegt also einerseits in der Verkleinerung durch den Faktor $\frac{1}{1+\varrho}$, andererseits auf der Tatsache, daß Ungleichung (2) nicht weiter verbessert werden kann.

⁷⁾ Vgl. L. Fejér, Über gewisse Minimumprobleme der Funktionentheorie, Math. Annalen 97 (1926), S. 104–123. Hier ist die einschlägige Literatur ausführlich besprochen. Insbesondere enthält die Fejérsche Arbeit für den im Texte angeführten Fall $0 < c_0 \leq c_1 \leq \dots \leq c_{n-1}$ einen elementaren, von der allgemeinen Theorie unabhängigen Beweis.

deren n ersten Koeffizienten die festen positiven Zahlen c_0, c_1, \dots, c_{n-1} sind, die untere Grenze des Modulmaximums für $|z| \leq 1$ gleich dem (jedenfalls positiven) absolut größten Werte, den die reelle quadratische Form

$$(4) \quad Q_n(x_0, x_1, \dots, x_{n-1}) = \sum_{\substack{p, q=0, 1, \dots, n-1 \\ 0 \leq p+q \leq n-1}} c_{n-(p+q)-1} x_p x_q \\ = \sum_{k=0}^{n-1} c_k (x_0 x_{n-k-1} + x_1 x_{n-k-2} + \dots + x_{n-k-1} x_0)$$

bei der Nebenbedingung

$$(4') \quad x_0^2 + x_1^2 + \dots + x_{n-1}^2 = 1$$

annimmt, also gleich der absolut größten Wurzel λ_n^* der charakteristischen Gleichung

$$(5) \quad D_n(\lambda) = \begin{vmatrix} c_{n-1} - \lambda & c_{n-2} & c_{n-3} & \dots & c_1 & c_0 \\ c_{n-2} & c_{n-3} - \lambda & & & c_1 & c_0 & 0 \\ c_{n-3} & & & & & & \vdots \\ \vdots & c_1 & & & & & \\ c_1 & c_0 & & & -\lambda & 0 \\ c_0 & 0 & \dots & 0 & -\lambda \end{vmatrix} = 0.$$

Für jede Funktion $f(z)$ der Gesamtheit (3) ist demnach

$$(6) \quad \text{Max}_{|z| \leq 1} |f(z)| \geq \text{Max}_{\sum_{v=0}^{n-1} x_v^2 = 1} |Q_n(x_0, x_1, \dots, x_{n-1})| = \lambda_n^*;$$

ist $(x_0^*, x_1^*, \dots, x_{n-1}^*)$ das einzige positive Wertsystem⁸⁾, für welches

⁸⁾ Bezüglich der Existenz eines positiven Wertsystems, für welches eine reelle quadratische Form mit nichtnegativen Koeffizienten ihr absolutes Maximum annimmt, vgl. Fejér, loc. cit. 7), S. 111.

Daß es nur ein Wertsystem gibt, für welches $Q_n(x_0, x_1, \dots, x_{n-1}) = \lambda_n^*$ wird, daß also die in Betracht kommenden Extremumprobleme nur eine Lösung besitzen, läßt sich — ohne Heranziehung der allgemeinen Theorie — folgendermaßen beweisen. Gibt es zwei linear unabhängige Wertsysteme, für welche Q_n ihr absolutes Maximum λ_n^* erreicht, so müssen auch alle Unterdeterminanten $n-1$ -ter Ordnung von

$$\begin{vmatrix} c_{n-1} - \lambda_n^* & c_{n-2} & c_{n-3} & \dots & c_1 & c_0 \\ c_{n-2} & c_{n-3} - \lambda_n^* & & & & \\ c_{n-3} & & & & & \vdots \\ \vdots & & & & & \\ c_1 & c_0 & & & -\lambda_n^* & 0 \\ c_0 & 0 & \dots & 0 & -\lambda_n^* \end{vmatrix}$$

(Fortsetzung der Fußnote ⁸⁾ auf der nächsten Seite.)

$Q_n(x_0^*, x_1^*, \dots, x_{n-1}^*) = \lambda_n^*$ wird, so gilt für die alsdann in $|z| \leq 1$ reguläre rationale Funktion und nur für diese

$$(7) \quad h^*(z) = \lambda_n^* \frac{x_{n-1}^* + x_{n-2}^* z + \dots + x_0^* z^{n-1}}{x_0^* + x_1^* z + \dots + x_{n-1}^* z^{n-1}}$$

das Gleichheitszeichen in (6).

§ 1 enthält die Berechnung und Diskussion der charakteristischen Gleichung für die Koeffizientenfolge $c_0 = c_1 = \dots = c_{n-1} = 1$, sowie für die allgemeinere Folge $c_0 = \varrho^{n-1}$, $c_1 = \varrho^{n-2}$, \dots , $c_{n-1} = 1$.

In § 2 berechne ich die explizite Lösung des Carathéodory-Fejérschen Extremumproblems für Potenzreihen mit den Anfangskoeffizienten $c_0 = c_1 = \dots = c_{n-1} = 1$, sowie mit

$$c_0 = \left(\frac{n}{n+1}\right)^{n-1}, \quad c_1 = \left(\frac{n}{n+1}\right)^{n-2}, \quad \dots, \quad c_{n-1} = 1.$$

§ 3 enthält die Bestimmung der oberen Grenze für den Absolutwert der Koeffizientensumme (sowie für die Summe der Absolutwerte der Koeffizienten) aller in $|z| \leq 1$ regulären Potenzreihen mit beschränktem Mittelmodul, wobei auch der Fall einer beliebigen gegebenen linearen Verbindung der Koeffizienten besprochen wird.

Im § 4 wird endlich der eingangs erwähnte Zusammenhang zwischen der Majorante $\mathfrak{M}(\varrho)$ und dem Mittelmodul $\frac{1}{2\pi} \int_0^{2\pi} |f(e^{it})| dt$ einer für $|z| \leq 1$ regulären Potenzreihe $f(z)$ hergeleitet.

§ 1.

Berechnung und Diskussion der charakteristischen Gleichung.

4. Ich betrachte zuerst die charakteristische Gleichung für die Koeffizientenfolge $c_0 = c_1 = \dots = c_{n-1} = 1$. In diesem Falle ist

verschwinden, es gibt also ein positives Wertsystem $(\xi_0^*, \xi_1^*, \dots, \xi_{n-3}^*)$ mit der Quadratsumme 1, für welches

$$Q_{n-2}(x_0, x_1, \dots, x_{n-3}) = \sum_{k=0}^{n-3} c_k (x_0 x_{n-k-3} + x_1 x_{n-k-4} + \dots + x_{n-k-3} x_0) = \lambda_n^*$$

wird. Dann nimmt aber $Q_n(x_0, \dots, x_{n-1})$ für das Wertsystem $x_0 = \xi_0^*, \dots, x_{n-3} = \xi_{n-3}^*$, $x_{n-2} = x_{n-1} = 0$ wegen $c_{n-1} \geq c_{n-2} \geq \dots \geq c_0 > 0$ einen größeren Wert an, entgegen der Annahme, daß λ_n^* ihr absolutes Maximum ist.

$$(8) \quad D_n(\lambda) = \begin{vmatrix} 1 - \lambda & 1 & 1 & \dots & 1 & 1 \\ 1 & 1 - \lambda & & & 1 & 0 \\ 1 & & & & & \vdots \\ \vdots & & & & & \vdots \\ 1 & 1 & & & -\lambda & 0 \\ 1 & 0 & \dots & 0 & -\lambda & \end{vmatrix}.$$

Um die Wurzeln dieser Gleichung n -ten Grades zu bestimmen, bilde ich $D_n(\lambda)D_n(-\lambda)$. Durch Anwendung der Multiplikationsregel für zwei Determinanten n -ter Ordnung ergibt sich

$$(9) \quad D_n(\lambda)D_n(-\lambda) = \begin{vmatrix} n - \lambda^2 & n - 1 & n - 2 & \dots & 2 & 1 \\ n - 1 & n - 1 - \lambda^2 & n - 2 & & 2 & 1 \\ n - 2 & n - 2 & n - 2 - \lambda^2 & & 2 & 1 \\ \vdots & & & & \vdots & \vdots \\ 2 & 2 & 2 & & 2 - \lambda^2 & 1 \\ 1 & 1 & 1 & \dots & 1 & 1 - \lambda^2 \end{vmatrix}.$$

Wird der Reihe nach die zweite Kolonne von der ersten, die dritte von der zweiten, ..., die n -te von der $n - 1$ -ten subtrahiert, während die n -te ungeändert bleibt, und in der so veränderten Determinante der Reihe nach die zweite Zeile von der ersten, die dritte von der zweiten, ..., die n -te von der $n - 1$ -ten subtrahiert, während die n -te ungeändert bleibt, so geht die Determinante (9) in die folgende über:

$$(10) \quad D_n(\lambda)D_n(-\lambda) = \begin{vmatrix} 1 - 2\lambda^2 & \lambda^2 & 0 & \dots & 0 \\ \lambda^2 & 1 - 2\lambda^2 & \lambda^2 & & 0 \\ 0 & \lambda^2 & 1 - 2\lambda^2 & & 0 \\ \dots & \dots & \dots & \dots & \dots \\ & & & 1 - 2\lambda^2 & \lambda^2 \\ 0 & 0 & 0 & \lambda^2 & 1 - \lambda^2 \end{vmatrix}.$$

Mit Benutzung der bekannten Relation⁹⁾

$$1) \quad \begin{vmatrix} \mu & 1 & 0 & \dots & 0 & 0 \\ 1 & \mu & 1 & & & \\ 0 & 1 & \mu & & & \\ \vdots & & & & & \vdots \\ & & & \mu & 1 & \\ 0 & 0 & \dots & 1 & \mu & \end{vmatrix} = \frac{\sin(n+1)\theta}{\sin\theta}, \quad \mu = 2\cos\theta$$

⁹⁾ Vgl. z. B. E. Pascal, Determinanten (Teubners Sammlung 3, 1900).

erhält man also

$$(12) \quad D_n(\lambda) D_n(-\lambda) = \lambda^{2n} \begin{vmatrix} \frac{1}{\lambda^2} - 2 & 1 & 0 & & 0 \\ 1 & \frac{1}{\lambda^2} - 2 & 1 & & 0 \\ 0 & 1 & \frac{1}{\lambda^2} - 2 & & \\ \vdots & & & & \vdots \\ & & & \frac{1}{\lambda^2} - 2 & 1 + 0 \\ 0 & 0 & & 1 & \frac{1}{\lambda^2} - 2 + 1 \end{vmatrix}$$

$$= \lambda^{2n} \left[\frac{\sin(n+1)\theta}{\sin\theta} + \frac{\sin n\theta}{\sin\theta} \right] = \lambda^{2n} \frac{\sin(2n+1)\frac{\theta}{2}}{\sin\frac{\theta}{2}}, \quad 2\cos\theta = \frac{1}{\lambda^2} - 2.$$

$D_n(\lambda) D_n(-\lambda)$ verschwindet demnach für

$$\theta = \pm \frac{2\pi}{2n+1}, \pm \frac{4\pi}{2n+1}, \dots, \pm \frac{2n\pi}{2n+1},$$

oder, da nach (12) $\lambda = \pm \frac{1}{2\cos\frac{\theta}{2}}$ ist, für

$$\lambda = \pm \frac{1}{2\cos\frac{\pi}{2n+1}}, \pm \frac{1}{2\cos\frac{2\pi}{2n+1}}, \dots, \pm \frac{1}{2\cos\frac{n\pi}{2n+1}}.$$

Die absolut größte Wurzel λ_n^* von $D_n(\lambda)$ ist also

$$(13) \quad \lambda_n^* = \frac{1}{2\cos\frac{n\pi}{2n+1}} = \frac{1}{2\sin\frac{\pi}{4n+2}} \sim \frac{2}{\pi} n.$$

5. Mit demselben Verfahren läßt sich auch die charakteristische Gleichung für die Koeffizientenfolge $c_0 = \varrho^{n-1}, c_1 = \varrho^{n-2}, \dots, c_{n-1} = 1$ ($0 < \varrho < 1$)

$$(14) \quad D_n(\lambda; \varrho) = \begin{vmatrix} 1 - \lambda & \varrho & \varrho^2 & \dots & \varrho^{n-2} & \varrho^{n-1} \\ \varrho & \varrho^2 - \lambda & & & \varrho^{n-1} & 0 \\ \varrho^2 & & & & & \\ \vdots & & & & & \vdots \\ \varrho^{n-2} & \varrho^{n-1} & & & -\lambda & 0 \\ \varrho^{n-1} & 0 & & & 0 & -\lambda \end{vmatrix} = 0$$

in eine geschlossene Form bringen. Es ergibt sich zunächst für $D_n(\lambda; \varrho) D_n(-\lambda; \varrho)$

$$\begin{aligned}
 (15) \quad & D_n(\lambda; \varrho) D_n(-\lambda; \varrho) = \\
 & \begin{vmatrix}
 \frac{1-\varrho^{2n}}{1-\varrho^2} - \lambda^2 & \varrho \frac{1-\varrho^{2(n-1)}}{1-\varrho^2} & \varrho^2 \frac{1-\varrho^{2(n-2)}}{1-\varrho^2} & \dots & \varrho^{n-2} \frac{1-\varrho^4}{1-\varrho^2} & \varrho^{n-1} \frac{1-\varrho}{1-\varrho^2} \\
 \varrho \frac{1-\varrho^{2(n-1)}}{1-\varrho^2} & \varrho^2 \frac{1-\varrho^{2(n-1)}}{1-\varrho^2} - \lambda^2 & & & \varrho^{n-1} \frac{1-\varrho^4}{1-\varrho^2} & \varrho^n \frac{1-\varrho^2}{1-\varrho^2} \\
 \dots & \dots & \dots & \dots & \dots & \dots \\
 \varrho^{n-2} \frac{1-\varrho^4}{1-\varrho^2} & \varrho^{n-1} \frac{1-\varrho^4}{1-\varrho^2} & & & & \\
 \varrho^{n-1} \frac{1-\varrho^2}{1-\varrho^2} & \varrho^n \frac{1-\varrho^2}{1-\varrho^2} & & & & \varrho^{2n-2} \frac{1-\varrho^2}{1-\varrho^2}
 \end{vmatrix} \\
 & = \varrho^2 \frac{1-\varrho^{2(n-2)}}{1-\varrho^2}
 \end{aligned}$$

Diese Determinante wird jetzt in folgenden Schritten umgestaltet:
 1. Aus der k -ten Kolonne wird der Faktor ϱ^k ($k=1, 2, \dots, n$), aus der h -ten Zeile aber der Faktor ϱ^h ($h=1, 2, \dots, n$) herausgehoben. 2. Die zweite Kolonne wird von der ersten, die dritte von der zweiten, ..., die n -te von der $n-1$ -ten subtrahiert, und die letzte bleibt ungeändert. 3. In der so entstandenen Determinante wird der Reihe nach die zweite Zeile von der ersten, die dritte von der zweiten, ..., die n -te von der $n-1$ -ten subtrahiert, während die n -te ungeändert bleibt. 4. Die k -te Zeile ($k=1, 2, \dots, n$) wird mit $\frac{\varrho^{2k-1}}{\lambda^2}$ multipliziert. Die Determinante (15) verwandelt sich dann in folgende

$$\begin{aligned}
 (16) \quad & D_n(\lambda; \varrho) D_n(-\lambda; \varrho) = \\
 & \frac{\lambda^{2n}}{\varrho^n} \begin{vmatrix}
 \frac{\varrho^{2n-1}}{\lambda^2} - \left(\varrho + \frac{1}{\varrho}\right) & \frac{1}{\varrho} & 0 & \dots & 0 & 0 \\
 \varrho & \frac{\varrho^{2n-1}}{\lambda^2} - \left(\varrho + \frac{1}{\varrho}\right) & \frac{1}{\varrho} & \dots & 0 & 0 \\
 0 & \varrho & & & & \\
 \vdots & & & & & \vdots \\
 0 & 0 & \dots & \frac{\varrho^{2n-1}}{\lambda^2} - \left(\varrho + \frac{1}{\varrho}\right) & \frac{1}{\varrho} & \\
 0 & 0 & \dots & \varrho & \frac{\varrho^{2n-1}}{\lambda^2} - \varrho &
 \end{vmatrix},
 \end{aligned}$$

also wieder mit Rücksicht auf (11)

$$\begin{aligned}
 (17) \quad & D_n(\lambda; \varrho) D_n(-\lambda; \varrho) = \frac{\lambda^{2n}}{\varrho^{n+1}} \cdot \frac{\varrho \sin(n+1)\theta + \sin n\theta}{\sin \theta}, \\
 & 2 \cos \theta = \frac{\varrho^{2n-1}}{\lambda^2} - \left(\varrho + \frac{1}{\varrho}\right).
 \end{aligned}$$

Hieraus erkennt man unschwer, daß $D_n(\lambda; \varrho) D_n(-\lambda; \varrho)$ für $2n-2$ (mod. 2π verschiedene) reelle Werte von θ

$$\theta = \pm \frac{\pi}{n + \frac{1}{2} \vartheta_1}, \pm \frac{2\pi}{n + \frac{1}{2} \vartheta_2}, \dots, \pm \frac{(n-1)\pi}{n + \frac{1}{2} \vartheta_{n-1}} \quad (0 \leq \vartheta_1, \vartheta_2, \dots, \vartheta_{n-1} \leq 1),$$

also, da nach (17) $\lambda^2 = \frac{\varrho^{2n}}{1 + 2\varrho \cos \theta + \varrho^2}$ ist, für

$$\lambda = \pm \frac{\varrho^n}{\sqrt{1 - 2\varrho \cos \frac{k\pi}{n + \frac{1}{2} \vartheta_k} + \varrho^2}} \quad (k = 1, 2, \dots, n-1)$$

verschwindet¹⁰⁾. Somit gilt für $n-1$ Wurzeln von $D_n(\lambda; \varrho)$ die Abschätzung

$$(18) \quad |\lambda| \leq \frac{\varrho^n}{1-\varrho} \quad (0 \leq \varrho < 1).$$

Für die absolut größte Wurzel von $D_n(\lambda; \varrho)$, die ich im folgenden mit $\lambda_n^*(\varrho)$ bezeichnen werde, gilt hingegen

$$(19') \quad |\lambda_n^*(\varrho)| = \frac{\varrho^n}{\sqrt{1 + 2\varrho \cos \frac{n\pi}{n + \frac{1}{2} \vartheta} + \varrho^2}} \leq \frac{\varrho^n}{1-\varrho} \quad (0 \leq \vartheta \leq 1) \text{ für } \frac{n}{n+1} \leq \varrho,$$

$$(19'') \quad |\lambda_n^*(\varrho)| > \frac{\varrho^n}{1-\varrho} \quad \text{für } \frac{n}{n+1} > \varrho.$$

Besteht in der rechtsseitigen Ungleichung von (19') Gleichheit, ist also $\varrho = \frac{n}{n+1}$, so verschwindet

$$D_n\left(\lambda; \frac{n}{n+1}\right) D_n\left(-\lambda; \frac{n}{n+1}\right) = \lambda^{2n} \frac{(n+1)^n}{n^{n+1}} \cdot \frac{n \sin(n+1)\theta + (n+1) \sin n \theta}{\sin \theta}$$

für $\theta = \pi$, also für

¹⁰⁾ Es ist nämlich

$$\operatorname{sgn} \left[\varrho \sin(n+1) \frac{k\pi}{n + \frac{1}{2}} + \sin n \frac{k\pi}{n + \frac{1}{2}} \right] = (-1)^k \operatorname{sgn} \left[\sin \frac{k\pi}{2n+1} (\varrho - 1) \right] = (-1)^{k+1}$$

$$(k = 1, 2, \dots, n-1, n),$$

$$\operatorname{sgn} \left[\varrho \sin(n+1) \frac{k\pi}{n} + \sin n \frac{k\pi}{n} \right] = \operatorname{sgn} \sin \left(k\pi + \frac{k}{n} \pi \right) = (-1)^k$$

$$(k = 1, 2, \dots, n-1)$$

und

$$\operatorname{sgn} D_n\left(\frac{\varrho^n}{1-\varrho}\right) D_n\left(\frac{-\varrho^n}{1-\varrho}\right) = \operatorname{sgn} \left[\varrho \frac{\sin(n+1)\pi}{\sin \pi} + \frac{\sin n \pi}{\sin \pi} \right] = \operatorname{sgn} [(n+1)\varrho - n] (-1)^n,$$

$$\operatorname{sgn} D_n(\infty) D_n(-\infty) = (-1)^{n+1}.$$

$$\lambda = \frac{e^n}{\sqrt{1+2\rho \cos \alpha + \rho^2}} = \frac{\left(\frac{n}{n+1}\right)^n}{1 - \frac{n}{n+1}} = \frac{n^n}{(n+1)^{n-1}}.$$

Die größte Wurzel von $D_n\left(\lambda; \frac{n}{n+1}\right) = 0$ ist demnach rational und hat den Wert

$$(20) \quad \lambda_n^*\left(\frac{n}{n+1}\right) = \frac{n^n}{(n+1)^{n-1}} \sim \frac{n}{e}.$$

6. Um das Verhalten der absolut größten Wurzel $\lambda_n^*(\rho)$ von $D_n(\lambda; \rho) = 0$ bei veränderlichen ρ und n genauer angeben können, schicke ich folgende Hilfssätze voraus:

Hilfssatz I. Die absolut größte Wurzel $\lambda_n^*(\rho)$ von $D_n(\lambda; \rho) = 0$ ist für jedes n und für jedes ρ ($0 \leq \rho \leq 1$) positiv, also

$$(21) \quad |\lambda_n^*(\rho)| = \lambda_n^*(\rho).$$

$\lambda_n^*(\rho)$ ist nämlich gleich dem absolut größten Werte, den die reelle quadratische Form $Q_n(\rho; x_0, x_1, \dots, x_{n-1}) = \sum_{k=0}^{n-1} \rho^k (x_0 x_k + x_1 x_{k-1} + \dots + x_k x_0)$ bei der Nebenbedingung $x_0^2 + x_1^2 + \dots + x_{n-1}^2 = 1$ annehmen kann, und dieser Wert ist, da die Koeffizienten von Q nichtnegativ sind, jedenfalls positiv¹¹⁾.

Hilfssatz II. Bei festem n wächst $\lambda_n^*(\rho)$ monoton mit ρ , d. h.

$$(22) \quad \lambda_n^*(\rho') > \lambda_n^*(\rho) \quad \text{für} \quad \rho' > \rho.$$

Es sei $x_0^*, x_1^*, \dots, x_{n-1}^*$ ein nichtnegatives Wertsystem⁸⁾, für welches $Q_n(\rho; x_0, x_1, \dots, x_{n-1})$ sein absolutes Maximum annimmt, also $Q_n(\rho; x_0^*, x_1^*, \dots, x_{n-1}^*) = \lambda_n^*(\rho)$. Dann ist für $\rho' > \rho$ offenbar

$$\lambda_n^*(\rho') \geq Q_n(\rho'; x_0^*, x_1^*, \dots, x_{n-1}^*) > Q_n(\rho; x_0^*, x_1^*, \dots, x_{n-1}^*) = \lambda_n^*(\rho),$$

was zu beweisen war.

Hilfssatz III. Bei festem ρ wächst $\lambda_n^*(\rho)$ monoton mit n , d. h.

$$(23) \quad \lambda_{n+1}^*(\rho) \geq \lambda_n^*(\rho).$$

Wenn nämlich $Q_n(\rho; x_0, x_1, \dots, x_{n-1})$ sein absolutes Maximum für das nichtnegative Wertsystem $x_0^*, x_1^*, \dots, x_{n-1}^*$ annimmt, so nimmt $Q_{n+1}(\rho; x_0, x_1, \dots, x_n)$ für das Wertsystem $x_0 = x_0^*, x_1 = x_1^*, \dots, x_{n-1} = x_{n-1}^*, x_n = 0$ (welches ebenfalls der Nebenbedingung $\sum_{r=0}^n x_r^2 = 1$ genügt) offenbar einen größeren Wert an, also

$$\lambda_{n+1}^*(\rho) \geq Q_{n+1}(\rho; x_0^*, x_1^*, \dots, x_{n-1}^*, 0) > Q_n(\rho; x_0^*, x_1^*, \dots, x_{n-1}^*) = \lambda_n^*(\rho)$$

wie behauptet wurde.

¹¹⁾ Vgl. Fejér, loc. cit. ⁷⁾, S. 110—111.

Hilfssatz IV. Bei festem ϱ ($0 \leq \varrho < 1$) und für unbegrenzt wachsendes n hat $\lambda_n^*(\varrho)$ den Grenzwert $\frac{1}{1-\varrho^2}$, d. h.

$$(24) \quad \lambda_n^*(\varrho) \rightarrow \frac{1}{1-\varrho^2} \quad \text{für } n \rightarrow \infty \quad (0 \leq \varrho < 1).$$

Die Summe der Wurzeln von $D_n(\lambda; \varrho) = 0$ ist offenbar gleich der Summe $1 + \varrho^2 + \varrho^4 + \dots + \varrho^{2\left[\frac{n-1}{2}\right]}$ der Hauptminoren erster Ordnung der Determinante (14), hat also für $\varrho < 1$ und $n \rightarrow \infty$ den Grenzwert $\frac{1}{1-\varrho^2}$; die Summe der absoluten Beträge von $n-1$ Wurzeln von $D_n(\lambda; \varrho) = 0$ bleibt nach (18) unterhalb der Größe $\frac{(n-1)\varrho^n}{1-\varrho}$, die für $\varrho < 1$ und $n \rightarrow \infty$ dem Grenzwert 0 zustrebt; folglich strebt die absolut größte Wurzel $\lambda_n^*(\varrho)$ (nach Hilfssatz III monoton wachsend) dem Grenzwerte $\frac{1}{1-\varrho^2}$ zu.

Die in diesem Paragraphen gewonnenen Ergebnisse lassen sich zusammenfassen im

Satz I. Die absolut größte Wurzel $\lambda_n^*(\varrho)$ der charakteristischen Gleichung n -ten Grades

$$(14) \quad D_n(\lambda; \varrho) = \begin{vmatrix} 1-\lambda & \varrho & \varrho^2 & \dots & \varrho^{n-2} & \varrho^{n-1} \\ \varrho & \varrho^2 - \lambda & & & \varrho^{n-1} & 0 \\ \varrho^2 & & & & & \\ \vdots & & & & & \vdots \\ \varrho^{n-2} & \varrho^{n-1} & & & -\lambda & 0 \\ \varrho^{n-1} & 0 & & \dots & 0 & -\lambda \end{vmatrix} = 0, \quad 0 \leq \varrho \leq 1,$$

ist positiv und wächst monoton mit n , sowie mit ϱ .

Es ist bei festem n ($n \geq 2$)

$$(25) \quad \begin{aligned} 1 < \lambda_n^*(\varrho) < \frac{n^n}{(n+1)^{n-1}} & \quad \text{für } 0 < \varrho < \frac{n}{n+1}, \\ \lambda_n^*\left(\frac{n}{n+1}\right) = \frac{n^n}{(n+1)^{n-1}} & \quad \text{'' } \varrho = \frac{n}{n+1}, \\ \frac{n^n}{(n+1)^{n-1}} < \lambda_n^*(\varrho) < \frac{1}{2 \cos \frac{n\pi}{2n+1}} & \quad \text{'' } \frac{n}{n+1} < \varrho < 1, \\ \lambda_n^*(1) = \frac{1}{2 \cos \frac{n\pi}{2n+1}} & \quad \text{'' } \varrho = 1. \end{aligned}$$

Bei festem $\varrho < 1$ strebt $\lambda_n^*(\varrho)$ für $n \rightarrow \infty$ monoton wachsend dem Grenzwerte $\frac{1}{1-\varrho^2}$ zu, während alle übrigen Wurzeln von $D_n(\lambda; \varrho) = 0$ dem Grenzwerte 0 zustreben.

$$f(z) = \left(\frac{n}{n+1}\right)^{n-1} + \left(\frac{n}{n+1}\right)^{n-2} z + \dots + 1 \cdot z^{n-1} + \gamma_n z^n + \dots,$$

ist für jedes $f(z)$

$$(30) \quad \max_{|z| \leq 1} |f(z)| \geq \frac{n^n}{(n+1)^{n-1}}$$

und Gleichheit gilt nur für die rationale Funktion

$$(31) \quad h^*(z) = \frac{n^n}{(n+1)^{n-1}} \cdot \frac{1 + 2z + 3z^2 + \dots + nz^{n-1}}{n + (n-1)z + \dots + z^{n-1}}.$$

§ 3.

Bestimmung der Grenzen der Koeffizientensummen für eine Potenzreihe, deren Mittelmodul eine gegebene Schranke hat.

9. Ich betrachte zuerst folgende allgemeinere Frage:

Gegeben sind die n positiven, monoton zunehmenden Zahlen

$$0 < c_0 \leq c_1 \leq \dots \leq c_{n-1};$$

welches ist die obere Grenze des Ausdrucks

$$(32) \quad L(f) = |c_{n-1}a_0 + c_{n-2}a_1 + \dots + c_0a_{n-1}|$$

im Bereiche der für $|z| \leq 1$ konvergenten Potenzreihen

$$(33') \quad f(z) = a_0 + a_1 z + a_2 z^2 + \dots,$$

die der Bedingung

$$(33'') \quad \frac{1}{2\pi} \int_0^{2\pi} |f(e^{it})| dt \leq 1$$

genügen?¹²⁾

¹²⁾ Die Frage nach der oberen Grenze des Ausdrucks $L(f) = \left| \sum_{\nu=0}^{n-1} c_{n-\nu-1} a_\nu \right|$ im Bereiche der für $|z| < 1$ konvergenten Potenzreihen $\sum_{\nu=0}^{\infty} a_\nu z^\nu$, die daselbst der Bedingung $\left| \sum_{\nu=0}^{\infty} a_\nu z^\nu \right| \leq 1$ genügen, wurde von Herrn O. Szász allgemein gelöst. (Korlátos hatványsorok együtthatóiról, Math. és term. ért. Budapest 43 (1926), S. 488–502.

Auf die Tatsache, daß die von ihm dort angewendete, einen ursprünglich von Herrn Landau herrührenden Gedanken benutzende Methode — entsprechend modifiziert — auch bei dem im Texte betrachteten, verwandten (namentlich die Nebenbedingung $\frac{1}{2\pi} \int_0^{2\pi} |f(e^{it})| dt \leq 1$ enthaltenden) Fejérschen Maximumprobleme zum Ziele

führt, bin ich durch eine mündliche Mitteilung von Herrn O. Szász aufmerksam gemacht worden. Man vgl. auch L. Fejér, loc. cit. S. 114.

Es ist nach der Cauchyschen Koeffizientenformel

$$\sum_{\nu=0}^{n-1} c_{n-\nu-1} a_{\nu} = \frac{1}{2\pi i} \int_{|z|=1} \frac{f(z)}{z^n} (c_0 + c_1 z + \dots + c_{n-1} z^{n-1} + \gamma_n z^n + \dots) dz,$$

wo $\gamma_n, \gamma_{n+1}, \dots$ beliebig sind; also, wenn $c_0 + \dots + c_{n-1} z^{n-1} + \gamma_n z^n + \dots$ mit $h(z)$ bezeichnet wird,

$$(34) \quad L(f) = \left| \sum_{\nu=0}^{n-1} c_{n-\nu-1} a_{\nu} \right| = \frac{1}{2\pi} \left| \int_0^{2\pi} \frac{f(e^{it}) h(e^{it})}{e^{(n-1)it}} dt \right| \leq \frac{1}{2\pi} \int_0^{2\pi} |f(e^{it}) h(e^{it})| dt \\ \leq \frac{1}{2\pi} \text{Max} |h(e^{it})| \int_0^{2\pi} |f(e^{it})| dt \leq \text{Max} |h(e^{it})|.$$

Sollen in diesen Ungleichungen für eine Funktion $f^*(z)$ der Gesamtheit (33) überall die Gleichheitszeichen gültig sein, so müssen folgende Bedingungen erfüllt werden:

I. $\text{sgn } f^*(e^{it}) h(e^{it}) = e^{(n-1)it}$,¹³⁾

II. $|h(e^{it})| = \text{Max} |h(e^{it})|$, $0 \leq t \leq 2\pi$,

III. $\frac{1}{2\pi} \int_0^{2\pi} |f^*(e^{it})| dt = 1$.

Nach Bedingung I. muß also $f^*(z) h(z)$ folgende Form haben¹⁴⁾:

$$(35) \quad f^*(z) h(z) = g(z) \cdot z^{n-1} \bar{g}\left(\frac{1}{z}\right), \quad g(z) \neq 0 \text{ in } |z| < 1,$$

wo $g(z)$ ein sonst beliebiges Polynom $n-1$ -ter Ordnung bedeutet.

¹³⁾ $\text{sgn } f(z) = \frac{f(z)}{|f(z)|}$, $\overline{\text{sgn } f(z)} = \text{sgn } \bar{f}(\bar{z})$. \bar{z} ist der konjugiert komplexe Wert

von z ; $\bar{f}(\bar{z})$ bedeutet diejenige Funktion, die aus $f(z)$ entsteht, wenn in ihr die Koeffizienten durch ihre konjugiert komplexen Werte ersetzt werden.

¹⁴⁾ Hat nämlich eine Funktion $F(z)$ die Form $g(z) z^{n-1} \bar{g}\left(\frac{1}{z}\right)$, wo $g(z)$ ein Polynom $n-1$ -ter Ordnung ist, so ist offenbar $\text{sgn } F(e^{it}) = \text{sgn } g(e^{it}) \bar{g}(e^{-it}) e^{(n-1)it} = e^{(n-1)it}$. Wenn umgekehrt für eine in $|z| \leq 1$ reguläre Funktion $F(z)$ $\text{sgn } F(e^{it}) = e^{(n-1)it}$ ist, so hat $F(z)$ im Einheitskreise genau $n-1$ Nullstellen. Es sei $z^{n-1} \bar{g}\left(\frac{1}{z}\right)$ das Polynom, dessen Nullstellen mit denjenigen von $F(z)$ in $|z| \leq 1$ zusammenfallen.

Dann ist der Imaginärteil der in $|z| \leq 1$ regulären Funktion $\log \frac{F(z)}{g(z) z^{n-1} \bar{g}\left(\frac{1}{z}\right)}$

gleich 0, also $\log \frac{F(z)}{g(z) z^{n-1} \bar{g}\left(\frac{1}{z}\right)}$ eine Konstante, d. h. $F(z) = C \cdot g(z) z^{n-1} \bar{g}\left(\frac{1}{z}\right)$,

$g(z) \neq 0$ in $|z| < 1$.

Nach II. hat die für $|z| \leq 1$ konvergente Potenzreihe $h(z)$ am Rande des Einheitskreises konstanten absoluten Betrag, ist also mit der zu den Koeffizienten c_0, c_1, \dots, c_{n-1} gehörigen Carathéodory-Fejérschen kanonischen Funktion

$$h^*(z) = \lambda_n^* \frac{x_{n-1}^* + x_{n-2}^* z + \dots + x_0^* z^{n-1}}{x_0^* + x_1^* z + \dots + x_{n-1}^* z^{n-1}} = \lambda^* \frac{z^{n-1} g^*\left(\frac{1}{z}\right)}{g^*(z)}$$

identisch.

Soll aber die in $|z| \leq 1$ reguläre Funktion $f^*(z)$ der Bedingung (35) genügen, also die Form haben

$$f^*(z) = \frac{1}{\lambda^*} \frac{g^*(z)}{z^{n-1} g^*\left(\frac{1}{z}\right)} g(z) z^{n-1} \bar{g}\left(\frac{1}{z}\right),$$

so muß — da die Nullstellen von $z^{n-1} g^*\left(\frac{1}{z}\right)$ alle im Einheitskreise liegen — $\bar{g}\left(\frac{1}{z}\right)$ mit $C g^*\left(\frac{1}{z}\right)$ zusammenfallen. Demnach ist

$$(36) \quad f^*(z) = \frac{C}{\lambda^*} (g^*(z))^2 = \frac{C}{\lambda^*} (x_0^* + x_1^* z + \dots + x_{n-1}^* z^{n-1})^2.$$

Aus III. bestimmt sich endlich die Konstante C ; es ist

$$\begin{aligned} \frac{1}{2\pi} \int_0^{2\pi} |f^*(e^{it})| dt &= \frac{1}{2\pi} \int_0^{2\pi} \frac{|C|}{\lambda^*} |x_0^* + x_1^* z + \dots + x_{n-1}^* z^{n-1}|_{z=e^{it}}^2 dt \\ &= \frac{|C|}{\lambda^*} \sum_0^{n-1} x_\nu^{*2} = \frac{|C|}{\lambda^*} = 1, \end{aligned}$$

d. h. $|C| = \lambda^*$.

Es ist also für $f^*(z)$

$$(37') \quad L(f^*) = \text{Max} |h^*(e^{it})| = |h^*(e^{it})| = \lambda^*$$

für alle anderen Funktionen $f(z)$ der Gesamtheit (33)

$$(37'') \quad L(f) < \lambda^*.$$

Mit Berücksichtigung der in der Einleitung Nr. 3 angegebenen Bedeutung von λ^* und x_0^*, \dots, x_{n-1}^* kann ich also das Ergebnis folgendermaßen aussprechen:

Satz IV. *Es seien c_0, c_1, \dots, c_{n-1} gegebene positive, monoton zunehmende Zahlen*

$$0 < c_0 \leq c_1 \leq \dots \leq c_{n-1},$$

dann ist für die Gesamtheit aller in $|z| \leq 1$ konvergenten Potenzreihen

$$(33') \quad f(z) = a_0 + a_1 z + a_2 z^2 + \dots,$$

die der Bedingung

$$(33'') \quad \frac{1}{2\pi} \int_0^{2\pi} |f(e^{it})| dt \leq 1$$

genügen, die obere Grenze des Ausdrucks

$$(32) \quad L(f) = |c_{n-1} a_0 + c_{n-2} a_1 + \dots + c_0 a_{n-1}|$$

gleich dem absolut größten Werte, den die reelle quadratische Form

$$(4) \quad Q_n(x_0, x_1, \dots, x_{n-1}) = \sum_{\substack{p, q=0, 1, \dots, n-1 \\ 0 \leq p+q \leq n-1}} c_{n-(p+q)-1} x_p x_q,$$

bei der Nebenbedingung $\sum_{p=0}^{n-1} x_p^2 = 1$ annimmt, also gleich der größten Wurzel der charakteristischen Gleichung (5).

Für jede Funktion der Gesamtheit (33) ist demnach

$$(37) \quad L(f) \leq \text{Max}_{x_0^2 + x_1^2 + \dots + x_{n-1}^2 = 1} |Q_n(x_0, x_1, \dots, x_{n-1})| = \lambda_n^*;$$

ist $x_0^*, x_1^*, \dots, x_{n-1}^*$ das einzige positive Wertsystem, für welches

$$Q_n(x_0^*, x_1^*, \dots, x_{n-1}^*) = \lambda_n^*$$

wird, so gilt für die rationale ganze Funktion $2n-2$ -ten Grades

$$(36) \quad f^*(z) = (x_0^* + x_1^* z + \dots + x_{n-1}^* z_{n-1})^2$$

und nur für diese das Gleichheitszeichen in (37).

10. Nach einem Fejérschen Satze¹⁵⁾ läßt sich hieraus die obere Grenze des Ausdrucks

$$(38) \quad \mathfrak{E}(f) = c_{n-1} |a_0| + c_{n-2} |a_1| + \dots + c_0 |a_{n-1}|$$

für die Gesamtheit (33) der Funktionen $f(z)$ unmittelbar ableiten. Nach diesem Satze ist nämlich, wenn $H^*(z)$ die zu den Koeffizienten $c_0 \overline{\text{sgn}} a_{n-1}, c_1 \overline{\text{sgn}} a_{n-2}, \dots, c_{n-1} \overline{\text{sgn}} a_0$ gehörige kanonische Funktion, $h^*(z)$ aber die zu den positiven Koeffizienten c_0, c_1, \dots, c_{n-1} gehörige kanonische Funktion bedeutet,

$$|h^*(e^{it})| \geq |H^*(e^{it})|,$$

also

$$(39) \quad \mathfrak{E}(f) = \sum_{\nu=0}^{n-1} c_{n-\nu-1} |a_\nu| = \frac{1}{2\pi} \int_0^{2\pi} \frac{f(e^{it}) H^*(e^{it})}{e^{(n-1)it}} dt \leq |H^*(e^{it})| \leq |h^*(e^{it})|.$$

Für die Funktion $f^*(z)$ in (36) wird aber, da ihre Koeffizienten positiv sind,

$$(39') \quad \mathfrak{E}(f^*) = L(f^*) = \lambda_n^* = |h^*(e^{it})|;$$

für die Funktion $f^*(z)$ gilt also das Gleichheitszeichen in (39).

Satz IV'. Bei denselben Bedingungen wie im Satze IV hat der Ausdruck

$$(38) \quad \mathfrak{E}(f) = c_{n-1} |a_0| + c_{n-2} |a_1| + \dots + c_0 |a_{n-1}|$$

¹⁵⁾ loc. cit. S. 120—121.

dieselbe obere Grenze wie der Ausdruck $L(f)$ in (32) und wird für dieselbe Funktion $f^*(z)$ erreicht.

11. Um die obere Grenze der Koeffizientensummen sämtlicher in $|z| \leq 1$ konvergenter Potenzreihen, deren Mittelmodul nicht größer als 1 ist, zu bestimmen, habe ich nur die Ergebnisse der vorigen Nummer für den Fall $c_0 = c_1 = \dots = c_{n-1} = 1$ in Anwendung zu bringen. In diesem Falle ist nach (13) $\lambda_n^* = \frac{1}{2 \cos \frac{n\pi}{2n+1}}$ und $x_p^* = \frac{2}{\sqrt{2n+1}} \sin \frac{(n-p)\pi}{2n+1}$ ($p = 0, 1, \dots, n-1$).

Also gilt

Satz V. Für die Gesamtheit aller in $|z| \leq 1$ konvergenter Potenzreihen

$$(33') \quad f(z) = a_0 + a_1 z + a_2 z^2 + \dots,$$

deren Mittelmodul der Ungleichung

$$(33'') \quad \frac{1}{2\pi} \int_0^{2\pi} |f(e^{it})| dt \leq 1$$

genügt, ist

$$(40) \quad |a_0 + a_1 + \dots + a_{n-1}| \leq \frac{1}{2 \cos \frac{n\pi}{2n+1}},$$

und Gleichheit gilt nur für die rationale ganze Funktion $2n-2$ -ten Grades

$$(41) \quad f^*(z) = \frac{4}{2n+1} \left(\sin \frac{n\pi}{2n+1} + \sin \frac{(n-1)\pi}{2n+1} z + \dots + \sin \frac{\pi}{2n+1} z^{n-1} \right)^2.$$

Für dieselbe Gesamtheit ist gleichzeitig (nach Satz IV')

$$(42) \quad |a_0| + |a_1| + \dots + |a_{n-1}| \leq \frac{1}{2 \cos \frac{n\pi}{2n+1}}$$

und Gleichheit gilt nur für die in (41) angegebene Funktion $f^*(z)$.

§ 4.

Über die Majorante von Potenzreihen, deren Mittelmodul eine gegebene Schranke hat.

12. Ich bestimme zuerst die obere Grenze des Ausdrucks

$$(43) \quad \mathfrak{C}(f) = |a_0| + |a_1| \varrho + \dots + |a_{n-1}| \varrho^{n-1}$$

bei festem ϱ ($0 \leq \varrho < 1$) für die Gesamtheit aller in $|z| \leq 1$ konvergenter Potenzreihen

$$(33') \quad f(z) = a_0 + a_1 z + a_2 z^2 + \dots,$$

deren Mittelmodul der Ungleichung genügt

$$(33'') \quad \frac{1}{2\pi} \int_0^{2\pi} |f(e^{it})| dt \leq 1.$$

Es ist nach Satz IV und IV'

$$(44) \quad \mathfrak{C}(f) = |a_0| + |a_1|\varrho + \dots + |a_{n-1}|\varrho^{n-1} \leq \lambda_n^*(\varrho),$$

wo $\lambda_n^*(\varrho)$ die absolut größte Wurzel der charakteristischen Gleichung (14) bedeutet; ist $x_0^*, x_1^*, \dots, x_{n-1}^*$ das einzige positive Wertsystem, für welches $Q_n(\varrho; x_0, \dots, x_{n-1})$ ihr absolutes Maximum erreicht, so gilt für die rationale ganze Funktion $2n-2$ -ten Grades

$$f_n^*(z) = C(x_0^* + x_1^*z + \dots + x_{n-1}^*z^{n-1})^2, \quad C(x_0^{*2} + \dots + x_{n-1}^{*2}) = 1$$

und nur für diese das Gleichheitszeichen in (44).

Nach Satz I ist aber

$$\lambda_n^*(\varrho) < \frac{1}{1-\varrho^2} \text{ für jedes } n \text{ und } \lambda_n^*(\varrho) \rightarrow \frac{1}{1-\varrho^2} \text{ für } n \rightarrow \infty,$$

demnach ist

$$(45) \quad \mathfrak{M}_f(\varrho) = |a_0| + |a_1|\varrho + |a_2|\varrho^2 + \dots \leq \frac{1}{1-\varrho^2}$$

und $\lambda_n^*(\varrho) < \mathfrak{M}_{f_n^*}(\varrho) = f_n^*(\varrho) < \lambda_{2n-1}^*(\varrho)$, d. h.

$$(46) \quad \mathfrak{M}_{f_n^*}(\varrho) = f_n^*(\varrho) \rightarrow \frac{1}{1-\varrho^2} \text{ für } n \rightarrow \infty.$$

Damit ist bewiesen der

Satz VI. Für die Gesamtheit aller in $|z| \leq 1$ konvergenten Potenzreihen

$$(33') \quad f(z) = a_0 + a_1z + a_2z^2 + \dots,$$

deren Mittelmodul der Ungleichung

$$(33'') \quad \frac{1}{2\pi} \int_0^{2\pi} |f(e^{it})| dt \leq 1$$

genügt, ist für jedes ϱ ($0 \leq \varrho < 1$)

$$(45) \quad \mathfrak{M}_f(\varrho) = |a_0| + |a_1|\varrho + |a_2|\varrho^2 + \dots \leq \frac{1}{1-\varrho^2}$$

und es gibt für jedes ϱ eine Folge von rationalen ganzen Funktionen $f_n^*(z)$ von der Beschaffenheit, daß

$$(46) \quad \mathfrak{M}_{f_n^*}(\varrho) = f_n^*(\varrho) \rightarrow \frac{1}{1-\varrho^2} \text{ für } n \rightarrow \infty.$$

Zusatz bei der Korrektur. Mit Benutzung einer von Herrn M. Fekete herrührenden Bemerkung läßt sich auch diejenige Potenzreihe $f^*(z)$ der Gesamtheit (33') bestimmen, für welche in (45) das Gleichheitszeichen gilt. Es ist nämlich für

$$f^*(z) = (1 - \varrho^2)(1 + 2\varrho z + 3\varrho^2 z^2 + \dots) = \frac{1 - \varrho^2}{(1 - \varrho z)^2},$$

$$\frac{1}{2\pi} \int_0^{2\pi} |f^*(e^{it})| dt = (1 - \varrho^2)(1 + \varrho^2 + \varrho^4 + \dots) = 1$$

und

$$\mathfrak{M}_{f^*}(\varrho) = f^*(\varrho) = (1 - \varrho^2)(1 + 2\varrho^2 + 3\varrho^4 + \dots) = \frac{1}{1 - \varrho^2}.$$

Über die Beziehungen der Eindeutigkeitsfragen in den Theorien der trigonometrischen Reihen und Integrale.

Von

A. Zygmund in Warschau.

Kapitel 1.

§ 1.

Wir nennen ein trigonometrisches Integral (kurz t. I.) jedes Integral der Gestalt

$$(1) \int_0^{\infty} (c_s \cos sx + d_s \sin sx) ds = \frac{1}{2} \int_{-\infty}^{+\infty} a_s e^{isx} ds \quad (a_s = c_s - id_s; a_{-s} = \bar{a}_s)^1,$$

wobei wir die rechte Seite der Gleichheit als $\lim_{\omega \rightarrow +\infty} \int_{-\omega}^{+\omega}$ betrachten, wo das Limeszeichen die übliche oder auch verallgemeinerte Bedeutung haben kann. Insbesondere kann jedes Fouriersche Integral in der Gestalt (1) geschrieben werden. In der letzten Zeit sind einige den t. I. gewidmete Arbeiten erschienen, die sich insbesondere mit der Eindeutigkeitsfrage der Darstellbarkeit einer Funktion durch ein t. I. befassen²⁾. Genauer gesagt, es wird gefragt: Bei welchen Voraussetzungen ist das t. I. (1) ein Fourierintegral, d. h. wann existiert eine Funktion $f(x)$ derart, daß für fast alle s :

$$a_s = \frac{1}{\pi} \int_{-\infty}^{+\infty} f(x) e^{-isx} dx.$$

Die t. I. haben eine große formelle Ähnlichkeit mit den trigonometrischen Reihen (kurz: t. R.)

$$(2) \frac{c_0}{2} + \sum_{n=1}^{\infty} c_n \cos nx + d_n \sin nx = \sum_{n=-\infty}^{+\infty} a_n e^{inx} \quad (a_n = c_n - id_n, a_{-n} = \bar{a}_n).$$

¹⁾ Mit \bar{z} bezeichnen wir mit z konjugierte Zahl.

²⁾ Vgl. S. Pollard [1], M. Jacob [1].

Es drängt sich daher die Vermutung auf, daß gewisse Sätze, die in der Theorie der t. R. gelten, nach entsprechenden Umänderungen für die t. I. bestehen und umgekehrt. Das Hauptziel der vorliegenden Arbeit ist die Begründung dieser Vermutung in der Theorie der Eindeutigkeit der Darstellung durch ein t. I. und eine t. R. in dem Sinne, daß die Eindeutigkeitsmengen bei t. I. und t. R., grob gesprochen, dieselben sind. Wegen genaueren Angaben verweisen wir auf die entsprechenden Stellen dieser Arbeit. Die hier angewandte Methode hat auch einen weiteren Anwendungsbereich. Sie beruht auf der Theorie der formalen Multiplikation, die von Herrn A. Rajchman³⁾ angebahnt wurde, und auf den Lokalisationsformeln von Riemann⁴⁾. Zur Bequemlichkeit des Lesers gebe ich die Beweise der Multiplikations- und Lokalisationstheoreme für t. I., obwohl sie sich von den entsprechenden, schon früher publizierten Sätzen für t. R.⁵⁾ nicht wesentlich unterscheiden. Es sei auch hier bemerkt, daß wir bei der Anwendung auf die Eindeutigkeitsfragen die Multiplikations- und Lokalisationsätze nur in dem Spezialfall benutzen, wo die Koeffizienten des Integrals bzw. der Reihe gegen Null streben.

§ 2.

In dem vorliegenden Paragraphen geben wir einige ganz elementare Hilfssätze an. Wir werden sagen, daß die Funktion a_s ($-\infty < s < \infty$) der Bedingung J genügt, wenn für jedes $\omega > 0$:

$$\int_{-\omega}^{+\omega} |a_s| ds < \infty.$$

(Wenn nicht ausdrücklich das Gegenteil vorausgesetzt wird, werden wir Integrale im Lebesgueschen Sinne verstehen.)

Lemma a . Genügt a_s der Bedingung J , gilt für $\mu \rightarrow \pm \infty$: $\int_{\mu}^{\mu+1} |a_s| ds \rightarrow 0$, und hat man außerdem

$$|a_s| < M \quad (M = \text{konst.}) \quad \text{und} \quad a_s = O(s^{-2}) \quad \text{für} \quad s \rightarrow \pm \infty,$$

so erfüllt die durch die Formel

$$(3) \quad A_{\omega} = \int_{-\infty}^{+\infty} a_s \alpha_{\omega-s} ds = \int_{-\infty}^{+\infty} a_{\omega-s} \alpha_s ds$$

definierte Funktion A_{ω} gleichfalls die Bedingung J und es gilt für $\omega \rightarrow \pm \infty$: $A_{\omega} \rightarrow 0$.

³⁾ Vgl. A. Rajchman [1], [3].

⁴⁾ B. Riemann [1].

⁵⁾ A. Rajchman [1], [3], A. Zygmund [1].

Lemma b. Erfüllt a_s die Bedingung J , gilt für $\mu \rightarrow \pm \infty$:

$\int_{\mu}^{\mu+1} |a_s| ds = o(\mu^{-\gamma})$ ($\gamma > 0$)⁶⁾, und hat man

$$|\alpha_s| < M; \quad \alpha_s = O(s^{-\gamma-1}) \quad \text{für } s \rightarrow \pm \infty,$$

so erfüllt die Funktion A_ω (3) die Bedingung J und es gilt: $A_\omega = o(\omega^{-\gamma})$ ($\omega \rightarrow \pm \infty$).

Lemma c. Erfüllt a_s die Bedingung J , gilt für $\mu \rightarrow \pm \infty$:

$\int_{\mu}^{\mu+1} |a_s| ds = o(\mu^\gamma)$ ($\gamma > 0$), und ist

$$(4) \quad |\alpha_s| < M; \quad \alpha_s = O(s^{-\gamma-2}),$$

so erfüllt die Funktion A_ω gleichfalls die Bedingung J und es ist: $A_\omega = o(\omega^\gamma)$.

Die Beweise dieser Hilfssätze sind ganz leicht und es kann auf ihre Angabe hier verzichtet werden.

Lemma d. Gilt $|\alpha_s| < M$ und für $s \rightarrow \pm \infty$: $\alpha_s = O(s^{-\gamma})$ ($\gamma > n + 1$; $n = 1, 2, \dots$), hat man außerdem für alle Werte von $k = 0, 1, 2, \dots, n - 1$:

$$(5) \quad \int_{-\infty}^{+\infty} \alpha_s \cdot s^k ds = 0,$$

und setzt man

$$(6) \quad \int_s^\infty \alpha_t dt = R_s^1; \quad \int_s^\infty R_t^i dt = R_s^{i+1} \quad (i = 1, 2, \dots),$$

so erhält man

$$(7) \quad R_s^k = O(s^{-\gamma+k}) \quad (k = 1, 2, \dots, n); \quad \int_{-\infty}^{+\infty} R_s^n ds = \frac{1}{n!} \int_{-\infty}^{+\infty} \alpha_s s^n ds.$$

Der Beweis wird durch vollständige Induktion geführt: 1. für $n = 1$ ist der Satz evident, denn gilt $\alpha_s = O(s^{-\gamma})$ ($\gamma > 2$) und $\int_{-\infty}^{+\infty} \alpha_s ds = 0$, so wird

$$R_s^1 = \int_s^\infty \alpha_t dt = O(s^{-\gamma+1})$$

und

$$\int_{-N}^{+N} R_s^1 ds = [s R_s^1]_{-N}^{+N} + \int_{-N}^{+N} s \alpha_s ds,$$

und es genügt jetzt den Übergang $N \rightarrow \infty$ zu machen.

2. Es sei der Hilfssatz für alle $n < m$ bewiesen. Wir haben zuerst

$$R_s^k = O(s^{-\gamma+k}) \quad (k = 1, 2, \dots, m - 1),$$

$$\int_{-\infty}^{+\infty} R_s^{m-1} ds = \frac{1}{(m-1)!} \int_{-\infty}^{+\infty} \alpha_s s^{m-1} ds = 0,$$

⁶⁾ Der Einfachheit halber schreiben wir hier (und analog fortan) $o(\mu^{-\gamma})$ statt $o(|\mu|^{-\gamma})$.

also $R_s^m = O(s^{-\gamma+m})$. Außerdem, nach partieller Integration,

$$\int_{-N}^{+N} R_s^m ds = \left[R_s^m \cdot s + R^{m-1} \cdot \frac{s^2}{2} + \dots + R^1 \cdot \frac{s^m}{m!} \right]_{-N}^{+N} + \frac{1}{m!} \int_{-N}^{+N} \alpha_s s^m ds,$$

und man braucht nur den Übergang $N \rightarrow \infty$ zu machen.

Folgerung. Ist $|\alpha_s| < M$; $\alpha_s = O(s^{-\gamma})$ ($\gamma > n+1$) und besteht (5) für $k = 0, 1, 2, \dots, n$, so ist

$$R_s^k = O(s^{-\gamma+k}) \quad (k = 0, 1, 2, \dots, n+1).$$

Lemma e. Ist $|\alpha_s| < M$, $\alpha_s = O(s^{-\gamma})$ ($\gamma > n+1$),

$$\int_{-\infty}^{+\infty} \alpha_s ds = \alpha; \quad \int_{-\infty}^{+\infty} \alpha_s \cdot s^k ds = 0 \quad (k = 1, 2, \dots, n-1)$$

und setzt man

$$\overset{*}{R}_s^1 = \int_s^{\infty} \alpha_t dt \quad (s \geq 0); \quad \overset{*}{R}_s^1 = \int_s^{\infty} \alpha_t dt - \alpha \quad (s < 0);$$

$$\overset{*}{R}_s^{i+1} = \int_s^{\infty} \overset{*}{R}_t^i dt \quad (i = 1, 2, \dots),$$

so erhält man:

$$(8) \quad \overset{*}{R}_s^k = O(s^{-\gamma+k}) \quad (k = 1, 2, \dots, n); \quad \int_{-\infty}^{+\infty} \overset{*}{R}_s^n ds = \frac{1}{n!} \int_{-\infty}^{+\infty} \alpha_s s^n ds.$$

Der Beweis unterscheidet sich nicht von dem des Lemmas d.

Folgerung. Bei den Voraussetzungen des Lemmas e gilt die erste der Formeln (8) auch für $k = n+1$.

Bemerkung. Sind in den Hilfssätzen a–e die Größen α_s und α_x Funktionen eines Parameters x und sind für einen gewissen Wertebereich dieses Parameters die entsprechenden Bedingungen gleichmäßig erfüllt, so sind auch die Behauptungen gleichmäßig erfüllt.

§ 3.

Das formale Produkt (kurz: f. P.) der t. I.

$$(9) \quad \frac{1}{2} \int_{-\infty}^{+\infty} \alpha_s e^{isx} ds \quad (\alpha_{-s} = \bar{\alpha}_s)$$

$$(10) \quad \frac{1}{2} \int_{-\infty}^{+\infty} \alpha_s e^{isx} ds \quad (\alpha_{-s} = \bar{\alpha}_s)$$

nennen wir das Integral

$$(11) \quad \frac{1}{2} \int_{-\infty}^{+\infty} A_s e^{isx} ds, \quad \text{wo} \quad A_s = \frac{1}{2} \int_{-\infty}^{+\infty} \alpha_t \alpha_{s-t} dt.$$

Es ist evident, daß $A_{-s} = \bar{A}_s$. Die Hilfssätze a, b, c sind Sätze über Koeffizienten des f. P. Das Integral

$$(12) \quad \int_0^{\infty} (c_s \sin sx - d_s \cos sx) ds = \frac{1}{2} \int_{-\infty}^{+\infty} b_s e^{isx} ds,$$

$$\text{wo für} \quad \begin{aligned} s > 0: & b_s = -d_s - i c_s = -i a_s, \\ s < 0: & b_s = d_s + i c_s = i a_s, \end{aligned}$$

nennen wir mit dem Integral (1) konjugiert. In dem vorliegenden Paragraphen geben wir einige Sätze an, die die f. P., wie auch die mit ihnen konjugierte Integrale betreffen.

Satz I. Ist 1. $\int_{\mu}^{\mu+1} |a_s| ds \rightarrow 0$, 2. konvergiert das Integral (10) für $x < E$ gegen Null und 3. ist $\alpha_s = O(s^{-3})$, so konvergiert das f. P. (11) der Integrale (9) und (10), wie auch das mit ihm konjugierte Integral, gleichmäßig in $x < E$, und zwar das erste gegen Null.

Bemerkung. Wir setzen immer stillschweigend voraus, daß a_s der Bedingung J genügt⁷⁾.

Beweis. Ich setze zur Vereinfachung $x = 0$ voraus:

$$\begin{aligned} S_{\omega} &= \frac{1}{2} \int_{-\omega}^{+\omega} A_s ds = \frac{1}{4} \int_{-\omega}^{+\omega} ds \int_{-\infty}^{+\infty} a_t \alpha_{s-t} dt = \frac{1}{4} \int_{-\infty}^{+\infty} a_t dt \left[\int_{-\omega-t}^{\omega-t} \alpha_s ds \right] \\ &= \frac{1}{4} \int_{-\infty}^{+\infty} a_t [R_{-\omega-t} - R_{\omega-t}] dt \rightarrow 0, \end{aligned}$$

wegen $R_s = \int_s^{\infty} \alpha_t dt = O(s^{-2})$ und nach dem Lemma a .

$$\begin{aligned} (13) \quad T_{\omega} &= \frac{1}{2} \int_{-\omega}^{+\omega} B_s ds = \frac{i}{2} \int_0^{\omega} (\bar{A}_s - A_s) ds = \frac{i}{4} \int_0^{\omega} \left[\int_{-\infty}^{+\infty} (\bar{a}_t \bar{\alpha}_{s-t} - a_t \alpha_{s-t}) dt \right] ds \\ &= \frac{i}{4} \int_{-\infty}^{+\infty} \bar{a}_t \left(\int_{-t}^{\omega-t} \bar{\alpha}_u du \right) dt - \frac{i}{4} \int_{-\infty}^{+\infty} a_t \left(\int_{-t}^{\omega-t} \alpha_u du \right) dt \\ &= \frac{i}{4} \int_{-\infty}^{+\infty} (\bar{a}_t \bar{R}_{-t} - a_t R_{-t}) dt - \frac{i}{4} \int_{-\infty}^{+\infty} (\bar{a}_t \bar{R}_{\omega-t} - a_t R_{\omega-t}) dt. \end{aligned}$$

Auf Grund des Lemmas a strebt der Subtrahend gegen Null.

⁷⁾ Es genügt hier und weiter (insbesondere in den Eindeutigkeitsätzen) voraussetzen, daß $\int_0^{\omega} (|c_s| + s |d_s|) ds < \infty$ für jedes $\omega > 0$ (vgl. (1)).

Ist $\int_{\mu}^{\mu+1} |a_s| ds = O(1)$, so sind offenbar die Teilintegrale $\int_{-\omega}^{+\omega} A_s ds$ und $\int_{-\omega}^{+\omega} B_s ds$ gleichmäßig in $x \in E$ beschränkt. Analog in weiteren Sätzen.

Satz II. Genügt a_s den im Satz I angegebenen Bedingungen, gilt weiter $\alpha_s = O(s^{-3})$ und ist

$$(14) \quad \frac{1}{2} \int_{-\infty}^{+\infty} \alpha_s e^{isx} ds = \alpha(x) \quad x \in E,$$

so konvergieren für $\omega \rightarrow \infty$ die Ausdrücke

$$(15) \quad \frac{1}{2} \int_{-\omega}^{+\omega} A_s e^{isx} ds - \frac{\alpha(x)}{2} \int_{-\omega}^{+\omega} a_s e^{isx} ds, \quad \frac{1}{2} \int_{-\omega}^{+\omega} B_s e^{isx} ds - \frac{\alpha(x)}{2} \int_{-\omega}^{+\omega} b_s e^{isx} ds,$$

gleichmäßig in E , und zwar der erste gegen Null.

Bemerkung. Ist $\alpha(x) \neq 0$, so sagt der Satz II aus, daß das f. P. (bzw. das damit konjugierte Integral) sich in derselben Weise verhält, wie das Integral (9) (bzw. das konjugierte). Z. B., ist das t. I. (9) konvergent, so konvergiert auch das Integral (11) und umgekehrt. Dasselbe gilt für die Cesàrosche Summabilität usw. Ist in E : $|\alpha(x)| \geq \varepsilon > 0$, so folgt aus der gleichmäßigen Konvergenz bzw. der Summabilität des t. I. (11) dasselbe für das Integral (9).

Beweis. Wir setzen wieder $x = 0$ und es sei $a(0) = \alpha$:

$$S_{\omega} = \frac{1}{4} \int_{-\infty}^{+\infty} a_t (R_{-\omega-t} - R_{\omega-t}) dt.$$

Ich setze

$$\overset{*}{R}_t = R_t \quad \text{für } t \geq 0 \quad \text{und} \quad \overset{*}{R}_t = R_t - 2\alpha \quad \text{für } t < 0.$$

Es ist leicht zu zeigen, daß

$$\begin{aligned} R_{-\omega-t} - R_{\omega-t} &= \overset{*}{R}_{-\omega-t} - \overset{*}{R}_{\omega-t} + 2\alpha && \text{(für } \omega < t \leq \omega) \\ &= \overset{*}{R}_{-\omega-t} - \overset{*}{R}_{\omega-t} && \text{(sonst).} \end{aligned}$$

Also

$$(16) \quad S_{\omega} = \frac{\alpha}{2} \int_{-\omega}^{+\omega} a_t dt + \frac{1}{4} \int_{-\infty}^{+\infty} a_t [\overset{*}{R}_{-\omega-t} - \overset{*}{R}_{\omega-t}] dt,$$

wobei der zweite Summand rechts gleichmäßig gegen Null strebt (Lemma a). Berücksichtigen wir analog, daß

$$R_{-t} - R_{\omega-t} = \begin{matrix} \overset{*}{R}_{-t} - \overset{*}{R}_{\omega-t} + 2\alpha & (\text{für } 0 < t \leq \omega) \\ \overset{*}{R}_{-t} - \overset{*}{R}_{\omega-t} & (\text{sonst}), \end{matrix}$$

so können wir den Ausdruck (13) für T_ω in der Gestalt umschreiben:

$$T_\omega = \frac{i}{4} \int_{-\infty}^{+\infty} [\bar{a}_t (\overset{*}{R}_{-t} - \overset{*}{R}_{\omega-t}) - a_t (\overset{*}{R}_{-t} - \overset{*}{R}_{\omega-t})] dt + \frac{\alpha i}{2} \int_0^\omega (\bar{a}_t - a_t) dt,$$

Jetzt genügt es zu bemerken, daß das letzte Glied rechts eben $\frac{\alpha}{2} \int_{-\infty}^{+\infty} b_t dt$ ist. Man sagt⁸⁾, das Integral

$$(17) \quad \int_0^\infty \alpha_u du$$

sei (C, γ) ($\gamma > 0$) zu s summierbar, wenn für $\omega \rightarrow \infty$

$$(18) \quad S_\omega^\gamma = \frac{1}{\omega^\gamma} \int_0^\omega \alpha_u (\omega - u)^\gamma du \rightarrow s.$$

Setzen wir $\int_0^\omega \alpha_u du = s_\omega$ und integrieren partiell, so können wir die Gleichheit (12) wie folgt umschreiben:

$$(19) \quad S_\omega^\gamma = \frac{\gamma}{\omega^\gamma} \int_0^\omega s_u (\omega - u)^{\gamma-1} du.$$

Diese Gleichheit stellt die Erklärung des (C, γ) -Grenzwertes der Funktion S_ω für $\omega \rightarrow \infty$ dar. Ist (18) erfüllt, so existiert für jedes $\gamma' > \gamma$ $\lim S_\omega^{\gamma'}$, der ebenfalls gleich s ist. Man sagt weiter, daß das Integral (17) bzw. die Funktion s_ω (s. o.) nach Poisson (kurz P.) zum Werte s summierbar (bzw. limitierbar) ist, wenn bei $0 < r < 1$

$$(20) \quad \lim_{r \rightarrow 1} \int_0^\infty \alpha_u r^u du = s \quad \text{bzw.} \quad \lim_{r \rightarrow 1} \lg r \int_0^\infty s_u r^u du = s$$

ist. (Wir setzen stillschweigend voraus, daß für jedes $0 < r < 1$ die in den Formeln (20) rechts stehenden Integrale konvergieren.) Konvergiert das Integral (17), bzw. die Funktion s_t , schlechthin gegen den Grenzwert s , so finden die Gleichheiten (20) statt. Für die (C, γ) -Summabilität ($\gamma > 0$) besteht der analoge Satz, wenn wir noch die Voraussetzung hinzufügen, daß für jedes $0 < r < 1$ die in den Formeln (20) rechts stehenden Integrale konvergieren, was gewiß in allen von uns betrachteten Fällen stattfindet.

⁸⁾ Vgl. z. B. Hardy und Riesz [1].

Satz III. Ist $\int_{\mu}^{\mu+1} |\alpha_s| ds = o(\mu^\gamma)$ ($\gamma > 0$)

$$(21) \quad |\alpha_s| < M, \quad \alpha_s = O(s^{-2\gamma-3})^9$$

und gilt für $x \in E$

$$\frac{1}{2} \int_{-\infty}^{+\infty} \alpha_s e^{isx} ds = \frac{i}{2} \int_{-\infty}^{+\infty} \alpha_s e^{isx} \cdot s \cdot ds = \dots = \frac{i^k}{2} \int_{-\infty}^{+\infty} \alpha_s e^{isx} s^k ds = 0,$$

wo k die kleinste ganze Zahl $\geq \gamma$ bedeutet, so ist das f. P. der Integrale (9) und (10), wie auch das konjugierte Integral, für $x \in E$ gleichmäßig (C, γ) -summierbar, und zwar das erste gegen Null.

Beweis.

$$S_\omega = \frac{1}{4} \int_{-\infty}^{+\infty} a_t R_{-\omega-t} dt - \frac{1}{4} \int_{-\infty}^{+\infty} a_t R_{\omega-t} dt$$

und es genügt z. B. zu zeigen, daß das zweite rechts stehende Integral (C, γ) gegen Null strebt, d. h.

$$\frac{\gamma}{\omega^\gamma} \int_0^\omega \left[\int_{-\infty}^{+\infty} a_t R_{s-t} dt \right] (\omega - s)^{\gamma-1} ds \rightarrow 0.$$

Der Ausdruck links ist gleich

$$(22) \quad \frac{\gamma}{\omega^\gamma} \int_{-\infty}^{+\infty} a_t dt \left[\int_0^\omega R_{s-t} (\omega - s)^{\gamma-1} ds \right].$$

(Die Vertauschung der Reihenfolge der Integrationen ist vollkommen berechtigt.) Integrieren wir partiell, so erhalten wir für $j < \gamma$ (vgl. (6), (7) und (21)):

$$(23) \quad \int_0^\omega R_{s-t} (\omega - s)^{\gamma-1} ds = [-R_{s-t}^1 (\omega - s)^{\gamma-1} + (\gamma - 1) R_{s-t}^2 (\omega - s)^{\gamma-2} \\ - \dots + (-1)^j R_{s-t}^j (\omega - s)^{\gamma-j} \cdot (\gamma - 1) \dots (\gamma - j + 1)]_0^\omega + \\ + (-1)^j (\gamma - 1) \dots (\gamma - j) \int_0^\omega R_{s-t}^j (\omega - s)^{\gamma-j-1} ds = O(\omega^{\gamma-1} t^{-2\gamma-3+j}) \\ + (-1)^j (\gamma - 1) \dots (\gamma - j) \int_0^\omega R_{s-t}^j (\omega - s)^{\gamma-j-1} ds = I + II.$$

Der Teil des Ausdruckes (22), welcher aus I hervorgeht, ist offenbar $o(1)$. Es gibt jetzt zwei Möglichkeiten: 1. $\gamma = k$, 2. $\gamma < k$. Betrachten wir

⁹⁾ Ersetzt man in (4) $-\gamma - 2$ durch $-\gamma - 1 - \xi$ ($\xi > 0$), so kann man auch (wie leicht aus dem Beweise folgt) in (21) $-2\gamma - 3$ durch $-2\gamma - 2$ ersetzen.

zuerst den Fall 1. und setzen wir $j = \gamma - 1$. Dann wird

$$\int_0^{\omega} R_{s-t}^j (\omega - s)^{\gamma-j-1} ds = \int_0^{\omega} R_{s-t}^{\gamma-1} ds = -R_{\omega-t}^{\gamma} + R_{-t}^{\gamma}$$

und der Ausdruck (22) erhält die Gestalt:

$$o(1) \pm \frac{\gamma!}{\omega^{\gamma}} \int_{-\infty}^{+\infty} a_t R_{-t}^{\gamma} dt \mp \int_{-\infty}^{+\infty} a_t R_{\omega-t}^{\gamma} dt = o(1) \mp \frac{\gamma}{\omega^{\gamma}} \int_{-\infty}^{+\infty} a_t R_{\omega-t}^{\gamma} dt.$$

Da aber $R_s^{\gamma} = O(s^{-\gamma-3})$ (Lemma *d*), so ist auch das letzte Glied $o(1)$ (Lemma *c*) und im Falle 1. ist der Satz bewiesen.

2. Wir setzen in der Formel (23) $j = [\gamma]$ und es sei $\gamma = j + \xi$ ($0 < \xi < 1$):

$$\begin{aligned} & \int_0^{\omega} R_{s-t}^j (\omega - s)^{\xi-1} ds = - \int_0^{\omega} \frac{d}{ds} R_{s-t}^{j+1} (\omega - s)^{\xi-1} ds = \\ & = - \int_0^{\omega} \frac{d}{ds} [R_{s-t}^{j+1} - R_{\omega-t}^{j+1}] (\omega - s)^{\xi-1} ds = - [(R_{s-t}^{j+1} - R_{\omega-t}^{j+1}) (\omega - s)^{\xi-1}]_0^{\omega} \\ & \quad + (1 - \xi) \int_0^{\omega} (R_{s-t}^{j+1} - R_{\omega-t}^{j+1}) (\omega - s)^{\xi-2} ds = A + B. \end{aligned}$$

Nun ist

$$A = (R_{-t}^{j+1} - R_{\omega-t}^{j+1}) \omega^{\xi-1},$$

daher genügt es, wegen

$$\frac{\omega^{\xi-1}}{\omega^{\gamma}} \int_{-\infty}^{+\infty} a_t (R_{-t}^{j+1} - R_{\omega-t}^{j+1}) dt = o(1)$$

$$(\text{denn } R_s^{j+1} = O(s^{-2\gamma-3+j+1}) = O(s^{-\gamma-2-\xi}))$$

den folgenden Ausdruck zu untersuchen:

$$\begin{aligned} & \omega^{-\gamma} \int_{-\infty}^{+\infty} a_t dt \int_0^{\omega} (R_{s-t}^{j+1} - R_{\omega-t}^{j+1}) (\omega - s)^{\xi-2} ds = \\ & = \omega^{-\gamma} \int_0^{\omega} (\omega - s)^{\gamma-2} ds \int_{-\infty}^{+\infty} (R_{s-t}^{j+1} - R_{\omega-t}^{j+1}) a_t dt = \omega^{-\gamma} \int_0^{\omega-1} + \omega^{-\gamma} \int_{\omega-1}^{\omega} = C + D. \end{aligned}$$

Da aber (Lemma *c*)

$$\int_{-\infty}^{+\infty} R_{s-t}^{j+1} a_t dt = o(s^{\gamma}); \quad \int_{-\infty}^{+\infty} R_{\omega-t}^{j+1} a_t dt = o(\omega^{\gamma}),$$

so ist

$$C = \omega^{-\gamma} \int_0^{\omega-1} (\omega - s)^{\xi-2} o(s^{\gamma}) ds + o(1) \int_0^{\omega-1} (\omega - s)^{\xi-2} ds = o(1).$$

Wenden wir jetzt auf die Differenz $R_{s-t}^{j+1} - R_{\omega-t}^{j+1}$ den Mittelwertsatz an:

$$R_{s-t}^{j+1} - R_{\omega-t}^{j+1} = (\omega - s) R_{\theta-t}^j$$

(wobei $\omega - 1 < s < \theta < \omega$; $\theta = \theta(s-t, \omega-t)$)

und beachten wir, daß, wie leicht zu beweisen (vgl. z. B. Lemma c),

$$\int_{-\infty}^{+\infty} R_{\theta-t}^j a_t dt = o(\omega^\gamma) \quad (\text{für } \omega - 1 < \theta < \omega),$$

so erhalten wir

$$D = \omega^{-\gamma} \int_{\omega-1}^{\omega} (\omega - s)^{\xi-1} ds \int_{-\infty}^{+\infty} a_t R_{\theta-t}^j dt = o(1) \int_{\omega-1}^{\omega} (\omega - s)^{\xi-1} ds = o(1),$$

w. z. b. w.

Analog wird der Satz von der konjugierten Reihe bewiesen.

Satz IV. *Genügen die a_s den Bedingungen des Satzes III und genügen die α_s außer (21) noch für $x < E$ den Bedingungen*

$$\frac{1}{2} \int_{-\infty}^{+\infty} \alpha_s e^{isx} ds = \alpha(x);$$

$$\frac{i}{2} \int_{-\infty}^{+\infty} s \alpha_s e^{isx} ds = \frac{i^2}{2} \int_{-\infty}^{+\infty} s^2 \alpha_s e^{isx} ds = \dots = \frac{i^k}{2} \int_{-\infty}^{+\infty} s^k \alpha_s e^{isx} ds = 0,$$

so sind die Ausdrücke (15) gleichmäßig in $x < E$ (C, γ)-limitierbar, und zwar der erste gegen Null.

Analog wie im Satze II erhalten wir (16) und es ist zu zeigen, daß der zweite Summand (C, γ) gegen Null strebt. Der Beweis dieser Tatsache ist im Grunde mit dem des Satzes III identisch (vgl. Lemma e).

Wir betrachten jetzt die Frage nach der Summabilität des m -mal differenzierten formalen Produktes der t. I. (9) und (10), d. h. des t. I.

$$\frac{i^m}{2} \int_{-\infty}^{+\infty} A_s \cdot s^m e^{isx} ds.$$

Wir beschränken uns auf den Fall $\gamma + m \geq 0$. Wir erhalten, wenn wir uns, wie oben, auf $x = 0$ beschränken,

$$\begin{aligned} S_\omega &= \frac{i^m}{2} \int_{-\infty}^{+\infty} A_s \cdot s^m ds = \frac{i^m}{4} \int_{-\infty}^{+\infty} s^m \left(\int_{-\infty}^{+\infty} a_t \alpha_{s-t} dt \right) ds \\ &= \frac{i^m}{4} \int_{-\infty}^{+\infty} a_t dt \int_{-\omega-t}^{\omega-t} \alpha_s (s+t)^m ds = \sum_{j=0}^m \binom{m}{j} \frac{i^m}{4} \int_{-\infty}^{+\infty} t^j a_t dt \cdot \int_{-\omega-t}^{\omega-t} s^{m-j} \alpha_s ds. \end{aligned}$$

Der Summand mit dem Index j rechts ist, bis auf einen konstanten Faktor, das Teilintegral (für $x = 0$) des f. P. der Integrale

$$1^\circ \frac{1}{2} \int_{-\infty}^{+\infty} t^j a_t \cdot i^j dt \text{ mit Gliedern } o(t^{\gamma+j}) \text{ und}$$

$$2^\circ \frac{1}{2} \int_{-\infty}^{+\infty} s^{m-j} i^{m-j} \alpha_s ds = \alpha^{(m-j)}(0)$$

(Leibnizsche Formel). Wir setzen voraus, daß

$$(24) \quad \alpha_s = O(s^{-\eta}), \quad \eta = \text{Max}(2(\gamma + m) + 3, m + 3),$$

k ganz, nicht negativ $\geq \gamma + m$.

Was kann man nun über die Summabilität des f. P. der Integrale

$$(25) \quad \frac{i^j}{2} \int_{-\infty}^{+\infty} t^j a_t e^{itx} dt \quad \text{und} \quad \frac{i^{m-j}}{2} \int_{-\infty}^{+\infty} s^{m-j} \alpha_s e^{isx} ds \quad (j = 0, 1, \dots, m)$$

aussagen? Man muß zwei Fälle unterscheiden: a) $\gamma + j \leq 0$, b) $\gamma + j > 0$. Im ersten Falle, wenn nur $\alpha_s \cdot s^{m-j} = O(s^{-3})$ (vgl. Satz II), also um so mehr, wenn $\alpha_s = O(s^{-\eta})$, ist das f. P. der t. I. (25) gegen Null konvergent, wenn nur $m - j > 0$ und verhält sich wie

$$(25a) \quad \alpha \frac{i^m}{2} \int_{-\infty}^{+\infty} t^m a_t e^{itx} dt$$

für $m - j = 0$. Im zweiten Fall dagegen, wenn nur $s^{m-j} \alpha_s = O(s^{-2(\gamma+j)-3})$ (vgl. den Satz III), also um so mehr, wenn $\alpha_s = O(s^{-\eta})$, ist dieses Produkt $(C, \gamma + j)$ gegen Null summierbar, wenn $m - j > 0$ und verhält sich $(C, \gamma + j)$ so wie (25a), wenn $j = m$. Fassen wir diese Betrachtungen zusammen und vervollständigen sie durch analoge Sätze über die konjugierten Integrale, so erhalten wir den folgenden

Satz V. Ist $\int_{\mu}^{\mu+1} |\alpha_s| ds = o(\mu^\gamma)$ (γ beliebig reell) und sind die k ersten Ableitungen der Funktion

$$\alpha(x) = \frac{1}{2} \int_{-\infty}^{+\infty} \alpha_s e^{isx} ds,$$

wobei (24) erfüllt ist¹⁰⁾, für $x \in E$ gleich Null, so sind die Ausdrücke

¹⁰⁾ Z. B. für $\gamma = -2$, $m = 2$: $k = 0$. Dieser Fall ist in den Anwendungen der wichtigste.

$$(26) \quad \frac{i^m}{2} \int_{-\infty}^{+\infty} A_s e^{isx} s^m ds - \frac{\alpha(x) i^m}{2} \int_{-\infty}^{+\infty} a_s e^{isx} s^m ds,$$

$$\frac{i^m}{2} \int_{-\infty}^{+\infty} B_s e^{isx} s^m ds - \frac{\alpha(x) i^m}{2} \int_{-\infty}^{+\infty} b_s e^{isx} s^m ds$$

gleichmäßig in $x \in E$ ($C, \gamma + m$) ($\gamma + m \geq 0$) limitierbar, und zwar das erste gegen Null¹¹⁾.

Aus dem letzten Satze werden wir leicht die, den Riemannschen analogen Lokalisationssätze erhalten. Betrachten wir zuerst den für die Anwendungen vollkommen ausreichenden Fall des Integrals

$$(27) \quad \frac{1}{2} \int_{-\infty}^{+\infty} a_s e^{isx} ds,$$

wo ' bei dem Integral hier und fortan bedeutet, daß für $|s| < 1$ $a_s = 0$ ist¹²⁾. Diese Beschränkung bezweckt die Erleichterung der formalen Integration.

Satz VI. Ist $\int_{\mu}^{\mu+1} |a_s| ds = o(\mu^\gamma)$ ($\gamma \geq 0$) und bezeichnen wir:

$$(28) \quad \Phi(x) = \frac{1}{2i^k} \int_{-\infty}^{+\infty} \frac{a_s}{s^k} e^{isx} ds,$$

wo k eine natürliche Zahl und $k - \gamma > 1$ ¹³⁾, so sind, wenn wir mit $\lambda(x)$ ($-\infty < x < \infty$) eine stetige, mit einer gewissen Zahl von Ableitungen versehene¹⁴⁾ Funktion bezeichnen, die außerhalb (a, b) Null und innerhalb (α, β) ($-\infty < a < \alpha < \beta < b < \infty$) Eins ist, die Ausdrücke

$$(29) \quad \frac{1}{2} \int_{-\infty}^{+\infty} a_s e^{isx} ds - \frac{(-1)^k}{\pi} \int_a^b \Phi(t) \lambda(t) \frac{d^k}{dt^k} \frac{\sin \omega(t-x)}{t-x} dt$$

$$\frac{1}{2} \int_{-\infty}^{+\infty} b_s e^{isx} ds - \frac{(-1)^k}{\pi} \int_a^b \Phi(t) \lambda(t) \frac{d^k}{dt^k} \frac{\sin \omega(t-x)}{t-x} dt$$

gleichmäßig in (α, β) (C, γ)-limitierbar, und zwar der erste gegen Null.

¹¹⁾ Für $\gamma + m < 0$ sind die Ausdrücke (26) offenbar konvergent.

¹²⁾ Wir bemerken noch, daß, wenn $\int_{-1}^{+1} |a_s| ds < \infty$, die Funktion $\int_{-1}^{+1} a_s e^{isx} ds$ Ableitungen aller Ordnung besitzt. Diese einfache Bemerkung werden wir noch benutzen.

¹³⁾ Damit das in (28) rechtsstehende Integral absolut konvergiert. Es genügt schwächere Voraussetzungen zu machen.

¹⁴⁾ Es genügt, wie aus dem Beweise folgt, daß die Koeffizienten des Fourierintegrals der Funktion $\lambda(x)$ von der Größenordnung $o(s^{-\eta})$ sind, wo $\eta = \text{Max}(2\gamma + 3, k + 3)$. Z. B. für $\gamma = 0$ kann man $k = 2$, also $\eta = 5$ setzen.

Mit $\overline{\Phi(t)\lambda(t)}$ bezeichnen wir hier die mit $\Phi(t)\lambda(t)$ konjugierte Funktion, d. h. den Wert des mit dem Fourierintegral der Funktion $\Phi(t)\lambda(t)$ konjugierten Integrales. Bei unseren Voraussetzungen über Φ und λ ist die Funktion $\overline{\Phi\lambda}$ stetig (vgl. den Beweis des Satzes VI).

Dem Beweise des Satzes VI schicken wir einige Bemerkungen über seine Bedeutung voraus. Er sagt aus, daß die Summabilität des t. I. (27) in (α, β) lediglich von dem Verhalten der Funktion $\Phi(x)$ in dem etwas größeren Intervall (a, b) abhängt. Haben wir das allgemeine Integral (1):

$$\frac{1}{2} \int_{-\infty}^{+\infty} a_s e^{isx} ds = \frac{1}{2} \int_{-1}^{+1} a_s e^{isx} ds + \frac{1}{2} \int_{-\infty}^{+\infty} a_s e^{isx} ds,$$

und definieren wir nach wie vor $\Phi(x)$ durch die Formel (28), so können wir auch jetzt behaupten, daß die (C, γ) -Summabilität (aber nicht mehr der Wert) des t. I. (1) in (α, β) nur von dem Verhalten der Funktion $\Phi(x)$ in (a, b) abhängt.

Beweis des Satzes VI. Es sei

$$(30) \quad \lambda(x) = \frac{1}{2} \int_{-\infty}^{+\infty} a_s e^{isx} ds.$$

Wir betrachten das f. P. (11) der Integrale (28) und (30). Auf Grund des Lemmas *b* ist $A_s = o(sr^{-k})$. Es ist leicht zu sehen, daß das f. P. das Fourierintegral der Funktion $\Phi\lambda$ ist, also

$$A_s = \frac{1}{\pi} \int_{-\infty}^{+\infty} F(t) \lambda(t) e^{-its} dt;$$

d. h.

$$\left(\frac{1}{2} \int_{-\infty}^{+\infty} A_s e^{isx} ds \right)^{(k)} = \frac{(-1)^k}{\pi} \int_a^b F(t) \lambda(t) \frac{d^k}{dt^k} \frac{\sin \omega(t-x)}{t-x} dt.$$

Wenden wir nun den Satz V an, so erhalten wir den ersten Teil des Satzes VI. Den zweiten Teil, der die konjugierten Integrale betrifft, beweisen wir analog.

Hätten wir die Beschränkung auf die Integrale (27) vermeiden wollen, so könnte man statt $\Phi(x)$ die Funktion

$$F(x) = \frac{1}{2i^k} \int_{-\infty}^{+\infty} a_s e^{isx} ds + \frac{1}{2i^k} \int_{-1}^{+1} a_s [e^{isx} - P_k(isx)] ds = \Phi(x) + \Psi(x)$$

betrachten¹⁵⁾, wo $P_k(u) = 1 + \frac{u}{1!} + \dots + \frac{u^{k-1}}{(k-1)!}$.

¹⁵⁾ H. Hahn [1], M. Jacob [1].

Satz VI'. Erfüllen a_s und $\lambda(x)$ die Voraussetzungen des vorigen Satzes, so sind die Ausdrücke

$$(31) \quad \frac{1}{2} \int_{-\omega}^{+\omega} a_s e^{isx} ds - \frac{(-1)^k}{\pi} \int_a^b F(t) \lambda(t) \frac{d^k}{dt^k} \frac{\sin \omega(t-x)}{t-x} dt$$

$$\frac{1}{2} \int_{-\omega}^{+\omega} b_s e^{isx} ds - \frac{(-1)^k}{\pi} \int_a^b \overline{F(t) \lambda(t)} \frac{d^k}{dt^k} \frac{\sin \omega(t-x)}{t-x} dt$$

in (α, β) gleichmäßig (C, γ) -limitierbar, und zwar der erste gegen Null.

Zum Beweise genügt es offenbar zu zeigen, daß die Ausdrücke

$$\frac{1}{2} \int_{-1}^{+1} a_s e^{isx} - \frac{(-1)^k}{\pi} \int_a^b \Psi(t) \lambda(t) \frac{d^k}{dt^k} \frac{\sin \omega(t-x)}{t-x} dt,$$

$$\frac{1}{2} \int_{-1}^{+1} b_s e^{isx} - \frac{(-1)^k}{\pi} \int_a^b \overline{\Psi(t) \lambda(t)} \frac{d^k}{dt^k} \frac{\sin \omega(t-x)}{t-x} dt$$

in (α, β) gleichmäßig konvergieren (der erste gegen Null). Wir betrachten nun den ersten, der nach einer k -maligen partiellen Integration die Gestalt annimmt

$$\frac{1}{2} \int_{-1}^{+1} a_s e^{isx} ds - \frac{1}{\pi} \int_a^b [\Psi(t) \lambda(t)]^{(k)} \frac{\sin \omega(t-x)}{t-x} dt;$$

aber für $j = 0, 1, 2, \dots, k-1$ gilt gleichmäßig in (α, β) :

$$\binom{k}{j} \frac{1}{\pi} \int_a^b \Psi^{(j)}(t) \lambda^{(k-j)}(t) \frac{\sin \omega(t-x)}{t-x} dt \rightarrow 0$$

$$\frac{1}{\pi} \int_a^b \Psi^{(k)}(t) \lambda(t) \frac{\sin \omega(t-x)}{t-x} dt \rightarrow \Psi^{(k)}(x) = \frac{1}{2} \int_{-1}^{+1} a_s e^{isx} ds,$$

w. z. b. w.¹⁶⁾

¹⁶⁾ Aus den Formeln (29) oder (31) folgt, vermittels einer zu der in Satz VI' angestellten analogen Betrachtung: wenn die Funktion

$$\int_0^{\infty} a_s e^{s^2 x} ds, \quad \left(\int_{\mu}^{\mu+1} |a_s| ds = o(\mu^\gamma), \quad \gamma \geq 0 \right),$$

die für $\Re z > 0$ holomorph ist, noch in einem geschlossenen und endlichen Intervall der Geraden $\Re z = 0$ sich regulär verhält, so ist das betrachtete Integral daselbst gleichmäßig (C, γ) summierbar.

Kapitel 2.

§ 4.

Bevor ich an den Beweis der im § 1 angesagten Sätze herangehe, erinnere ich an einige bekannte daran anknüpfende Definitionen. Die Menge $E \subset (0, 2\pi)$ nennen wir eine *Eindeutigkeitsmenge* oder Menge U für t. R., wenn für jede t. R.

$$(32) \quad \frac{1}{2} \sum_{n=-\infty}^{+\infty} a_n e^{in\theta} \quad a_{-n} = \bar{a}_n,$$

die überall in $C E$ (d. h. in der Komplementärmenge von E) gegen Null konvergiert, die Bedingung $a_n = 0$ ($n = 0, \pm 1, \pm 2, \dots$) zutrifft. Sonst nennen wir die Menge E eine *Mehrdeutigkeitsmenge* (M -Menge). Analog werden definiert die entsprechenden Mengen für die t. I.

$$(33) \quad \frac{1}{2} \int_{-\infty}^{+\infty} a_s e^{isx} ds, \quad a_{-s} = \bar{a}_s,$$

mit dem Unterschied, daß wir hier die ergänzende Bedingung

$$(34) \quad \lim_{\mu \rightarrow +\infty} \int_{\mu}^{\mu+1} |a_s| ds = 0$$

hinzufügen. Die Menge E liegt hier in $(-\infty, +\infty)$ und wir fordern $a_s = 0$ für fast alle s . Beschränken wir uns auf meßbare U -Mengen, so gilt der Satz, daß jede U -Menge — gleich ob für Reihen oder Integrale — vom Maße Null ist. Die Umkehrung ist nicht richtig¹⁷⁾.

Diese Definitionen knüpfen an den Begriff der Konvergenz der Reihen bzw. der Integrale an, und daher könnten wir zwecks größerer Deutlichkeit die in Betracht kommenden Mengen U_C resp. M_C nennen. Betrachten wir dagegen statt der Konvergenz die verallgemeinerten Summierungsverfahren, z. B. das Poissonsche, so gelangen wir zu den Definitionen der Mengen U_P und M_P . Damit diese Definitionen zweckmäßig seien, muß man für den Fall der Reihen die Voraussetzung $a_n \rightarrow 0$ hinzufügen. Für die Integrale bleibt die Voraussetzung (34) bestehen. Man vermutet, was nur in dem Spezialfall der geschlossenen Mengen bewiesen ist¹⁸⁾, daß überhaupt die Klassen der Mengen U_P und U_C identisch sind.

Satz I. *Konvergiert das t. I. (33) mit der Nebenbedingung (34) in einem endlichen Intervall (A, B) gegen Null, so ist die Konvergenz gleichmäßig in jedem Intervall $(A + \varepsilon, B - \varepsilon)$ ($\varepsilon > 0$).*

¹⁷⁾ D. Menchoff [1].

¹⁸⁾ Vgl. z. B. Nina Bary [1].

Beweis. Es sei $\lambda(x)$ ($-\infty < x < \infty$) eine stetige mit einer gewissen Zahl von Ableitungen versehene Funktion, die in $(A + \varepsilon, B - \varepsilon)$ gleich 1, außerhalb (A, B) gleich 0 ist. Wir entwickeln $\lambda(x)$ in ein Fourierintegral

$$(35) \quad \lambda(x) = \frac{1}{2} \int_{-\infty}^{+\infty} \alpha_s e^{isx} ds.$$

Das formale Produkt der Integrale (33) und (35) ist ein t. I.

$$(36) \quad \frac{1}{2} \int_{-\infty}^{+\infty} A_s e^{isx} ds, \quad (A_s \rightarrow 0),$$

das überall gegen Null konvergiert, also sind fast alle $A_s = 0$ ¹⁹⁾, die Konvergenz ist also gleichmäßig. Daraus folgt (Kap. 1, Satz II) die Behauptung.

Folgerung. Konvergiert das Integral (33) in dem Intervall (A, B) gegen eine in ein gleichmäßig konvergentes Fourierintegral entwickelbare Funktion $f(x)$ ²⁰⁾, so konvergiert das Integral (33) gleichmäßig in $(A + \varepsilon, B - \varepsilon)$. Es genügt nämlich für $x < (A, B)$ $f^*(x) = f(x)$ und sonst $f^*(x) = 0$ zu setzen und dann die Differenz des t. I. (33) und des Fourierintegrals von $f^*(x)$ zu betrachten.

Bemerkung. Stützen wir uns auf den Satz II (s. u.), so ist leicht zu sehen, daß die Voraussetzung der Poissonschen Summierbarkeit statt der Konvergenz genügt.

Satz II. Jede leere Menge ist für die t. I. ein U_P .

Bemerkung. Im Beweise werden wir uns auf den Satz stützen, daß für die t. R. die leere Menge ein U_P ist²¹⁾. Die im folgenden angewandte Methode gibt aber noch mehr; sie zeigt nämlich für eine umfangreiche Klasse von Summierungsverfahren T , daß wenn die leere Menge ein U_T für t. R. ist, das gleiche auch für die t. I. gilt.

Lemma α . Ist eine t. R. (32) mit nach Null strebenden Koeffizienten in (A, B) ($0 \leq A < B \leq 2\pi$) P -summierbar zu einer in eine Fourierreihe entwickelbaren Funktion $f(x)$, so ist diese Reihe (32) innerhalb (A, B) konvergent.

¹⁹⁾ S. Pollard [1], M. Jacob [1].

²⁰⁾ Darunter verstehen wir folgendes: wenn wir eine beliebige, in $(-\infty, +\infty)$ absolut integrierbare Funktion $f^*(x)$ nehmen, wobei $f(x) = f^*(x)$ in (A, B) , so ist das Fourierintegral von $f^*(x)$ gleichmäßig in $(A + \varepsilon, B - \varepsilon)$ ($\varepsilon > 0$) konvergent. Analog für t. R.

²¹⁾ A. Rajchman [2], A. Rajchman und A. Zygmund [1], A. Zygmund [2].

Der Beweis ist dem des Satzes I analog. Zuerst betrachten wir den Fall, wo $f(x) = 0$ in (A, B) ist. Multiplizieren wir die Reihe (32) mit der Fourierreihe einer Lokalisationsfunktion $\lambda(x)$ von der Periode 2π , die in $(A + \varepsilon, B - \varepsilon)$ gleich 1, außerhalb $(A, B) \pmod{2\pi}$ gleich 0 ist und eine gewisse Anzahl stetiger Ableitungen besitzt, so erhalten wir eine t. R. mit nach Null strebenden Koeffizienten, die überall zur Null P -summierbar ist. Daher ist das f. P. auch konvergent, also ist die Reihe (32) in $(A + \varepsilon, B - \varepsilon)$ konvergent (sogar gleichmäßig). Im allgemeineren Fall genügt es, von der Reihe (32) die Fourierreihe der Funktion $f^*(x)$, die in (A, B) gleich $f(x)$, sonst Null ist, zu subtrahieren.

Lemma β . Ist das t. I.

$$(37) \quad \frac{1}{2} \int_{-\infty}^{+\infty} a_s e^{isx} ds, \quad \left(a_{-s} = \bar{a}_s, \quad \int_{\mu}^{\mu+1} |a_s| ds \rightarrow 0 \right)^{22)}$$

P -summierbar zu einer Funktion, die z. B. eine stetige Ableitung besitzt, so ist es sogar konvergent (gleichmäßig in jedem endlichen Intervall).

Es sei $\lambda(x)$ eine in $(-\infty, +\infty)$ definierte Funktion, gleich 1 in $(\delta, 2\pi - \delta)$, Null außerhalb $(0, 2\pi)$ und samt einer gewissen Anzahl ihrer Ableitungen stetig. Es sei (35) das Fourierintegral von $\lambda(x)$. Das f. P. (36) der t. I. (35) und

$$-\frac{1}{2} \int_{-\infty}^{+\infty} \frac{a_s}{s^2} e^{isx} ds = F(x)$$

ist das Fourierintegral der Funktion $F(x)\lambda(x)$ und dabei ist (vgl. Lemma b) $A_s = o(s^{-2})$. Wir entwickeln die Funktion $F(x)\lambda(x)$ in $(0, 2\pi)$ in eine Fourierreihe

$$(38) \quad F(x)\lambda(x) \sim \frac{1}{2} \sum_{n=-\infty}^{+\infty} b_n e^{inx},$$

wo

$$b_n = \frac{1}{\pi} \int_0^{2\pi} F(x)\lambda(x) e^{-inx} dx = \frac{1}{\pi} \int_{-\infty}^{+\infty} F(x)\lambda(x) e^{-inx} dx = A_n = o(n^{-2}).$$

Wir differenzieren die Reihe (38) formal. Wir erhalten dann die Reihe $-\frac{1}{2} \sum n^2 b_n e^{inx}$, deren Koeffizienten gegen Null streben, wobei

$$-\frac{1}{2} \sum_{n=-N}^{+N} n^2 b_n e^{inx} - \frac{1}{2\pi} \int_0^{2\pi} F(t)\lambda(t) \frac{d^2}{dt^2} \frac{\sin\left(N + \frac{1}{2}\right)(x-t)}{\sin \frac{x-t}{2}} dt = 0.$$

²²⁾ ' bedeutet, daß $a_s = 0$ für $|s| < 1$.

Andererseits ist nach der Lokalisationsformel²³⁾

$$\frac{1}{2} \int_{-N^{-1/2}}^{N+1/2} a_s e^{isx} ds - \frac{1}{\pi} \int_0^{2\pi} F(t) \lambda(t) \frac{d^2}{dt^2} \frac{\sin\left(N + \frac{1}{2}\right)(x-t)}{x-t} dt \rightarrow 0$$

$$(x \in (\delta, 2\pi - \delta)).$$

Ich behaupte, daß

$$(39) \quad \frac{1}{\pi} \int_0^{2\pi} F(t) \lambda(t) \frac{d^2}{dt^2} \left[\left(\frac{1}{2 \sin \frac{x-t}{2}} - \frac{1}{x-t} \right) \sin\left(N + \frac{1}{2}\right)(x-t) \right] dt \rightarrow 0$$

$$(x \in (\delta, 2\pi - \delta)).$$

In der Tat, setzen wir

$$\frac{1}{2 \sin \frac{u}{2}} - \frac{1}{u} = \psi(u),$$

so können wir die linke Seite der Formel (39) in der Gestalt

$$(40) \quad - \frac{\left(N + \frac{1}{2}\right)^2}{\pi} \int_0^{2\pi} F(t) \lambda(t) \psi(x-t) \sin\left(N + \frac{1}{2}\right)(x-t) dt -$$

$$- \frac{2N+1}{\pi} \int_0^{2\pi} F(t) \lambda(t) \psi'(x-t) \cos\left(N + \frac{1}{2}\right)(x-t) dt +$$

$$+ \frac{1}{\pi} \int_0^{2\pi} F(t) \lambda(t) \psi''(x-t) \sin\left(N + \frac{1}{2}\right)(x-t) dt$$

umschreiben. Da jedoch für jedes x die Funktion

$$\lambda^*(t) = \lambda(t) \psi^{(i)}(x-t) \begin{matrix} \sin \frac{t}{2} \\ \cos \frac{t}{2} \end{matrix} \quad (i = 0, 1, 2)$$

sich genügend regulär verhält, so behält sie die Ordnung der Koeffizienten bei und die Funktion $F(t) \lambda^*(t)$ hat Fourierkoeffizienten $o(N^{-2})$, also strebt jeder der Summanden in dem Ausdruck (40) gegen Null. Wir haben gezeigt, daß

$$\left(-\frac{1}{2} \sum_{n=-N}^{+N} n^2 b_n e^{inx} \right) - \frac{1}{2} \int_{-N^{-1/2}}^{+N+1/2} a_s e^{isx} ds \rightarrow 0,$$

²³⁾ Vgl. Kap. I, Satz VI.

also, da $\int_{\mu}^{\mu+1} |a_s| ds \rightarrow 0$, wird auch

$$(41) \quad \Delta_n = \left(-\frac{1}{2} \sum_{n=-N}^{+N} n^2 b_n e^{inx} \right) - \frac{1}{2} \int_{-N}^{+N} a_s e^{isx} ds \rightarrow 0 \quad (x \in (\delta, 2\pi - \delta)).$$

Da Δ_n eine Nullfolge ist, so ist sie auch zur Null nach Poisson limitierbar, also

$$(42) \quad \lim_{r \rightarrow 1} \left[\left(-\frac{1}{2} \sum_{n=-\infty}^{+\infty} n^2 r^{|n|} b_n e^{inx} \right) - \frac{1}{2} \sum_{n=0}^{\infty} \left(\int_n^{n+1} a_s e^{isx} ds + \int_{-n-1}^{-n} a_s e^{isx} ds \right) r^{n+1} \right] = 0.$$

Da aber, wie leicht zu zeigen, bei der Voraussetzung $\int_{\mu}^{\mu+1} |a_s| ds \rightarrow 0$, auch

$$\frac{1}{2} \sum_{n=0}^{\infty} \left(\int_n^{n+1} a_s e^{isx} ds + \int_{-n-1}^{-n} a_s e^{isx} ds \right) r^{n+1} - \frac{1}{2} \int_{-\infty}^{+\infty} a_s e^{isx} r^{|s|} ds \rightarrow 0 \quad (\text{für } r \rightarrow 1)$$

gilt, so folgt aus (42), daß

$$(43) \quad \lim_{r \rightarrow 1} \left[\left(-\frac{1}{2} \sum_{n=-\infty}^{+\infty} n^2 b_n r^{|n|} e^{inx} \right) - \frac{1}{2} \int_{-\infty}^{+\infty} a_s e^{isx} r^{|s|} ds \right] = 0, \\ (x \in (\delta, 2\pi - \delta)).$$

Die Formeln (41) und (43) werden wir oft benutzen^{23a)}. Ist das Integral (37) in (A, B) P -summierbar gegen $f(x)$, so ist es wegen der Formel (43) auch die Reihe $-\frac{1}{2} \sum n^2 b_n e^{inx}$. Nach dem Lemma α ist diese Reihe also gegen $f(x)$ konvergent, daher ist es auch, nach der Formel (41), das Integral (37) in $(\delta, 2\pi - \delta)$. Da aber statt $(\delta, 2\pi - \delta)$ ein beliebiges Intervall von der Länge $< 2\pi$ genommen werden kann, so ist das Lemma bewiesen. Nun ist der Beweis des Satzes II unmittelbar. Ist das Integral (33) mit der Nebenbedingung (34) zur Null P -summierbar, so setzen wir $a_s = a'_s + a''_s$, wo $a'_s = 0$ für $|s| < 1$ und $a'_s = a_s$ für $|s| \geq 1$,

^{23a)} Es sei $S(x)$ eine t. R., deren Glieder gegen Null streben, und $J(x)$ ein t. I. (33), wobei auch $a_s \rightarrow 0$. $\bar{S}(x)$ entstehe aus $S(x)$ durch zweimalige gliedweise Integration. Analog sei $\bar{I}(x)$ definiert. Die beim Beweise des Lemmas β benutzten Überlegungen ergeben: wenn für $a \leq x \leq b$

$$\bar{I}(x) = \bar{S}(x)$$

ist, so sind in (a, b) die Reihe $S(x)$ und das Integral $I(x)$ äquikonvergent, d. h. es ist in leichtverständlicher Bezeichnungsweise: $I(x) - S(x) = 0$ (sogar gleichmäßig in $(a + \varepsilon, b - \varepsilon)$).

und erhalten

$$\frac{1}{2} \int_{-\infty}^{+\infty} a_s e^{isx} ds = \frac{1}{2} \int_{-\infty}^{+\infty} a'_s e^{isx} ds + \frac{1}{2} \int_{-\infty}^{+\infty} a''_s e^{isx} ds.$$

Also ist das erste Integral links P -summierbar zum Werte

$$(44) \quad f(x) = -\frac{1}{2} \int_{-1}^{+1} a_s e^{isx} ds,$$

wobei die Funktion $f(x)$ stetige Ableitung besitzt. Nach dem Lemma β ist es also gegen $f(x)$ konvergent. Also konvergiert das t. I. (33) gegen Null, d. h. $a_s = 0$ für fast alle s , w. z. b. w.

Satz III. Jede U_C -Menge für t. R. ist es auch für t. I. und umgekehrt.

Satz III'. Jede U_P -Menge für t. R. ist es auch für t. I. und umgekehrt.

Die Sätze III und III' werden in derselben Weise bewiesen. Wir beweisen also nur den Satz III. Die Aussage, „die Menge E , die ein U_C für Integrale ist, ist ein U_C für Reihen“, ist in dem Sinne zu deuten, daß jedes Stück der Menge E , das in einem Intervall von der Länge $< 2\pi$ enthalten ist, ein U_C für Reihen bildet. Das folgende Lemma klärt genauer den Sinn der Sätze III und III'.

Lemma γ . Liegt E in $(-\infty, +\infty)$ und ist jedes endliche Stück von E ein U_C für die t. I., so ist auch die ganze Menge E ein U_C für Integrale. Dasselbe gilt für U_P .

In der Tat, würde es ein t. I. (33) mit der Bedingung (34) geben, das außerhalb E , aber nicht überall, gegen Null konvergiert, so würde es einen Punkt x_0 geben, in welchem das betrachtete Integral nicht gegen Null konvergiert. Nehmen wir jetzt eine Lokalisationsfunktion $\lambda(x)$, positiv in $(x_0 - \delta, x_0 + \delta)$ und Null außerhalb, multiplizieren wir das Integral (33) in das Fourierintegral der Funktion $\lambda(x)$, so erhalten wir jetzt ein t. I. (36), das außerhalb des Stückes $(x_0 - \delta, x_0 + \delta)$ der Menge E , aber nicht überall gegen Null konvergiert. Das ist aber nach Voraussetzung unmöglich.

Bemerkung. Ein analoges Lemma gilt — ähnlich bewiesen — für t. R.

Lemma δ . Ist eine t. R. außerhalb einer U_C -Menge gegen eine Funktion $f(x)$ — mit überall konvergenter Fourierentwicklung — konvergent, so ist diese Reihe die Fourierreihe von $f(x)$.

Dies ist evident, denn subtrahiert man von der betrachteten Reihe die Fourierreihe von $f(x)$, so erhält man eine überall außerhalb einer U_C -Menge gegen Null konvergente Reihe.

Beweis des Satzes III. 1. Es sei E ein U_C für t. R. Wir zeigen, daß es auch ein U_C für t. I. ist. Ohne die Allgemeinheit zu beschränken, kann man voraussetzen, daß E auf der Strecke von der Länge $< 2\pi$, also z. B. auf $(2\delta, 2\pi - 2\delta)$ liegt. (Vgl. das Lemma γ und Bemerkung.) Es sei (33) ein t. I. mit der Eigenschaft (34), das überall außerhalb E gegen Null konvergiert. Also konvergiert das t. I. (37) gegen die Funktion (44) in $C E$. Gehen wir von dem Integral (36) aus und stellen wir eine der im Lemma β angewandten ähnliche Betrachtung an, so erhalten wir die Reihe

$$(45) \quad -\frac{1}{2} \sum_{n=-\infty}^{+\infty} n^2 b_n e^{in x},$$

wobei noch die Relation (41) stattfindet, also wird die Reihe (45) überall in $(\delta, 2\pi - \delta)$ außerhalb E gegen $f(x)$ konvergieren. Es sei $\lambda_1(x)$ eine Funktion mit der Periode 2π , die samt einer gewissen Zahl von Ableitungen stetig ist, gleich 1 in $(2\delta, 2\pi - 2\delta)$ und Null außerhalb $(\delta, 2\pi - \delta) \pmod{2\pi}$. Wir multiplizieren die Reihe (45) mit der Fourierreihe der Funktion $\lambda_1(x)$. Wir erhalten eine t. R. mit nach Null strebenden Koeffizienten, die gegen $f(x)\lambda(x)$ überall in $C E$ konvergiert, also, nach dem Lemma δ , überall. Also konvergiert die Reihe (45), damit auch auf Grund der Formel (41) das Integral (37), überall in $(2\delta, 2\pi - 2\delta)$ gegen $f(x)$, also konvergiert (33) gegen Null, w. z. b. w.

2. Es sei E ein U_C für t. I., das in $(\delta, 2\pi - \delta)$ enthalten ist. Wir zeigen, daß es auch ein U_C für t. R. ist. Es sei

$$(46) \quad \frac{1}{2} \sum_{n=-\infty}^{\infty} b_n e^{in x}$$

eine t. R. die in $C E$ gegen Null konvergiert. Wir setzen

$$(47) \quad F(x) = \frac{b_0 x^2}{2} - \sum_{n=-\infty}^{+\infty} \frac{b_n}{n^2} e^{in x} = \frac{b_0 x^2}{2} + \Phi(x).$$

Es sei $\lambda(x)$ eine Funktion von der Periode 2π , die in $(\delta, 2\pi - \delta)$ Eins, in den Punkten $0, 2\pi$ Null ist und eine gewisse Anzahl von Ableitungen besitzt, die auch in den Endpunkten des Intervalles verschwinden.

Lemma ϵ . Es ist

$$\frac{1}{\pi} \int_0^{2\pi} F(t) \lambda(t) e^{-i\omega t} dt = o(\omega^{-2}).$$

Nun ist

$$\frac{1}{\pi} \int_0^{2\pi} F(t) \lambda(t) e^{-i\omega t} dt = \frac{b_0}{2\pi} \int_0^{2\pi} t^2 \lambda(t) e^{-i\omega t} dt + \frac{1}{\pi} \int_0^{2\pi} \Phi(t) \lambda(t) e^{-i\omega t} dt.$$

Offenbar ist der erste Summand rechts $o(\omega^{-2})$. Ist ω ganzzahlig: $\omega = N$, so ist das auch für den zweiten Summanden evident, da das betrachtete Integral den N -ten Fourierkoeffizienten der Funktion $\Phi(t) \lambda(t)$ darstellt, wobei die Fourierreihe der Funktion $\Phi(t) \lambda(t)$ durch formale Ausmultiplizierung der Fourierreihen der Funktionen Φ und λ entsteht. Verhält sich also $\lambda(x)$ genügend regulär, so sind die Fourierkoeffizienten des f. P., auf Grund eines dem Lemma b analogen Lemmas, von derselben Ordnung wie diejenigen von $\Phi(x)$, d. h. $o(N^{-2})$. Um also das Lemma allgemein zu beweisen, genügt es zu zeigen, daß für beschränktes ν ($|\nu| \leq \nu_0$) gleichmäßig

$$\int_0^{2\pi} \Phi(t) \lambda(t) [e^{-iNt} - e^{-i(N+\nu)t}] dt = \int_0^{2\pi} \Phi(t) \lambda(t) [1 - e^{-i\nu t}] e^{-iNt} dt = o(N^{-2})$$

wird. Es sei

$$\varphi(t) = 1 - \cos t \quad \text{oder} \quad \varphi(t) = \sin t.$$

Ich setze $\lambda^*(t) = \lambda(t) \varphi(\nu t)$. Wie leicht ersichtlich, wird für $|\nu| \leq \nu_0$ die Funktion $\lambda^*(t)$ gleichmäßig kleine (z. B. gleichmäßig $o(N^{-2})$) Fourierkoeffizienten haben. Da die Fourierreihe der Funktion $\Phi \lambda^*$ durch formale Multiplikation der Reihen Φ und λ^* entsteht, erhält man daher

$$\int_0^{2\pi} \Phi(t) \lambda^*(t) e^{-iNt} dt = o(N^{-2}),$$

gleichmäßig für: $|\nu| \leq \nu_0$, w. z. b. w.

Auf Grund der Riemannschen Formel gilt:

$$\frac{1}{2} \sum_{n=-N}^{+N} b_n e^{inx} - \frac{1}{2\pi} \int_0^{2\pi} F(t) \lambda(t) \frac{d^2}{dt^2} \frac{\sin\left(N + \frac{1}{2}\right)(x-t)}{\sin \frac{x-t}{2}} dt \rightarrow 0$$

$$(\delta \leq x \leq 2\pi - \delta).$$

Daraus folgt leicht (vgl. die entsprechenden Betrachtungen beim Beweise des Lemmas β), daß

$$\frac{1}{2} \sum_{n=-N}^{+N} b_n e^{inx} - \frac{1}{\pi} \int_0^{2\pi} F(t) \lambda(t) \frac{d^2}{dt^2} \frac{\sin\left(N + \frac{1}{2}\right)(x-t)}{x-t} dt \rightarrow 0$$

$$(\delta \leq x \leq 2\pi - \delta).$$

Wir behaupten, daß allgemein für stetiges $\omega \rightarrow \infty$:

$$(48) \quad \frac{1}{2} \sum_{-\omega < n < \omega} b_n e^{in x} - \frac{1}{\pi} \int_0^{2\pi} F(t) \lambda(t) \frac{d^2}{dt^2} \frac{\sin \omega(x-t)}{x-t} dt \rightarrow 0$$

$$(\delta \leq x \leq 2\pi - \delta).$$

Dazu genügt es, wie leicht zu sehen, zu zeigen, daß für $N \rightarrow \infty$ und für beschränktes, nicht notwendig ganzzahliges ν ,

$$(49) \quad \int_0^{2\pi} F(t) \lambda(t) \frac{d^2}{dt^2} \frac{\sin(N+\nu)(x-t) - \sin N(x-t)}{x-t} dt \rightarrow 0.$$

Setzt man

$$\varphi(t) = \frac{1 - \cos t}{t} \quad \text{resp.} \quad \varphi(t) = \frac{\sin t}{t}$$

und beachtet man, daß z. B.

$$\frac{d^2}{dt^2} [\sin Nt \varphi(\nu(x-t))] = -N^2 \sin Nt \varphi(\nu(x-t))$$

$$- 2N\nu \cos Nt \varphi'(\nu(x-t)) + \nu^2 \sin Nt \varphi''(\nu(x-t)),$$

so ist der Beweis von (49) dem der Relation (39) ganz analog. Aus der Formel (48) erhält man, daß in den Punkten des Intervalles $(\delta, 2\pi - \delta)$, die CE angehören,

$$(50) \quad \frac{1}{\pi} \int_0^{2\pi} F(t) \lambda(t) \frac{d^2}{dt^2} \frac{\sin \omega(x-t)}{x-t} dt \rightarrow 0$$

gilt. Es sei (33) das zweimal formal differenzierte Fourierintegral der Funktion $F(x) \lambda(x)$. Nach dem Lemma ε ist $a_\varepsilon \rightarrow 0$. Die Formel (50) gibt, daß das Integral (33) in allen zu CE gehörenden Punkten des Intervalls $(\delta, 2\pi - \delta)$ gegen Null strebt. Wir behaupten, daß das in dem ganzen Intervall stattfindet. In der Tat, es sei $\lambda_1(x)$ eine Funktion, die gleich Null außerhalb $(\delta, 2\pi - \delta)$ und positiv innerhalb $(\delta, 2\pi - \delta)$ ist. Das f. P. des t. I. (33) in das Fourierintegral der Funktion $\lambda_1(x)$ ergibt ein t. I., das gegen Null in CE konvergiert, also, da E ein U_C ist, überall gegen Null konvergiert. Daraus folgt, daß das Integral (33) überall gegen Null in $(\delta, 2\pi - \delta)$ konvergiert. Stützen wir uns jetzt auf die Formel (41) und beachten wir, daß $E \subset (\delta, 2\pi - \delta)$, so erhalten wir, daß die Reihe (46) überall in $(0, 2\pi)$ gegen Null konvergiert, also $b_n = 0$ ($n = 0, \pm 1, \pm 2, \dots$), w. z. b. w.

§ 5.

Wir nennen eine in $(0, 2\pi)$ gelegene Menge E ein U'_C für t. R., wenn jede t. R. die überall in $C E$ gegen eine integrierbare Funktion $f(x)$ konvergiert, die Fourierreihe der Funktion $f(x)$ ist. Es ist wahrscheinlich, daß die Klassen der Mengen U_C und U'_C identisch sind, es ist aber nur für geschlossene Mengen bewiesen²⁴⁾. Wir nennen eine Menge $E \subset (0, 2\pi)$ ein U''_C , wenn jede t. R. die in $C E$ gegen eine positive (oder von unten beschränkte) Funktion $f(x)$ konvergiert, eine Fourierreihe ist. Wie Steinhäus gezeigt hat, ist die leere Menge ein U''_C ²⁵⁾. Des weiteren ist bekannt, daß alle abzählbaren und gewisse nicht abzählbare (z. B. vom Typus H) Mengen U''_C sind. Auch hier drängt sich die Vermutung auf, daß die Klasse der U''_C mit derjenigen U_C identisch ist. Analoge Definitionen geben wir für t. I. Eine Menge $E \subset (-\infty, +\infty)$ ist ein U'_C (bzw. ein U''_C), wenn jedes t. I. (33), wobei (34) in $C E$ gegen eine in jedem endlichen Intervalle integrierbare (bzw. von unten beschränkte) Funktion $f(x)$ konvergiert, ein Fourierintegral von $f(x)$ ist. Der Begriff des Fourierintegrals ist hier im verallgemeinerten Sinne benutzt, d. h. die Koeffizienten a_s sind durch die Formeln

$$a_s = (C, 1) \frac{1}{\pi} \int_{-\infty}^{+\infty} f(x) e^{-isx} dx = \lim_{\lambda \rightarrow \infty} \frac{1}{\pi \lambda} \int_0^\lambda d\omega \int_{-\infty}^{+\infty} f(x) e^{-isx} dx$$

gegeben²⁶⁾.

Analog kann man die Mengen U'_P und U''_P für t. R. und t. I. definieren.

Satz IV. Die leere Menge ist ein U'_P und U''_P für t. I.

Bemerkung. Daß die leere Menge ein U'_C ist, hat Pollard (loc. cit.) gezeigt; daß sie auch U''_C ist, folgt leicht aus den Pollardschen Betrachtungen.

Satz V. Jede Menge, die ein U'_C bzw. ein U'_P für t. R. ist, ist es auch für t. I. und umgekehrt.

Satz V'. Jede Menge, die ein U''_C bzw. ein U''_P für t. R. ist, ist es auch für t. I. und umgekehrt.

Bemerkung. Der hier auftretende Integralbegriff kann sowohl im Lebesgueschen wie auch im Denjoy-Perronschen Sinne gedeutet werden.

Satz IV erweist sich als ein Spezialfall der Sätze V und V', wenn man beachtet, daß die leere Menge ein U'_P und U''_P für t. R. ist. Der Be-

²⁴⁾ Vgl. z. B. A. Zygmund [14], p. 94.

²⁵⁾ H. Steinhäus [1].

²⁶⁾ S. Pollard [1].

weis des Satzes V' unterscheidet sich nicht von dem des Satzes V, deswegen beschränken wir uns auf den Beweis des letzteren. Wir werden nur die Mengen U'_C betrachten. Die Beweise für U'_P unterscheiden sich nur durch Anwendung der Formel (43) statt (41). Außerdem können wir uns auf die Betrachtung des Integrals (37) beschränken, wo $a_s = 0$ für $|s| < 1$, denn die t. I. (33) und (37) unterscheiden sich um

$$\frac{1}{2} \int_{-1}^{+1} a_s e^{isx} ds,$$

was gewiß ein Fourierintegral im Pollardschen Sinne ist. Da, wie leicht zu zeigen (der Beweis ist dem des Lemmas γ analog), die Menge E dann und nur dann ein U'_C für t. I. ist, wenn jedes endliche Stück dieselbe Eigenschaft besitzt (analog für t. R.), so können wir uns auf Mengen beschränken, die in einem kleineren Intervall als 2π gelegen sind.

1. Es sei eine in $(\delta, 2\pi - \delta)$ gelegene Menge E ein U'_C für t. R. Es ist zu zeigen, daß sie auch ein U_C für t. I. ist. Es sei

$$\frac{1}{2} \int_{-\infty}^{+\infty} a_s e^{isx} ds \quad \left(a_s = 0 \text{ für } |s| < 1; \int_{\mu}^{\mu+1} |a_s| ds \rightarrow 0 \right)$$

das t. I., das überall in $C E$ gegen eine in jedem endlichen Intervall integrierbare Funktion $f(x)$ konvergiert. Ich zeige: setzt man

$$(51) \quad F(x) = -\frac{1}{2} \int_{-\infty}^{+\infty} \frac{a_s}{s^2} e^{isx} ds,$$

so hat man

$$(52) \quad F(x) = \int_0^x dy \int_0^y f(t) dt + Ax + B \quad (-\infty < x < \infty; A, B \text{ konst.}).$$

Zu diesem Zwecke multiplizieren wir, analog dem Beweise des ersten Teiles des Satzes III, das t. I. (51) mit dem Fourierintegral einer Lokalisationsfunktion $\lambda(x)$, die in $(\delta, 2\pi - \delta)$ 1 und außerhalb $(0, 2\pi)$ 0 ist. Die so erhaltene Funktion entwickeln wir in eine Fourierreihe (46), die wir zweimal differenzieren. Stützen wir uns auf die Formel (41), so folgern wir, daß die t. R. $-\frac{1}{2} \sum n^2 b_n e^{inx}$ in dem Intervall $(\delta, 2\pi - \delta)$ außerhalb einer U'_C -Menge gegen eine integrierbare Funktion $f(x)$ konvergiert. Daraus folgt²⁷⁾, daß die Reihe (46) in $(\delta, 2\pi - \delta)$ gegen die

²⁷⁾ Wir stützen uns hier auf den folgenden leicht beweisbaren Hilfssatz: Konvergiert eine t. R. (46) in einem Intervall (a, b) überall außerhalb einer U'_C -Menge gegen eine endliche integrierbare Funktion $f(x)$, so ist die Summe $F(x)$ (vgl. (47))

(Fortsetzung der Fußnote ²⁷⁾ auf nächster Seite)

Funktion $\int_0^x dy \int_0^y f(t) dt + Ax + B$ konvergiert; also haben wir, da doch in $(\delta, 2\pi - \delta)$ $\lambda(x) = 1$, die Formel (52) für $x \in (\delta, 2\pi - \delta)$ bewiesen. Da die hier auf $(\delta, 2\pi - \delta)$ angewandte Betrachtung auf jedes Intervall von der Länge $< 2\pi$ sich anwenden läßt, so findet die Formel (52) für jedes solche Intervall statt, wobei, wie man zunächst vermuten kann, die Zahlen A und B vom Intervall abhängig sind. Beachtet man aber, daß $F'(x)$ stetig ist, so zeigt sich, daß A und B überall denselben Wert haben, und die Formel (52) ist damit bewiesen. Da aber das Integral (51) absolut konvergiert, so gilt für fast alle Werte von s ²⁸⁾:

$$(C, 1) \frac{1}{\pi} \int_{-\infty}^{+\infty} F(x) e^{-isx} ds = \lim_{\lambda \rightarrow +\infty} \frac{1}{\pi\lambda} \int_0^\lambda d\omega \int_{-\infty}^{+\infty} F(x) e^{-isx} dx = -\frac{a_s}{s^2} \quad (|s| \geq 1)$$

$$(C, 1) \frac{1}{\pi} \int_{-\infty}^{+\infty} F(x) e^{-isx} ds = 0 \quad (|s| < 1).$$

Aber

$$\int_{-\infty}^{+\infty} F(x) e^{-isx} dx = \left[i F(x) \frac{e^{-isx}}{s} + F'(x) \frac{e^{-isx}}{s^2} \right]_{-\infty}^{+\infty} - \frac{1}{s^2} \int_{-\infty}^{+\infty} f(x) e^{-isx} ds.$$

Es ist leicht zu verifizieren, daß $\lim_{\omega \rightarrow \pm\infty} F(\omega) = 0 = (C, 1) \lim_{\omega \rightarrow \pm\infty} F'(\omega) e^{-is\omega}$.

Also haben wir für fast alle s

$$(52a) \quad (C, 1) \frac{1}{\lambda} \int_{-\infty}^{+\infty} f(x) e^{-isx} dx = \begin{matrix} a_s & (|s| \geq 1) \\ 0 & (|s| < 1) \end{matrix} \quad \text{w. z. b. w.}^{29)}.$$

der zweimal integrierten Reihe

$$F(x) = \int_a^x dy \int_a^y f(t) dt + Ax + B; \quad x \in (a, b).$$

Beweis. Es sei $\frac{1}{\lambda} \sum b_n' e^{inx}$ (46') das f. P. der Reihe (46) in die Fourierreihe einer Lokalisationsfunktion $\lambda(x)$, die 1 in $(a + \varepsilon, b - \varepsilon)$ ($\varepsilon > 0$) und 0 außerhalb (a, b) (mod 2π) ist, und es sei $F_1(x)$ die Summe der zweimal integrierten Reihe (46'). Da (46') die Fourierreihe einer Funktion $f_1(x)$ ist, so hat offenbar $F_1(x)$ die Gestalt

$$F_1(x) = \int_a^x dy \int_a^y f_1(t) dt + Ax + B.$$

Nun genügt es zu bemerken, daß die Differenz der Reihen (46) und (46') gegen Null im Intervall $(a + \varepsilon, b - \varepsilon)$ konvergiert, also ist $F - F_1$ eine lineare Funktion in $(a + \varepsilon, b - \varepsilon)$ und $f(x) = f_1(x)$ in dem gleichen Intervalle.

²⁸⁾ S. Pollard [1].

²⁹⁾ Hätten wir noch die absolute Integrierbarkeit von $f(x)$ im Intervall $(-\infty, +\infty)$ vorausgesetzt, so könnten wir in den Formeln (52a) die Zeichen $(C, 1)$ unterdrücken.

2. Wir setzen nun voraus, daß $E \subset (2\delta, 2\pi - 2\delta)$ ein U'_G für t. I. ist. Es gilt zu zeigen, daß E ein U'_G für t. R. ist. Es sei

$$(53) \quad \frac{1}{2} \sum_{n=-\infty}^{+\infty} b_n e^{inx}$$

eine in $(0, 2\pi)$, außerhalb E , gegen $f(x)$ konvergierende Reihe, wobei $f(x)$ endlich und integrierbar ist. Bezeichnen wir mit $F(x)$ (vgl. (47)) die zweimal formal integrierte Reihe (53), und mit $\lambda(x)$ eine Lokalisationsfunktion von der Periode 2π , gleich 1 in $(2\delta, 2\pi - 2\delta)$ und 0 außerhalb $(\delta, 2\pi - \delta) \pmod{2\pi}$. Bezeichnen wir mit (33) das zweimal differenzierte Fourierintegral der Funktion, die in $(0, 2\pi)$ $F(x)\lambda(x)$, sonst Null ist, so finden wir, daß dieses Integral überall in $(2\delta, 2\pi - 2\delta)$ außerhalb einer U'_G -Menge gegen eine integrierbare Funktion $f(x)$ konvergiert. Daraus folgt, daß in $(2\delta, 2\pi - 2\delta)$, also auch in jedem Intervall von der Länge $< 2\pi$,

$$F(x) = \int_0^x dy \int_0^y f(t) dt + Ax + B.$$

Daraus ist sehr leicht

$$b_n = \frac{1}{\pi} \int_0^{2\pi} f(x) e^{-inx} dx \quad (n = 0, \pm 1, \pm 2, \dots)$$

zu erhalten, w. z. b. w.

Bemerkungen. 1. Da die abzählbaren und H -Mengen U''_G (wie auch U''_P) für t. R. sind, so folgt aus dem Satze V', daß sie es auch für t. I. sind. Es ist aber bewiesen, daß im Fall der t. R. für die genannten Mengen statt der Beschränktheit von unten, schon die Beschränktheit von unten durch eine integrierbare und endliche Funktion ausreicht³⁰⁾. Offenbar gilt dasselbe für t. I.

2. Bekanntlich gibt es eine Anzahl von Eindeutigkeitsätzen in der Theorie der t. R., in denen man statt Konvergenz oder Summierbarkeit gewisse Voraussetzungen über die Unbestimmtheitsgrenzen macht³¹⁾. Offenbar lassen sich vermittels der benutzten Methode die entsprechenden Sätze für t. I. beweisen.

3. Herr Pollard³²⁾ hat die bekannte Mellinsche Formel für die Umkehrung einiger Integrale auf die Eindeutigkeitsfrage für t. R. zurückgeführt. Die angegebenen Sätze erlauben einige seiner Voraussetzungen abzuschwächen.

³⁰⁾ Für abzählbare Mengen vgl. z. B. A. Zygmund [1].

³¹⁾ Vgl. z. B. de la Vallée Poussin [1], A. Rajchman [2], A. Zygmund [2].

³²⁾ S. Pollard [1].

Verzeichnis der zitierten Arbeiten.

1. N. Bary, [1], Sur l'unicité du développement trigonométrique, *Fund. Math.* 9 (1927).
2. H. Hahn, [1], Über die Methode der arithmetischen Mittel in der Theorie der verallgemeinerten Fourierschen Integrale. *Sitzb. d. Ak. d. Wiss. Wien* (1925).
3. Hardy and Riesz, [1], *The general theory of Dirichlet's series*, Cambridge (1915), pp. 1–79.
4. M. Jacob, [1], Über den Eindeutigkeitssatz in der Theorie der trigonometrischen Integrale. *Math. Annalen* 98 (1927).
5. D. Menchoff, [1], Sur l'unicité du développement trigonométrique, *C. R.* 163 (1916), pp. 433–436.
6. S. Pollard, [1], On the identification of coefficients in the trigonometrical integral, *Proc. Lond. Math. Soc.* (1926).
7. A. Rajchman, [1], Sur le principe de localisation de Riemann (polnisch), *C. R. de la Soc. Scient. de Varsovie* 11 (1918), pp. 115–152.
8. A. Rajchman, [2], Sur les séries trigonométriques sommables par le procédé de Poisson (polnisch), *Prace Mat.-Fiz.* 30 (1919).
9. A. Rajchman, [3], Sur l'unicité du développement trigonométrique, *Math. Annalen* 96 (1925).
10. A. Rajchman und A. Zygmund, [1], Sur la possibilité d'appliquer la méthode de Riemann aux séries trigonométriques sommables par le procédé de Poisson, *Math. Zeitschr.* 25 (1926).
11. B. Riemann, [1], Über die Darstellbarkeit einer Funktion durch eine trigonometrische Reihe, *Ges. Abhandlungen*.
12. H. Steinhaus, [1], Quelques propriétés des séries trigonométriques et des séries de Fourier (polnisch), *C. R. de l'Acad. de Cracovie* (1915).
13. Ch. de la Vallée Poussin [1], Sur l'unicité du développement trigonométrique, *C. R.* 156 (1912), pp. 952–956.
14. A. Zygmund, [1], Sur la théorie riemannienne des séries trigonométriques, *Math. Zeitschr.* 23 (1925).
15. A. Zygmund, [2], Sur les séries trigonométriques sommables par le procédé de Poisson, *Math. Zeitschr.* 25 (1926).

(Eingegangen am 25. 7. 1927.)

Die Funktionaldeterminante als Deformationsmaß einer Abbildung und als Kriterium der Abhängigkeit von Funktionen.

Von

Gustav Doetsch in Stuttgart.

§ 1.

Der Zusammenhang zwischen der Funktionaldeterminante als Deformationsmaß und als Abhängigkeitskriterium bei n Funktionen von n Variablen.

Vor kurzem haben Knopp und R. Schmidt¹⁾ den bekannten Satz, der die *Abhängigkeit* von zwei Funktionen u, v zweier Variablen x, y (entsprechend auch bei n Funktionen von n Variablen) mit dem Verschwinden der *Funktionaldeterminante* in Zusammenhang bringt, zum ersten Male in einem für die praktischen Anwendungen brauchbaren Umfang bewiesen. Während nach den üblichen Beweisen das identische Verschwinden der Funktionaldeterminante in einem offenen Gebiet G die Abhängigkeit der Funktionen nur in einer hinreichend kleinen Umgebung jedes Punktes garantiert, zeigen K.-S., daß die Funktionen in jedem abgeschlossenen Teilbereich B abhängig sind, mit einer einheitlichen Abhängigkeitsfunktion. Dabei wird der Begriff der Abhängigkeit von ihnen dahin präzisiert, daß das durch die Transformation $u = u(x, y)$, $v = v(x, y)$ vermittelte Bild B' des Bereiches B eine in der uv -Ebene nirgends dichte Punktmenge sein soll. (Dann läßt sich nämlich leicht eine Funktion $F(u, v)$ konstruieren, die auf der Menge B' den Wert 0 hat, ohne identisch zu verschwinden, so daß $F(u(x, y), v(x, y)) = 0$ in B ist.) Die partiellen Ableitungen erster Ordnung von u und v werden von K.-S. in G als stetig vorausgesetzt.

¹⁾ Knopp und R. Schmidt, Funktionaldeterminanten und Abhängigkeit von Funktionen, Math. Zeitschr. 25 (1926), S. 373–381. Im folgenden als K.-S. zitiert.

Im Anschluß an diese Untersuchung hat dann Ostrowski²⁾ gezeigt, daß sich das obige Resultat noch verallgemeinern läßt. Um nämlich die Abhängigkeit zwischen den Funktionen in einem abgeschlossenen Bereich zu erschließen, braucht man nur das Verschwinden der Funktionaldeterminante in dem Bereich B selbst und sehr viel weniger als die Stetigkeit der Ableitungen vorauszusetzen, es genügt schon, daß die Funktionen in B ein erstes vollständiges Differential im Sinne von Stolz³⁾ besitzen, d. h. daß die partiellen ersten Ableitungen $u_x, u_y; v_x, v_y$ existieren und die Funktionen in einer Umgebung jedes Punktes (x_0, y_0) in der Form darstellbar sind:

$$u(x, y) = u(x_0, y_0) + u_x(x_0, y_0)(x - x_0) + u_y(x_0, y_0)(y - y_0) \\ + \varepsilon_1(x, y)(|x - x_0| + |y - y_0|),$$

$$v(x, y) = v(x_0, y_0) + v_x(x_0, y_0)(x - x_0) + v_y(x_0, y_0)(y - y_0) \\ + \varepsilon_2(x, y)(|x - x_0| + |y - y_0|)$$

mit $\varepsilon_1(x, y) \rightarrow 0, \varepsilon_2(x, y) \rightarrow 0$ für $|x - x_0| + |y - y_0| \rightarrow 0$.

Das Wesentliche bei K.-S. und O. liegt nicht in dem Nachweis, daß aus der Tatsache der Abhängigkeit das Verschwinden der Funktionaldeterminante folgt (dies ergibt sich aus bekannten Sätzen sehr leicht), sondern darin, daß umgekehrt aus dem Verschwinden der Funktionaldeterminante in B auf die funktionale Abhängigkeit in dem ganzen Bereich geschlossen wird, was durch den Nachweis geschieht, daß B' das Maß 0 hat. Der eigentliche Zweck dieses ersten Paragraphen ist nun, darauf hinzuweisen, daß gerade dieser Schluß ein unmittelbar und anschaulich einleuchtender ist. Bekanntlich gibt ja die Funktionaldeterminante das *flächenhafte Vergrößerungsverhältnis* an, das der ursprüngliche Bereich in seinen kleinsten Teilen bei der durch die Funktionen vermittelten Abbildung erfährt. Wenn nun bei einer Transformation dieses Verhältnis *an jeder Stelle* (im Kleinen) 0 ist, so ist es auch *im Großen* 0, d. h. B' hat den Flächeninhalt oder das Maß 0. So erschließt sich der Abhängigkeitssatz ganz intuitiv.

Nun ist allerdings zu sagen, daß in den Lehrbüchern dieser für die Lehre vom Doppelintegral, also für die ganze Analysis so wichtige Satz, daß das Deformationsmaß gleich der Funktionaldeterminante $D(x, y)$ ist, immer nur für den Fall $D \neq 0$ bewiesen wird. Das Äußerste, was man z. B. aus den mit größter Sorgfalt durchgeführten Beweisen in den Büchern

²⁾ Ostrowski, Funktionaldeterminanten und Abhängigkeit von Funktionen (Mathematische Miscellen VIII), Jahresber. d. Deutsch. Math.-Verein. 36 (1927), S. 129–134. Im folgenden als O. zitiert.

³⁾ Stolz, Grundzüge der Differential- und Integralrechnung. I. Reelle Veränderliche und Funktionen, S. 133. Leipzig 1893.

von Kowalewski⁴⁾ bei hinreichend weitgehender Ausdeutung entnehmen kann, ist folgender

Satz I. Die Funktionen $u = u(x, y)$ und $v = v(x, y)$ sollen in einem Punkte (x_0, y_0) und einer gewissen Umgebung U eindeutig definiert sein und in U partielle Ableitungen erster Ordnung $u_x, u_y; v_x, v_y$ besitzen, die im Punkte (x_0, y_0) stetig sind. Die Funktionaldeterminante

$$D(x, y) = \begin{vmatrix} u_x & u_y \\ v_x & v_y \end{vmatrix}$$

sei im Punkte (x_0, y_0) von 0 verschieden. Dann kann man zunächst um (x_0, y_0) ein Teilgebiet U_1 von U abgrenzen, in dem die durch die Funktionen u und v vermittelte Abbildung der xy - auf die uv -Ebene eineindeutig ist. Bedeutet nun Q ein ganz zu U_1 gehöriges, achsenparalleles Quadrat mit dem Mittelpunkt (x_0, y_0) , so enthält die ihm zugeordnete Punktmenge Q' der uv -Ebene eine volle Umgebung des Punktes $u_0 = u(x_0, y_0)$, $v_0 = v(x_0, y_0)$ und ist quadrierbar; für die Flächeninhalte M von Q und Q' gilt die Grenzwertbeziehung:

$$\lim \frac{M(Q')}{M(Q)} = |D(x_0, y_0)|,$$

wenn Q sich auf den Punkt (x_0, y_0) zusammenzieht.

Die Arbeiten von K.-S. und O. kann man nun so interpretieren, daß sie zunächst die *Ergänzung dieses Satzes für den Fall $D = 0$* liefern, bei K.-S. (S. 379 oben bis S. 381 Zeile 12) unter der Voraussetzung der Stetigkeit der Ableitungen im Punkte (x_0, y_0) , bei O. (S. 133 Zeile 17 bis S. 134 Zeile 9) unter der schwächeren Annahme, daß die Funktionen in (x_0, y_0) ein erstes vollständiges Differential im Sinne von Stolz haben. (Der Beweis von O. ist, obwohl allgemeiner, doch einfacher als der von K.-S.) Dann wird der oben angedeutete Schluß von dem Verschwinden des Bildinhaltes im Kleinen auf das Verschwinden im Großen exakt durchgeführt: unter der Voraussetzung von K.-S. auf sehr einfachem Wege (S. 381), da man wegen der Stetigkeit der Ableitungen die Menge B mit endlich vielen Quadraten überdecken kann, für die sämtlich dieselbe Abschätzung gilt (so daß man mit dem Jordan-Peanoschen Maß auskommt), unter der Voraussetzung von O. (S. 134) mit etwas komplizierteren Hilfsmitteln, nämlich einem Satz aus dem Ideenkreis des Vitalischen Überdeckungssatzes (wobei man das Lebesguesche Maß heranziehen muß).

Wenn nun auch die Lehrbücher den Fall $D = 0$ außer acht lassen, so lag doch schon vor den Arbeiten von K.-S. und O. eine Behandlung dieses

⁴⁾ Kowalewski, Das Integral und seine geometrischen Anwendungen, S. 63–71. Leipzig 1910; Einführung in die Determinantentheorie, S. 300–306. Leipzig 1909.

Falles vor, und zwar hat Rademacher⁵⁾ im Jahre 1918 den fraglichen Satz bereits in dem von O. benutzten Umfange bewiesen. Ich gebe im folgenden einen Beweis, der eine Kleinigkeit einfacher als der von Rademacher ist und im wesentlichen mit dem von O. übereinstimmt, nur in einer mehr geometrischen Deutung und in einer Form, die die Verallgemeinerungsfähigkeit auf beliebig viele Dimensionen etwas deutlicher hervortreten läßt. Es handelt sich um folgenden

Satz II. Die Funktionen $u = u(x, y)$ und $v = v(x, y)$ seien in einem Punkte (x_0, y_0) und einer Umgebung U eindeutig definiert; sie sollen im Punkte (x_0, y_0) ein erstes vollständiges Differential im Sinne von Stolz besitzen, und die Funktionaldeterminante $D(x, y)$ sei im Punkte (x_0, y_0) gleich 0. Grenzt man um (x_0, y_0) als Mittelpunkt ein achsenparalleles, ganz zu U gehörendes Quadrat Q von der Seitenlänge q ab und bezeichnet das äußere Maß der ihm entsprechenden Punktmenge Q' der uv -Ebene mit $M(Q')$, so gilt:

$$\lim \frac{M(Q')}{q^2} = 0 \quad (= D(x_0, y_0)),$$

wenn Q sich auf den Punkt (x_0, y_0) zusammenzieht.

Zunächst erledigen wir den Spezialfall, daß

$$v_x(x_0, y_0) = v_y(x_0, y_0) = 0$$

ist. Für ihn gilt:

$$v(x, y) - v(x_0, y_0) = \varepsilon_2(x, y)(|x - x_0| + |y - y_0|),$$

d. h. das Bild Q' des Quadrates Q in der uv -Ebene ist, selbst im Verhältnis zu $|x - x_0| + |y - y_0|$ oder, was dasselbe ist, zur Quadratseite q , in der v -Richtung auf ein sehr kleines Intervall reduziert, während es in der u -Richtung wegen

$$u(x, y) - u(x_0, y_0) = u_x(x_0, y_0)(x - x_0) + u_y(x_0, y_0)(y - y_0) + \varepsilon_1(x, y)(|x - x_0| + |y - y_0|)$$

im Verhältnis zu q wenigstens beschränkt bleibt, so daß es also eine Fläche bedeckt, die im Verhältnis zu q^2 sehr klein ist. Genauer: Wird

⁵⁾ Rademacher, Über partielle und totale Differenzierbarkeit von Funktionen mehrerer Variablen und über die Transformation der Doppelintegrale, Math. Annalen 79 (1919), S. 340–359 [S. 350, 351]. Hier ist übrigens auch der Satz I dahin verallgemeinert, daß in dem betr. Punkt nur Stolz'sche Differenzierbarkeit vorausgesetzt wird, wobei aber natürlich von umkehrbarer Eindeutigkeit in einem Gebiet U_1 sowie von Quadrierbarkeit von Q' keine Rede mehr ist. — Daß übrigens die bloße Existenz der ersten partiellen Ableitungen, sogar in einem ganzen Gebiet um (x_0, y_0) herum, nicht genügt, den Satz noch zu sichern, hat Rademacher durch ein Beispiel gezeigt: Eineindeutige Abbildungen und Meßbarkeit, Monatshefte f. Math. u. Phys. 27 (1916), S. 183–291 [S. 289].

eine beliebig kleine positive Zahl δ vorgegeben, so kann man ein q_0 so bestimmen, daß für

$$|x - x_0| = \frac{q_0}{2}, \quad |y - y_0| = \frac{q_0}{2}$$

gilt:

$$|\varepsilon_1(x, y)| < 1, \quad |\varepsilon_2(x, y)| < \delta.$$

Sobald also die Quadratseite $q \leq q_0$ ist, erhalten wir für die Breitenausdehnung des Bildes die Abschätzung:

$$|u(x, y) - u(x_0, y_0)| < (|u_x(x_0, y_0)| + |u_y(x_0, y_0)| + 2) \frac{q}{2} = C \frac{q}{2}$$

und für die Höhenausdehnung:

$$|v(x, y) - v(x_0, y_0)| < \delta q.$$

Das Bild Q' liegt also in einem Rechteck mit den Seitenlängen Cq und $2\delta q$, sein äußeres Maß $M(Q')$ ist mithin kleiner als $2C\delta q^2$ und der Quotient $M(Q') : q^2$ folglich beliebig klein, w. z. b. w.

Den allgemeinen Fall, wo $v_x(x_0, y_0)$ und $v_y(x_0, y_0)$ nicht beide verschwinden, können wir nun sofort auf den behandelten Spezialfall zurückführen. Aus dem Verschwinden der Funktionaldeterminante folgt nämlich, daß die Gleichungen

$$\begin{aligned} \lambda_1 u_x(x_0, y_0) + \lambda_2 v_x(x_0, y_0) &= 0, \\ \lambda_1 u_y(x_0, y_0) + \lambda_2 v_y(x_0, y_0) &= 0 \end{aligned}$$

eine von der trivialen ($\lambda_1 = 0, \lambda_2 = 0$) verschiedene Lösung λ_1, λ_2 besitzen. Es ist sicher $\lambda_1 \neq 0$, denn andernfalls würde, da v_x und v_y nicht beide verschwinden, auch noch $\lambda_2 = 0$ folgen. Wir gehen nun von der uv -Ebene durch die lineare Transformation

$$\begin{aligned} u^* &= v \\ v^* &= \lambda_1 u + \lambda_2 v \end{aligned}$$

zu einer u^*v^* -Ebene über, wobei das Maß von Q' sich nur mit der von 0 verschiedenen Konstanten $|\lambda_1|$ multipliziert. Für die Abbildung der xy -Ebene auf die u^*v^* -Ebene treffen jetzt aber dieselben Verhältnisse zu wie in dem vorausbehandelten Spezialfall; es ist nämlich:

$$\begin{aligned} v_x^*(x_0, y_0) &= \lambda_1 u_x(x_0, y_0) + \lambda_2 v_x(x_0, y_0) = 0, \\ v_y^*(x_0, y_0) &= \lambda_1 u_y(x_0, y_0) + \lambda_2 v_y(x_0, y_0) = 0. \end{aligned}$$

Die Behauptung unseres Satzes folgt also allgemein.

Geometrisch liegt der Reduktion des allgemeinen auf den speziellen Fall folgender Gedanke zugrunde: Betrachtet man die Linienelemente durch (x_0, y_0) in Richtung der x - und y -Achse ($y = y_0 = \text{konst. bzw.}$

$x = x_0 = \text{konst.}$), so entspricht dem ersten in der uv -Ebene das Linien-
element mit den Komponenten $du = u_x(x_0, y_0) dx$, $dv = v_x(x_0, y_0) dx$,
dem zweiten $du = u_y(x, y) dy$, $dv = v_y(x, y) dy$. Wegen des Verschwindens
der Funktionaldeterminante sind die beiden Elemente linear abhängig,
fallen also in dieselbe Richtung. Diese braucht man nur durch eine ge-
eignete affine Transformation zur u -Richtung zu machen, um auf den
Spezialfall zurückzukommen, in dem $v_x = v_y = 0$ war, in dem also jene
Richtung von vornherein mit der Richtung der u -Achse zusammenfiel.

Die *Verallgemeinerung* auf mehr als zwei Dimensionen ist jetzt un-
mittelbar einleuchtend. Hat man z. B. drei Funktionen u, v, w von drei
Variablen x, y, z , so nimmt man den Fall voraus, daß $w_x = w_y = w_z = 0$
ist. Hier werden die Linienelemente in Richtung der x -, y - und z -Achse
offenbar in solche übergeführt, die sämtlich in der uv -Ebene liegen, also
ein verschwindendes Parallelepiped aufspannen. Sind dagegen die partiellen
Ableitungen von w nicht sämtlich 0, so zeigt doch das Verschwinden der
Funktionaldeterminante, daß die Bilder der Linienelemente linear abhängig
sind, also in einer Ebene liegen. Diese macht man durch eine affine,
nicht ausgeartete Transformation zur uv -Ebene, und zwar folgendermaßen:
Aus dem Verschwinden der Determinante folgt, daß die linearen Gleichungen

$$\begin{aligned} \lambda_1 u_x(x_0, y_0) + \lambda_2 v_x(x_0, y_0) + \lambda_3 w_x(x_0, y_0) &= 0 \\ \lambda_1 u_y(x_0, y_0) + \lambda_2 v_y(x_0, y_0) + \lambda_3 w_y(x_0, y_0) &= 0 \\ \lambda_1 u_z(x_0, y_0) + \lambda_2 v_z(x_0, y_0) + \lambda_3 w_z(x_0, y_0) &= 0 \end{aligned}$$

eine von der trivialen $(0, 0, 0)$ verschiedene Lösung $\lambda_1, \lambda_2, \lambda_3$ besitzen.
Dabei muß mindestens eine der Zahlen $\lambda_1, \lambda_2 \neq 0$ sein. Denn wären
beide gleich 0, so würde auch noch $\lambda_3 = 0$ folgen, da w_x, w_y und w_z nicht
sämtlich verschwinden. Es sei z. B. $\lambda_1 \neq 0$. Dann leistet die Transformation

$$\begin{aligned} u^* &= u \\ v^* &= v \\ w^* &= \lambda_1 u + \lambda_2 v + \lambda_3 w \end{aligned}$$

das Verlangte.

§ 2.

Abhängigkeit zwischen n Funktionen von m Variablen.

In § 1 wurde ein System von Funktionen betrachtet, deren Anzahl
gleich der Anzahl der unabhängigen Variablen war. Wir wollen nun den
allgemeineren Fall behandeln, daß diese Gleichheit nicht zutrifft. Zunächst
ergibt sich sehr leicht der

Satz III. *Ist die Anzahl der im Stolzischen Sinne differenzierbaren Funktionen größer als die Anzahl der unabhängigen Variablen, so liegt stets eine Abhängigkeit vor.*

Handelt es sich z. B. um drei Funktionen u, v, w von zwei Variablen x, y , die in einem Bereich B der xy -Ebene definiert sind, so brauchen wir zu x und y nur noch eine dritte Variable z etwa mit dem Variabilitätsbereich $-1 \leq z \leq +1$ hinzuzunehmen und u, v, w als Funktionen dieser drei Variablen zu betrachten, die aber in Wahrheit gar nicht von z abhängen, so daß $u_z = v_z = w_z = 0$ ist. Die Funktionaldeterminante

$$\begin{vmatrix} u_x & u_y & u_z \\ v_x & v_y & v_z \\ w_x & w_y & w_z \end{vmatrix} = \begin{vmatrix} u_x & u_y & 0 \\ v_x & v_y & 0 \\ w_x & w_y & 0 \end{vmatrix}$$

verschwindet identisch, folglich sind die drei Funktionen in dem durch B und das Intervall $-1 \leq z \leq +1$ bestimmten Bereich des xyz -Raumes, also um so mehr in dem Bereich B der xy -Ebene abhängig.

Wir gehen nun zu dem Fall über, daß die Anzahl n der Funktionen u_1, u_2, \dots, u_n kleiner ist als die Anzahl m der unabhängigen Variablen x_1, x_2, \dots, x_m . Zu einem solchem System gehört eine *Funktionalmatrix*

$$M = \begin{vmatrix} u_{11} & u_{12} & \dots & u_{1m} \\ \cdot & \cdot & \cdot & \cdot \\ u_{n1} & u_{n2} & \dots & u_{nm} \end{vmatrix}, \quad u_{ik} = \frac{\partial u_i}{\partial x_k},$$

deren Verhalten die Abhängigkeit oder Unabhängigkeit der Funktionen bestimmt. Diese Matrix hat an jeder Stelle $P(x_1, x_2, \dots, x_m)$ einen bestimmten Rang $r = r(P)$, der höchstens gleich n sein kann.

Ich behaupte zunächst den

Satz IV. *Wenn die Funktionen in einem beschränkten, offenen Gebiet G im Stolzischen Sinne differenzierbar und abhängig sind, so ist der Rang von M notwendig dort überall kleiner als n , d. h. sämtliche n -reihigen Determinanten aus M verschwinden identisch.*

Um dies zu beweisen, greifen wir einen beliebigen Punkt $P_0(x_1^0, x_2^0, \dots, x_m^0)$ des Gebietes heraus und legen durch ihn eine n -dimensionale lineare Mannigfaltigkeit, indem wir $m - n$ der Variablen, z. B. $x_{n+1}, x_{n+2}, \dots, x_m$, konstant halten. Diese Mannigfaltigkeit

$$x_{n+1} = x_{n+1}^0, \dots, x_m = x_m^0$$

hat, da P_0 ein innerer Punkt in G ist, mit G in der Umgebung von P_0 ein n -dimensionales Gebiet g gemein, in dem die n Funktionen von n Variablen $u_i(x_1, \dots, x_n, x_{n+1}^0, \dots, x_m^0)$ im Stolzischen Sinne differenzierbar

und abhängig sind. Also ist nach dem von O. bewiesenen Satz (es genügt schon sein Hilfssatz 4, a. a. O. S. 131) die Funktionaldeterminante dieser Funktionen in P_0 gleich 0⁶⁾. Diese lautet aber:

$$\begin{vmatrix} u_{11}(x_1^0, \dots, x_m^0) & u_{12} & \dots & u_{1n} \\ \dots & \dots & \dots & \dots \\ u_{n1} & \dots & \dots & u_{nn} \end{vmatrix}.$$

Da dies für jeden Punkt von G gilt, so verschwindet die Determinante in G identisch. Genau ebenso beweist man das Verschwinden der übrigen n -reihigen Determinanten.

Wir wollen nun umgekehrt eine *hinreichende* Bedingung angeben, unter der die n Funktionen von m Variablen abhängig sind. Da hier jedoch die für $n = m$ und $n > m$ benutzten Methoden, wie man sofort sieht, versagen, sind wir nicht imstande, etwa obige notwendige Bedingung ohne weitere Zusatzbedingungen als hinreichend nachzuweisen. Zunächst wollen wir von jetzt an voraussetzen, daß die sämtlichen in der Funktionalmatrix vorkommenden Ableitungen in dem betrachteten Gebiet *stetig* sind. Dann lassen sich die Punkte des Gebietes in zwei Klassen einteilen, in denen die Matrix verschiedenes Verhalten zeigt. Ist nämlich der Rang von M in einem Punkte P_0 gleich r_0 , so bedeutet das, daß mindestens eine r_0 -reihige Determinante aus M in P_0 von 0 verschieden ist; wegen der Stetigkeit der Elemente verschwindet diese Determinante dann in einer ganzen Umgebung von P_0 nicht, es ist also dort der Rang $r(P) \geq r_0$. Wenn sich nun um P_0 eine Umgebung so abgrenzen läßt, daß überall in ihr $r(P) = r_0$ ist, so wollen wir den Punkt P_0 *regulär* nennen, im anderen Falle, d. h. wenn es in jeder Umgebung Punkte gibt, wo $r(P) > r_0$ ist, soll P_0 *singulär* heißen. Bei Verwendung dieser Terminologie gilt folgender

Satz V. Die Funktionen u_1, \dots, u_n sollen in einem offenen, beschränkten Gebiet G des $x_1 \dots x_m$ -Raumes definiert sein und dort stetige partielle Ableitungen erster Ordnung besitzen; der Rang der Funktionalmatrix der u sei überall in G kleiner als n . Dann sind die Funktionen in jedem abgeschlossenen Teilbereich B von G , der nur aus regulären Punkten besteht, abhängig.

Zum Beweise stützen wir uns auf einen Abhängigkeitssatz „im Kleinen“ der aussagt, daß, wenn der Rang von M in P gleich r und P ein regulärer Punkt ist, in einer gewissen hinreichend kleinen Umgebung U von P genau r der Funktionen u unabhängig sind, während die $n - r$ übrigen

⁶⁾ Den Hilfssatz 4 von O. kann man statt aus einem allgemeinen Satz von W. H. Young, wie O. es tut, auch aus dem Rademacherschen Beweis des verallgemeinerten Satzes I (loc. cit. ⁵⁾, S. 352) entnehmen.

($n - r$ ist ≥ 1) sich als Funktionen jener r unabhängigen darstellen lassen⁷⁾. Das Abbild von U im $u_1 \dots u_n$ -Raum ist also eine r -dimensionale Mannigfaltigkeit, die sich, wie man des näheren aus dem zitierten Beweise ersieht, eindeutig auf eine r -dimensionale „Ebene“ (die lineare Mannigfaltigkeit der unabhängigen u) beziehen läßt und also im n -dimensionalen Raum das Maß 0 hat und nirgends dicht liegt. Ordnen wir nun jedem Punkt des Bereiches B eine solche Umgebung U zu, so können wir, da B abgeschlossen ist, diesen Bereich nach dem Borelschen Überdeckungssatz auch schon mit endlich vielen der Gebiete U überdecken. Die Gesamtheit ihrer Bilder im u -Raum enthält sicher das Bild von B und ist als Vereinigungsmenge von endlich vielen nirgends dichten Mengen selbst nirgends dicht; das gleiche gilt also für das Bild von B .

Natürlich kann auch Abhängigkeit in Bereichen bestehen, die singuläre Punkte enthalten. Eine hinreichende Bedingung, unter der dies stattfindet, wird gegeben durch

Satz VI. *Unter den Voraussetzungen von Satz V sind die Funktionen abhängig in jedem abgeschlossenen Bereich, dessen singuläre Punkte bei der Abbildung in eine Nullmenge übergeführt werden.*

Vorbemerkung. Es ist klar, daß die singulären Punkte im Bereich B eine abgeschlossene Menge S bilden, d. h. daß jeder Häufungspunkt von singulären Punkten selbst singulär ist. Denn in der Umgebung eines regulären Punktes liegen nur reguläre Punkte. Ferner ist S in dem (größeren) Gebiet G , also auch in dem Bereich B , nirgends dicht, d. h. die inneren Punkte von $G - S$ oder, da $G - S$ als offene Menge nur aus inneren Punkten besteht: alle Punkte von $G - S$ (das sind die regulären Punkte von G) liegen in G überall dicht. Angenommen nämlich, es gäbe in G eine volle m -dimensionale Kugel mit dem Mittelpunkt P_0 , in der kein regulärer Punkt läge, so wäre P_0 singulär, man könnte also in der Kugel einen Punkt P_1 finden, wo der Rang $r(P_1)$ der Funktionalmatrix größer als in P_0 wäre; da P_1 ebenfalls singulär ist, so könnte man in seiner Umgebung, noch innerhalb der Kugel, einen Punkt P_2 finden, wo $r(P_2) > r(P_1)$ wäre usw. Man erhielte so eine Punktfolge mit der Eigenschaft

$$r(P_0) < r(P_1) < r(P_2) < \dots,$$

was unmöglich ist. — Da wir es mit einer stetigen Abbildung zu tun haben, so ist das Bild von S auch abgeschlossen (daher sicher meßbar) und nirgends dicht. Weil es aber abgeschlossene, nirgends dichte Mengen von positivem Maß gibt, so ist die Voraussetzung, daß das Bild von S eine Nullmenge sein soll, nicht trivial.

⁷⁾ Beweis bei Kowalewski, Einführung in die Determinantentheorie, S. 306–311. Leipzig 1910.

Nunmehr gehen wir zum Beweis unseres Satzes VI über. Wir betrachten die Menge der regulären Punkte von B , d. h. $B - S$. Zu jedem solchen Punkte gehört eine Umgebung U , die auf eine Menge vom Maße 0 abgebildet wird (s. S. 598). Nach dem Lindelöf'schen Überdeckungssatz können wir $B - S$ durch höchstens abzählbar unendlich viele solcher Umgebungen U vollständig überdecken. Die Bilder dieser U überdecken dann das Bild von $B - S$; ihre Vereinigungsmenge hat als Summe von abzählbar vielen Nullmengen das Maß 0, folglich auch das Bild von $B - S$. Das Bild von S ist nach Voraussetzung eine Nullmenge, mithin hat das Bild von $B = (B - S) + S$ das Maß 0. Da aber B abgeschlossen und wegen der Stetigkeit der Abbildung auch sein Bild abgeschlossen ist, so folgt daraus, daß das Bild von B nirgends dicht ist.

Man kann *allgemeine Fälle* angeben, in denen die Voraussetzungen von Satz VI erfüllt sind, in denen also vor allem die singulären Punkte in eine Nullmenge transformiert werden. Um uns kurz ausdrücken zu können, führen wir folgende Bezeichnungsweise ein. Im m -dimensionalen Raum sei eine p -dimensionale Mannigfaltigkeit dadurch gegeben, daß $m - p$ der Variablen als Funktionen der p übrigen dargestellt sind. Wir können die Numerierung der Variablen so treffen, daß die Darstellung so lautet:

$$\begin{aligned}
 x_{p+1} &= x_{p+1}(x_1, \dots, x_p), \\
 &\dots \dots \dots \dots \dots \dots \dots \\
 x_m &= x_m(x_1, \dots, x_p).
 \end{aligned}$$

Wenn dann die Funktionen x_{p+1}, \dots, x_m in einem Bereich D der Variablen x_1, \dots, x_p stetige partielle Ableitungen erster Ordnung besitzen, so wollen wir die Mannigfaltigkeit „glatt“ nennen.

Dann gilt der

Satz VII. *Es mögen die Voraussetzungen von Satz V gelten. Ferner sei B ein abgeschlossener Teilbereich von G , dessen singuläre Punkte auf endlich vielen oder abzählbar unendlich vielen glatten, höchstens n -dimensionalen Mannigfaltigkeiten liegen. Dann sind die Funktionen in B abhängig.*

Wir können annehmen, daß jede der fraglichen Mannigfaltigkeiten genau n -dimensional ist, da man eine niedrigerdimensionale Mannigfaltigkeit stets zu einer n -dimensionalen erweitern kann, z. B. indem man unter ihren Gleichungen so viele streicht, daß $m - n$ übrigbleiben. Ferner genügt es nachzuweisen, daß jede n -dimensionale glatte Mannigfaltigkeit, soweit sie in B liegt, in eine Nullmenge übergeführt wird. Denn dann hat auch das Bild der abzählbar vielen Mannigfaltigkeiten das Maß 0 und folglich auch das Bild der Menge der auf ihnen liegenden Punkte. Damit ist nach Satz VI die Abhängigkeit der Funktionen gesichert.

Die Mannigfaltigkeit habe also die Gleichungen:

$$\begin{aligned}x_{n+1} &= x_{n+1}(x_1, \dots, x_n), \\ \dots & \dots \dots \dots \dots \dots \dots \dots \\ x_m &= x_m(x_1, \dots, x_n).\end{aligned}$$

Die ihr zugeordneten Punkte im u -Raum werden dargestellt durch:

$$\begin{aligned}u_i &= u_i(x_1, \dots, x_n, x_{n+1}(x_1, \dots, x_n), \dots, x_m(x_1, \dots, x_n)) \\ &= u_i^*(x_1, \dots, x_n) \quad (i = 1, \dots, n).\end{aligned}$$

Dies sind n Funktionen von n Variablen, die in der Projektion G^* von G auf den $x_1 \dots x_n$ -Raum definiert sind und dort stetige Ableitungen erster Ordnung besitzen, die folgende Gestalt haben:

$$\frac{\partial u_i^*}{\partial x_k} = \frac{\partial u_i}{\partial x_k} + \frac{\partial u_i}{\partial x_{n+1}} \frac{\partial x_{n+1}}{\partial x_k} + \frac{\partial u_i}{\partial x_{n+2}} \frac{\partial x_{n+2}}{\partial x_k} + \dots + \frac{\partial u_i}{\partial x_m} \frac{\partial x_m}{\partial x_k}$$

oder

$$u_{ik}^* = u_{ik} + u_{i, n+1} x_{n+1, k} + u_{i, n+2} x_{n+2, k} + \dots + u_{i, m} x_{m, k}.$$

Ihre Funktionaldeterminante

$$\begin{vmatrix} u_{11} + u_{1, n+1} x_{n+1, 1} + \dots + u_{1, m} x_{m, 1} & \dots & u_{1, n} + u_{1, n+1} x_{n+1, n} + \dots + u_{1, m} x_{m, n} \\ u_{21} + u_{2, n+1} x_{n+1, 1} + \dots + u_{2, m} x_{m, 1} & \dots & u_{2, n} + u_{2, n+1} x_{n+1, n} + \dots + u_{2, m} x_{m, n} \\ \dots & \dots & \dots \\ u_{n1} + u_{n, n+1} x_{n+1, 1} + \dots + u_{n, m} x_{m, 1} & \dots & u_{n, n} + u_{n, n+1} x_{n+1, n} + \dots + u_{n, m} x_{m, n} \end{vmatrix}$$

läßt sich auflösen in eine Summe von n -reihigen Determinanten, die teilweise eo ipso den Wert 0 haben, da sie proportionale Kolonnen besitzen, während die übrigen nach Voraussetzung verschwinden, da der Rang der Funktionalmatrix der u in G überall kleiner als n sein sollte. Nach dem von K.-S. bewiesenen Satze⁸⁾ hat also das durch die u^* von der Projektion der Mannigfaltigkeit, soweit sie in B liegt, oder, was dasselbe ist, das durch die u von unserer Mannigfaltigkeit selbst entworfene Bild das Maß 0.

Der Satz VII ist vor allem dann leicht anzuwenden, wenn $m = n + 1$ ist und die singulären Punkte auf glatten Mannigfaltigkeiten liegen. Denn diese können dann in der Tat höchstens die Dimension n haben.

Ein spezielles Beispiel: $m = 3$, $n = 2$; zwei Funktionen u, v von drei Variablen x, y, z . Ist der Rang der Funktionalmatrix M überall in dem betrachteten Gebiet kleiner als zwei, so kommen als singuläre Punkte nur solche in Betracht, wo der Rang von M gleich 0 ist, d. h. wo sämtliche erste Ableitungen von u und v verschwinden und mithin beide Funktionen ausgezeichnete Werte besitzen (stationär sind). Von diesen fallen noch diejenigen als regulär weg, zu denen eine volle Umgebung gleich-

⁸⁾ Loc. cit. ¹⁾. Dieser Satz reicht hier wegen der Stetigkeit der $\frac{\partial u_i^*}{\partial x_k}$ aus.

artiger Punkte existiert (Gebiete, wo u und v beide konstant sind), denn damit ein Punkt singular ist, muß jede Umgebung Stellen enthalten, wo der Rang von M gleich eins ist, d. h. wo mindestens eine der Ableitungen von 0 verschieden ist. Satz VII liefert also folgendes: Wenn bei Stetigkeit der partiellen Ableitungen erster Ordnung die drei Determinanten

$$\begin{vmatrix} u_x & u_y \\ v_x & v_y \end{vmatrix}, \quad \begin{vmatrix} u_y & u_z \\ v_y & v_z \end{vmatrix}, \quad \begin{vmatrix} u_x & u_z \\ v_x & v_z \end{vmatrix}$$

in einem offenen Gebiet G verschwinden und die Stellen, wo beide Funktionen zugleich stationär sind, d. h. wo die sechs Gleichungen für drei Unbekannte x, y, z : $u_x = u_y = u_z = v_x = v_y = v_z = 0$ erfüllt sind, sich, wenn überhaupt vorhanden, zu dreidimensionalen Gebieten zusammenschließen, bzw. sich auf höchstens abzählbar vielen zweidimensionalen Flächen mit stetig veränderlicher Tangentialebene lokalisieren lassen, so sind die beiden Funktionen in jedem abgeschlossenen Teilbereich von G abhängig.

(Eingegangen am 31. 10. 1927.)

Zur Theorie der Differentialgleichungen.

Von

E. Kamke in Tübingen.

In der Theorie der reellen Differentialgleichungen begegnet man vielfach folgendem Satz: Ist die rechte Seite der Differentialgleichung

$$(1) \quad y' = f(x, y)$$

etwa in einem achsenparallelen Quadrat \mathfrak{Q} stetig und beschränkt und ist für sie überdies die Lipschitzbedingung erfüllt, so lassen sich die Integralcurven von (1) durch Auflösung einer Gleichung

$$\psi(x, y) = c$$

gewinnen, wobei $\psi(x, y)$ eine passend zu wählende, im ganzen Quadrat erklärte Funktion und c ein beliebiger Wert eines Intervalls $c_1 < c < c_2$ ist.

In der Literatur pflegt diese Tatsache, wenn überhaupt, so folgendermaßen begründet zu werden: Ist

$$y = \varphi(x, \xi, \eta)$$

die durch den Punkt ξ, η von \mathfrak{Q} gehende Integralkurve, so halte man $x = x_0$ fest und setze

$$\psi(\xi, \eta) = \varphi(x_0, \xi, \eta).$$

Nun gibt es unter den angegebenen Voraussetzungen zwar durch jeden Punkt ξ, η des Quadrats \mathfrak{Q} eine Integralkurve von (1), aber diese Integralkurven brauchen keineswegs sämtlich eine feste Ordinate mit der Abszisse x_0 zu schneiden. Ein Beispiel hierfür liefert die Differentialgleichung

$$y' = \frac{y^2}{x} \quad \text{in der Halbebene } x > 0$$

mit

$$\varphi(x, \xi, \eta) = \frac{\eta}{1 + \eta \log \frac{\xi}{x}}.$$

Man erhält daher durch Auflösung der Gleichung

$$\psi(\xi, \eta) = \varphi(x_0, \xi, \eta) = y$$

nach ξ, η keineswegs stets *sämtliche* Integralkurven¹⁾.

Gleichwohl ist der eingangs angeführte Satz richtig, und zwar trivial, da die Menge der Integralkurven, wie leicht zu sehen, von der Mächtigkeit des Kontinuums ist, also einem beliebigen Intervall $c_1 < c < c_2$ eindeutig zugeordnet werden kann. Die so entstehende Funktion $\psi(\xi, \eta)$ kann natürlich recht kompliziert sein.

In den Anwendungen braucht man den Satz allerdings in schärferer Form; man braucht nämlich die Differenzierbarkeit von $\psi(\xi, \eta)$ und noch verschiedene andere Eigenschaften. Ein derartiger Satz ist zuerst von I. Bendixson²⁾ bewiesen, aber, da $\psi(\xi, \eta) = \varphi(x_0, \xi, \eta)$ gesetzt wird, auch nur für ein Teilgebiet von \mathfrak{D} , das erst übersehen werden kann, wenn die Differentialgleichung (1) schon wirklich gelöst ist. Die Gültigkeit dieses Satzes nun wird im folgenden auf ein nahezu beliebiges Teilgebiet des Ausgangsgebietes ausgedehnt (§ 1—3). Damit ist dann — und hierin liegt die Bedeutung dieses Satzes — für die einfachste partielle lineare Differentialgleichung ein Existenzsatz für ein von vornherein angebbares großes Gebiet gewonnen, während die bisherigen Beweise die Existenz nur in einer gewissen Umgebung eines Punktes erkennen ließen. Die Einzelheiten finden sich neben anderen Zusätzen im § 4.

§ 1.

Formulierung des Hauptsatzes.

Von einem offenen Gebiet g sage ich, es *liege total* in einem offenen Gebiet \mathfrak{G} , wenn nicht nur g , sondern auch jeder (endliche) Randpunkt von g zu \mathfrak{G} gehört. Hiernach liegt z. B. der Streifen $|x| < 1 - e^{-x^2}$ total in dem Streifen $|x| < 1$ und die ganze Ebene total in sich selber.

Satz 1 (Hauptsatz). *In dem offenen Gebiet \mathfrak{G} möge die Funktion $f(x, y)$ nebst ihrer partiellen Ableitung $f'_y(x, y)$ vorhanden und stetig sein. Dann gibt es für jedes offene und einfach zusammenhängende Gebiet g ,*

¹⁾ Auch eine von Herrn G. Hoheisel [Gewöhnliche Differentialgleichungen, Berlin u. Leipzig 1926, S. 29] angegebene Überlegung ist nicht überzeugend. Denn u. a. wird daraus, daß $y = y(x)$ ein Integral der Differentialgleichung $y' = f(x, y) + N$ ist, geschlossen, daß $\bar{y} = y(x) - Nx$ ein Integral von $y' = f(x, y)$ sei. Hierfür müßte $\bar{y}' = f(x, \bar{y})$ sein, während sich nur $\bar{y}' = y'(x) - N = f(x, y)$ ergibt.

²⁾ Bulletin de la Société Math. de France 23 (1895), p. 220—225; für ein allgemeineres Ausgangsgebiet und zugleich für Systeme von Differentialgleichungen: E. Lindelöf, Journal de Math. (5) 10 (1900), p. 423 ff.

das total in \mathcal{G} liegt und in dem $f(x, y)$ beschränkt ist³⁾, eine Funktion $\psi(x, y)$ mit folgenden Eigenschaften:

a) $\psi(x, y)$ hat g als Definitionsbereich;

b) $\psi(x, y)$ hat für alle Punkte einer und derselben Integralkurve der Differentialgleichung

$$(1) \quad y' = f(x, y)$$

einen konstanten Wert;

c) $\psi(x, y)$ hat stetige partielle Ableitungen erster Ordnung, und zwar ist

$$(2) \quad \psi'_y(x, y) > 0,$$

$$(3) \quad \psi'_x(x, y) + f(x, y) \psi'_y(x, y) \equiv 0.$$

Aus diesen Behauptungen folgt sofort weiter:

d) Der Wertevorrat von $\psi(x, y)$ erfüllt genau ein offenes Intervall $a < z < b$ ($a = -\infty$, $b = +\infty$ zugelassen).

Denn sicher erfüllt er ein Intervall, da ψ wegen der aus c) folgenden Stetigkeit jeden Zwischenwert annimmt. Das Intervall ist aber auch offen; denn wäre z. B. b ein Maximum, etwa $\psi(x_0, y_0) = b$, so gäbe es wegen (2) einen ebenfalls in g gelegenen Punkt x_0, y_1 , für den $\psi(x_0, y_1) > b$ wäre.

e) Für jedes $a < z_0 < b$ liefert die Auflösung von

$$(4) \quad \psi(x, y) = z_0$$

eine oder mehrere, jedoch höchstens abzählbar viele Integralkurven von (1) völlig; diese liegen „nebeneinander“, das soll heißen: keine zwei schneiden dieselbe ganz in g verlaufende vertikale Strecke.

Genügt nämlich ein Punkt einer Integralkurve der Gleichung (4), so wird sie wegen b) auch von allen Punkten x, y dieser Integralkurve erfüllt. Da sie wegen d) und der Voraussetzung $a < z_0 < b$ mindestens eine Lösung besitzt, wird somit (4) von den Punkten mindestens einer Integralkurve erfüllt. Genügen der Gleichung die Punkte von zwei Integralkurven und würde für zwei ihrer Punkte x_0, y_1 und x_0, y_2 auch noch ihre Verbindungsstrecke zu g gehören, so würde wegen (2) und (4) sich der Widerspruch $z_0 = \psi(x_0, y_1) \geq \psi(x_0, y_2) = z_0$ ergeben. Daraus folgt nun weiter: wenn man bei jeder Integralkurve, die eine Lösung von (4) ist, um irgendeinen ihrer Punkte als Mittelpunkt ein in g liegendes

³⁾ Also sicher in jedem beschränkten einfach zusammenhängenden Gebiet g , das total in \mathcal{G} liegt.

achsenparalleles Rechteck konstruiert, derart, daß nur seine beiden *vertikalen* Seiten von der Integralkurve geschnitten werden, so liegen diese Rechtecke getrennt, sind daher höchstens abzählbar viele, und somit gibt es auch höchstens abzählbar viele Integralkurven von (1), die (4) bei festem z_0 erfüllen.

Hieraus folgt nun schließlich:

f) *Läßt man z das offene Intervall (a, b) durchlaufen und löst man für jedes dieser z die Gleichung $\psi(x, y) = z$, so erhält man die Gesamtheit der in g verlaufenden Integralkurven von (1), und zwar jede genau einmal.*

Denn, wenn x, y ein Punkt einer Integralkurve ist, gehört er zum Definitionsbereich von ψ , liefert also einen Wert $z = \psi(x, y)$ mit $a < z < b$. Man erhält aber jede Integralkurve auch nur einmal, da $\psi(x, y)$ eine eindeutige Funktion ist.

§ 2.

Topologischer Teil des Beweises.

1. Die Funktion $\psi(x, y)$, deren Existenz nachzuweisen ist, wird im § 3 schrittweise konstruiert werden. Zur Vorbereitung dieser Konstruktion zeige ich hier, daß sich das Gebiet g in ganz bestimmter Weise in abzählbar vielen Schritten aufbauen läßt. Mit Rücksicht auf eine Bemerkung im § 4 sei schon hier darauf hingewiesen, daß die Entwicklungen dieses Paragraphen nicht an die Existenz einer stetigen Ableitung $f'_y(x, y)$ gebunden sind. Es genügt vielmehr reichlich für diesen Paragraphen, wenn $f(x, y)$ in g stetig ist und die Lipschitzbedingung erfüllt. Es wird nämlich nur benutzt werden, daß durch jeden Punkt von g genau eine Integralkurve von (1) geht, daß die Integralkurven stetig vom Anfangspunkt abhängen und daß sie in jeder der beiden Richtungen der x -Achse entweder unbeschränkte Ordinaten haben oder beliebig nahe an den Rand von g herankommen.

2. Es bezeichne \mathfrak{s} eine endliche vertikale offene Strecke in g und $v(\mathfrak{s})$ die Menge der Punkte in g , die den durch \mathfrak{s} gehenden Integralkurven der Differentialgleichung (1) angehören. *Jedes $v(\mathfrak{s})$ ist ein offenes Gebiet⁴⁾.*

Denn erstens ist zu zeigen, daß mit jedem Punkt P von $v(\mathfrak{s})$ auch eine gewisse Umgebung von P zu $v(\mathfrak{s})$ gehört. Wäre dieses nicht der Fall, so gäbe es, da P zu g gehört, eine in g , aber nicht in $v(\mathfrak{s})$ liegende Punktfolge $P_n \rightarrow P$. Da P zu $v(\mathfrak{s})$ gehört, geht die durch P laufende

⁴⁾ Übrigens sogar ein einfach zusammenhängendes Gebiet, und zwar auch dann, wenn g nicht einfach zusammenhängend ist.

Integralkurve von (1) durch einen (inneren) Punkt von \mathfrak{s} . Da die Integralkurven stetig vom Anfangspunkt abhängen⁵⁾, gäbe es dann wegen $P_n \rightarrow P$ ein solches P_n , daß die durch P_n gehende Integralkurve bis zur Abszisse von \mathfrak{s} in g läge und durch einen Punkt von \mathfrak{s} ginge. Dann würde aber P_n entgegen der Annahme doch zu $\mathfrak{o}(\mathfrak{s})$ gehören.

Zweitens lassen sich auch je zwei Punkte P und Q von $\mathfrak{o}(\mathfrak{s})$ durch eine stetige Kurve in $\mathfrak{o}(\mathfrak{s})$ verbinden; denn von P und Q führen ja Integralkurven innerhalb $\mathfrak{o}(\mathfrak{s})$ zur Strecke \mathfrak{s} .

3. Es seien P und Q zwei senkrecht übereinander liegende Punkte von $\mathfrak{o}(\mathfrak{s})$, und es möge die Strecke \overline{PQ} ganz zu g gehören; dann gehört die Strecke \overline{PQ} auch ganz zu $\mathfrak{o}(\mathfrak{s})$.

Die durch P und Q gehenden Integralkurven (Fig. 1) führen nämlich zu zwei Punkten P_1 und Q_1 auf \mathfrak{s} . Ist P der höher gelegene der beiden Punkte P und Q , so liegt die durch P gehende Integralkurve ständig höher als die durch Q gehende, da sonst durch einen Punkt von g zwei Integralkurven gingen. Es wird also durch die Kurvenstücke PP_1 und QQ_1 und durch die vertikalen Strecken \overline{PQ} und $\overline{P_1Q_1}$ ein Bereich b abgegrenzt, der ganz zu g gehört, da g einfach zusammenhängend ist. Es sei nun R ein Punkt zwischen P und Q . Durch den Punkt R gibt es, da er zu g gehört, eine wohlbestimmte Integralkurve \mathfrak{R} , von der ich nur zu zeigen brauche, daß sie durch \mathfrak{s} geht. Da sowohl unter den Voraussetzungen des Hauptsatzes wie auch unter den auf der vorigen Seite angeführten jede Integralkurve in jeder der beiden Richtungen der x -Achse beliebig nahe an den Rand von g herankommt, schneidet die bei R in den Bereich b eintretende Kurve \mathfrak{R} dessen Rand nochmals. Da \mathfrak{R} als Integralkurve mit den Integralkurven PP_1 und QQ_1 keinen Punkt gemeinsam haben kann, schneidet somit \mathfrak{R} die Strecke P_1Q_1 ; w. z. b. w.

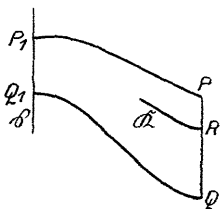


Fig. 1.

4. Unter den $\mathfrak{o}(\mathfrak{s})$ gibt es eine Folge $\mathfrak{o}_1, \mathfrak{o}_2, \dots$ mit den Eigenschaften:

a) $\mathfrak{o}_1 + \mathfrak{o}_2 + \dots = g$;

β) kein \mathfrak{o}_v ist in $\mathfrak{g}_{v-1} = \mathfrak{o}_1 + \dots + \mathfrak{o}_{v-1}$ enthalten;

γ) jedes \mathfrak{o}_v hat mit \mathfrak{g}_{v-1} mindestens einen Punkt gemeinsam.

Zu jedem Punkt P von g gibt es nämlich mindestens eine ihn enthaltende Menge $\mathfrak{o}(\mathfrak{s})$. Da die $\mathfrak{o}(\mathfrak{s})$ nach 2. offene Mengen sind, gibt es nach dem Überdeckungssatz von Lindelöf unter ihnen eine höchstens

⁵⁾ In der hier erforderlichen Allgemeinheit und Exaktheit bewiesen von E. Lindelöf, a. a. O.

abzählbare Teilmenge, etwa $\mathfrak{D}_1, \mathfrak{D}_2, \dots$, die in ihrer Gesamtheit ebenfalls g enthalten, deren Summe also in dem vorliegenden Fall gleich g ist.

Ich setze nun $g_1 = v_1 = \mathfrak{D}_1$.

Bezüglich \mathfrak{D}_2 unterscheide ich drei Fälle:

Entweder ist \mathfrak{D}_2 in g_1 enthalten. Dann gehe ich zu \mathfrak{D}_3 über.

Oder \mathfrak{D}_2 ist nicht in g_1 enthalten, hat aber mit g_1 mindestens einen Punkt gemeinsam. Dann setze ich $v_2 = \mathfrak{D}_2$.

Oder \mathfrak{D}_2 hat mit g_1 keinen Punkt gemeinsam. Dann werde in g_1 ein Punkt P und in \mathfrak{D}_2 ein Punkt Q ausgewählt. Beide Punkte lassen sich in g durch einen einfachen Polygonzug verbinden. Dieser wird nach dem Borelschen Überdeckungssatz schon durch endlich viele \mathfrak{D}_ν überdeckt, etwa durch $\mathfrak{D}_{\mu_1}, \dots, \mathfrak{D}_{\mu_r}$, wobei unter diesen Mengen (ev. überflüssigerweise) auch \mathfrak{D}_2 sein soll. Diese endlich vielen Mengen können so geordnet werden, daß in der Folge

$$(5) \quad v_1, \mathfrak{D}_{\mu_1}, \dots, \mathfrak{D}_{\mu_r}$$

jede Menge mit der Summe der ihr vorhergehenden mindestens einen Punkt gemeinsam hat. Das sei in (5) schon geschehen. Dann streiche ich in (5), von links anfangend, jede Menge, die in der Summe der ihr vorhergehenden enthalten ist. Die übrigbleibenden Mengen werden mit

$$(6) \quad v_1, v_2, \dots, v_{k_2}$$

bezeichnet.

Nun kommt \mathfrak{D}_3 an die Reihe. Es sind wieder drei Fälle zu unterscheiden:

Entweder ist \mathfrak{D}_3 in der Summe g_{k_2} der bisher schon festgelegten Folge (6) enthalten. Dann gehe ich zu \mathfrak{D}_4 über.

Oder \mathfrak{D}_3 ist nicht in g_{k_2} enthalten, hat aber mit g_{k_2} mindestens einen Punkt gemeinsam. Dann setze ich $v_{k_2} = v_{k_2+1} = \mathfrak{D}_3$.

Oder \mathfrak{D}_3 hat mit g_{k_2} keinen Punkt gemeinsam. Dann wird in g_{k_2} ein Punkt P und in \mathfrak{D}_3 ein Punkt Q ausgewählt. Beide Punkte werden wieder durch einen in g verlaufenden einfachen Polygonzug verbunden. Der Borelsche Überdeckungssatz führt jetzt zu einer Folge

$$(7) \quad v_1, \dots, v_{k_2}, \mathfrak{D}_{\nu_1}, \dots, \mathfrak{D}_{\nu_l},$$

in der \mathfrak{D}_3 vorkommt und jede Menge mit der Summe der ihr vorhergehenden mindestens einen Punkt gemeinsam hat. In (7) wird nun wieder jede Menge gestrichen, die in der Summe der ihr vorhergehenden enthalten ist. Die übrig bleibenden Mengen seien mit

$$v_1, \dots, v_{k_2}, \dots, v_{k_3}$$

bezeichnet.

So kann man fortfahren und erhält ersichtlich eine höchstens abzählbare Folge mit den oben formulierten Eigenschaften.

5. Jedes g_n ist ein offenes Gebiet⁶⁾.

Da g_n als Summe von offenen Mengen offen ist, brauche ich nur zu zeigen, daß sich je zwei Punkte P und Q von g_n durch eine stetige Kurve in g_n verbinden lassen. Es sei R ein beliebiger Punkt von v_1 . Wenn ich zeigen kann, daß sich jeder der beiden Punkte P und Q durch eine in g_n verlaufende stetige Kurve mit R verbinden läßt, bin ich sicher fertig. Es genügt, eine solche Verbindung zwischen P und R nachzuweisen. P möge zu v_{κ_1} gehören. Dann hat v_{κ_1} mit g_{κ_1-1} , also mit einem v_{κ_2} ($1 \leq \kappa_2 < \kappa_1$) einen Punkt P_1 gemeinsam. Die Menge v_{κ_2} hat, falls $\kappa_2 > 1$ ist, mit einem v_{κ_3} ($1 \leq \kappa_3 < \kappa_2$) einen Punkt P_2 gemeinsam; usw. Nach höchstens κ_1 Schritten ist ein v_{κ_λ} erreicht, das mit v_1 einen Punkt P_λ gemeinsam hat. Nun läßt sich P nach 2. mit P_1 innerhalb v_{κ_1} , P_1 mit P_2 innerhalb v_{κ_2} , usw., P_λ mit R innerhalb v_1 verbinden, womit der angekündigte Beweis erbracht ist.

6. Es seien P und Q zwei senkrecht übereinander liegende Punkte von g_n , und es möge die Strecke \overline{PQ} ganz zu g gehören; dann gehört sie auch ganz zu g_n .

Es sei S (Fig. 2) ein beliebiger Punkt im Innern der Strecke \overline{PQ} . Ich muß zeigen, daß er auch zu g_n gehört. Wie in 5. läßt sich aus den v_x ($x \leq n$) eine Kette $v_{\kappa_1}, \dots, v_{\kappa_\lambda}$ bilden, so daß je zwei aufeinander folgende dieser Mengen einen Punkt gemeinsam haben, etwa die Punkte $R_1, \dots, R_{\lambda-1}$, und daß P zu v_{κ_1} , Q zu v_{κ_λ} gehört. Es möge dabei v_{κ_ν} durch die Strecke ξ_ν erzeugt sein. Von P kann man dann innerhalb g_n auf folgendem Wege zu Q gelangen: Von P auf der durch P gehenden Integralkurve bis ξ_1 und auf dieser Strecke bis zu der durch R_1 gehenden Integralkurve. Auf dieser bis ξ_2 und auf dieser Strecke bis zu der durch R_2 gehenden Integralkurve; usw. Schließlich auf ξ_λ bis zu der durch Q gehenden Integralkurve und auf dieser bis Q . Diese Integralkurvenstücke schneiden \overline{PQ} endlich oft. Liegt S in einem dieser Schnittpunkte oder auf einem der ξ_ν , so bin ich fertig. Trifft das nicht zu, so seien T und U die ersten Schnittpunkte oberhalb bzw. unterhalb S mit der vertikalen Geraden durch S .

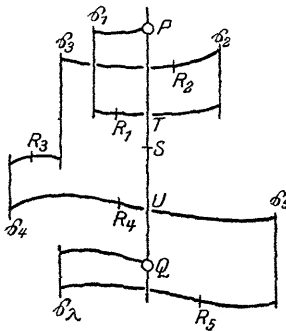


Fig. 2.

⁶⁾ Übrigens, wie aus 6. leicht folgt, sogar einfach zusammenhängend.

Man betrachte nun den zwischen T und U durchlaufenen Linienzug, der überdies zu einem doppelpunktsfreien schon vereinfacht gedacht werden kann (Fig. 3). Ergänzt man ihn durch die Strecke \overline{TU} , so entsteht ein einfacher geschlossener Linienzug, der aus vertikalen Strecken und Teilen von Integralkurven besteht. Durch ihn wird, weil g einfach zusammenhängend ist, ein ganz zu g gehöriger Bereich \mathfrak{h} begrenzt, dessen Rand im Innern von g liegt. Verfolgt man nun die bei S in \mathfrak{h} eintretende Integralkurve, so bemerkt man wie in 3., daß sie den Rand von \mathfrak{h} nochmals schneidet. Das kann aber wieder nur in einer vertikalen Strecke, also einem \mathfrak{s}_1 , geschehen; w. z. b. w.

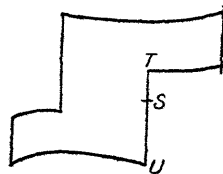


Fig. 3.

7. Für die Konstruktion von $\psi(x, y)$ ist es zweckmäßig, die Folge der \mathfrak{o}_n noch etwas zu verändern:

Es gibt eine Folge q_1, q_2, \dots mit folgenden Eigenschaften:

$\alpha)$ jedes $q_n = q(t_n)$ ist die Menge der Punkte, die auf den durch eine offene endliche vertikale Strecke t_n in g gehenden Integralkurven von (1) liegen;

$\beta)$ jedes $\mathfrak{h}_n = q_1 + \dots + q_n$ ist ein offenes Gebiet;

$\gamma)$ $q_1 + q_2 + \dots = g$;

$\delta)$ die gemeinsamen Punkte von t_n und \mathfrak{h}_{n-1} bilden genau eine offene Strecke und t_n ragt mit genau einem Ende über \mathfrak{h}_{n-1} hinaus.

Zunächst kann nämlich $\mathfrak{h}_1 = q_1 = \mathfrak{o}_1$ gesetzt werden. Ist \mathfrak{s}_2 die \mathfrak{o}_2 bestimmende vertikale Strecke, so folgt aus 6., daß der Durchschnitt von \mathfrak{s}_2 mit g_1 genau eine offene Strecke ist. Wird diese fortgelassen, so bleiben wegen 4. $\beta)$ eine oder zwei halboffene Strecken von \mathfrak{s}_2 übrig. Im ersten Fall wird $t_2 = \mathfrak{s}_2$ und $q_2 = \mathfrak{o}_2$ gesetzt. Im zweiten Fall wird jede der halboffenen Teilstrecken in das Innere von \mathfrak{s}_2 hinein etwas vergrößert, so daß aus ihnen zwei getrennte offene Strecken t_2 und t_3 entstehen. Es wird nun $q_2 = q(t_2)$ und $q_3 = q(t_3)$ gesetzt. In derselben Weise kann $\mathfrak{o}_3, \mathfrak{o}_4, \dots$ behandelt und eine Folge q_1, q_2, \dots aufgebaut werden, für die sich die Eigenschaften $\alpha), \gamma), \delta)$ dann unmittelbar aus der Konstruktion ergeben. Schließlich ergibt sich auch $\beta)$ so, wie in 5. für g , statt \mathfrak{h} , dargestellt ist.

§ 3.

Abschluß des Beweises.

1. Die Bezeichnungen von § 2, 7. werden beibehalten. Ferner soll x , die Abszisse von t , bezeichnen und

$$y = \varphi(x, \xi, \eta)$$

die durch den Punkt ξ, η von \mathfrak{G} gehende Integralkurve der Differentialgleichung (1). Wird $x = x_1$ festgehalten, so ist $\varphi(x_1, \xi, \eta)$ sicher in den Punkten ξ, η des Gebiets \mathfrak{h}_1 definiert. Wird nun

$$\psi(\xi, \eta) = \varphi(x_1, \xi, \eta)$$

in \mathfrak{h}_1 gesetzt, so hat $\psi(\xi, \eta)$ das Gebiet \mathfrak{h}_1 als genauen Definitionsbereich, ist für alle Punkte einer und derselben Integralkurve von (1) konstant und hat nach einem Satz von I. Bendixson²⁾ stetige partielle Ableitungen, und zwar ist

$$\psi'_\eta(\xi, \eta) = e^{\int_{\xi}^{x_1} f'_\varphi(x, \varphi(x, \xi, \eta)) dx}, \quad \text{also } > 0,$$

und

$$\psi'_\xi(\xi, \eta) + f(\xi, \eta) \psi'_\eta(\xi, \eta) \equiv 0;$$

d. h. $\psi(\xi, \eta)$ erfüllt für \mathfrak{h}_1 statt g die allein zu beweisenden drei Behauptungen a) bis c) des Hauptsatzes.

2. Um nun die Existenz eines $\psi(\xi, \eta)$ für das ganze Gebiet g zu beweisen, nehme ich an, es sei schon für \mathfrak{h}_{n-1} ein $\psi(\xi, \eta)$ konstruiert, so daß diese Funktion für \mathfrak{h}_{n-1} statt g die Behauptungen des Hauptsatzes erfüllt und überdies für $\nu < n$ in jedem $\mathfrak{h}_\nu - \mathfrak{h}_{\nu-1}$ eine lineare Funktion von $\varphi(x_\nu, \xi, \eta)$, etwa

$$\psi(\xi, \eta) = a_\nu + b_\nu \varphi(x_\nu, \xi, \eta)$$

mit $b_\nu > 0$ ist. Wegen § 2, 7. γ) bin ich mit dem Beweise fertig, wenn ich zeigen kann, daß es auch ein Zahlenpaar a_ν, b_ν gibt, so daß $b_\nu > 0$ und die durch

$$\psi(\xi, \eta) = a_\nu + b_\nu \varphi(x_\nu, \xi, \eta)$$

in $\mathfrak{h}_\nu - \mathfrak{h}_{\nu-1}$ für $\nu = 1, 2, \dots, n$ definierte Funktion die Eigenschaften a) bis c) des Hauptsatzes im Gebiet \mathfrak{h}_n aufweist.

3. Nach § 2, 7. δ) ragt t_n mit genau einem Ende, etwa dem oberen, über \mathfrak{h}_{n-1} hinaus. Die Endpunkte des halboffenen über \mathfrak{h}_{n-1} hinausragenden Teilintervalls mögen die Ordinaten V_n, W_n haben (Fig. 4). Da die Ordinate, auf welcher dieses Intervall liegt, schon durch den Index angedeutet ist, will ich dieses Intervall kurz mit $\langle V_n, W_n \rangle$ bezeichnen und in ähnlichen Fällen entsprechend verfahren. Das untere Stück von t_n gehört bis zum Punkt w_n exkl. ganz zu \mathfrak{h}_{n-1} . Dieses Stück hat nach § 2, 3. mit jedem der Gebiete q_1, \dots, q_{n-1} , wenn über-

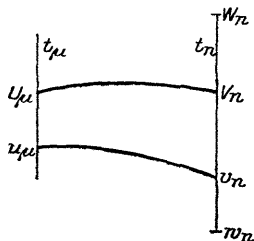


Fig. 4.

haupt Punkte, so genau ein offenes Intervall gemeinsam. Daher gibt es ein kleinstes μ ($1 \leq \mu < n$), so daß für ein $v_n < V_n$ das auf t_n liegende Inter-

vall (v_n, V_n) ganz zu q_μ und auch zu $h_\mu - h_{\mu-1}$ gehört. Die durch das Intervall (v_n, V_n) gehenden Integralkurven laufen also durch t_μ und schneiden aus dieser endlichen Strecke ein offenes Intervall (u_μ, U_μ) aus. Dabei ist U_μ innerer oder Randpunkt von g , also jedenfalls innerer Punkt von \mathcal{G} . Durch diesen Punkt gibt es daher eine Integralkurve

$$(8) \quad y = \varphi(x, x_\mu, U_\mu),$$

die sicher in einer gewissen Umgebung von x_μ existiert.

4. Aber die Kurve (8) existiert sogar bis $x = x_n$. Um die Ausdrucksweise zu erleichtern, werde etwa $x_\mu < x_n$ angenommen. Würde (8) nicht im ganzen abgeschlossenen Intervall $\langle x_\mu, x_n \rangle$ existieren, so würde (8), soweit sie in diesem Intervall existiert, entweder nicht beschränkt sein oder zwar beschränkt sein, jedoch beliebig nahe an den Rand von \mathcal{G} herankommen.

Da die Integralkurven stetig vom Anfangspunkt abhängen, würde es im ersten Fall auf den durch das Intervall (u_μ, U_μ) gehenden Integralkurven im Intervall $\langle x_\mu, x_n \rangle$ Punkte mit beliebig großen Ordinaten geben. Dann wären aber die Ableitungen dieser Kurvenfunktionen $\varphi(x)$ nicht gleichmäßig beschränkt in g , und das ist, da jedes dieser $\varphi(x)$ ein Integral von (1) ist, ein Widerspruch gegen die für g vorausgesetzte Beschränktheit von $f(x, y)$.

Daß auch der zweite Fall nicht zutreffen kann, erkennt man, wenn man die Kurve (8), soweit sie in $\langle x_\mu, x_n \rangle$ existiert, in ein abgeschlossenes Quadrat Ω einschließt, dessen Seite von der Kurve etwa einen Abstand > 1 haben mögen. Die Ränder von \mathcal{G} und g , soweit sie in Ω liegen, seien mit \mathfrak{R} und r bezeichnet. Diese beiden Mengen sind abgeschlossen und haben, da g total in \mathcal{G} liegt, keinen Punkt gemeinsam, also, da sie beschränkt sind, einen positiven Abstand ϱ voneinander. Da die Kurve (8) ganz in Ω liegt und beliebig nahe an den Rand von \mathcal{G} kommt, hat (8) von \mathfrak{R} den Abstand 0. Wegen der stetigen Abhängigkeit der Integralkurven vom Anfangspunkt gibt es daher eine durch das Intervall (u_μ, U_μ) gehende Integralkurve, die für $x_\mu \leq x \leq x_n$ ganz in Ω liegt und von \mathfrak{R} einen Abstand $< \frac{1}{2}\varrho$ hat. Dann würde aber auch r von \mathfrak{R} einen Abstand $< \frac{1}{2}\varrho$ haben, womit der gewünschte Widerspruch erzielt ist.

5. Das halboffene Intervall (u_μ, U_μ) wird nun nach oben zu einer offenen Strecke \mathfrak{s} erweitert, die noch ganz zu \mathcal{G} gehört. Dann existiert $\varphi(x_\mu, \xi, \eta)$ u. a. in $\mathfrak{o} = \mathfrak{o}(\mathfrak{s})$ und hat dort stetige partielle Ableitungen nach ξ und η . Da die Kurve (8) zu \mathfrak{o} gehört, existieren diese Ableitungen insbesondere in den Punkten der Kurve (8). Ich kann daher definieren:

$$(9) \quad b_n = b_\mu \cdot \varphi'_\eta(x_\mu, x_n, V_n),$$

$$(10) \quad a_n = a_\mu + b_\mu U_\mu - b_n V_n,$$

$$(11) \quad \psi(\xi, \eta) = a_n + b_n \varphi(x_n, \xi, \eta) \text{ in } h_n - h_{n-1}.$$

Damit ist nun ψ in \mathfrak{h}_n eindeutig definiert und auch in $\mathfrak{h}_n - \mathfrak{h}_{n-1}$ eine lineare Funktion von $\varphi(x_n, \xi, \eta)$, wobei auch $b_n > 0$ ist, weil b_μ wie φ'_η positiv ist. Ferner ergibt sich aus (11), daß $\psi(\xi, \eta)$ auch für alle Punkte einer Integralkurve denselben Wert hat. Der Induktionsbeweis und damit der Beweis des Hauptsatzes ist daher beendet, sobald noch die Behauptung c) des Hauptsatzes für \mathfrak{h}_n statt \mathfrak{g} bewiesen ist.

Diese Behauptung trifft teils nach meiner Annahme, teils wegen (11) sicher zu für \mathfrak{h}_{n-1} und für die inneren Punkte von $\mathfrak{h}_n - \mathfrak{h}_{n-1}$. Die zu $\mathfrak{h}_n - \mathfrak{h}_{n-1}$ gehörigen Randpunkte von \mathfrak{h}_{n-1} werden von der Kurve (8) gebildet, soweit sie in \mathfrak{g} liegt. Nur mit diesen Punkten habe ich mich nun noch zu beschäftigen.

Ich betrachte dazu die Funktionen

$$(12) \quad a_\mu + b_\mu \varphi(x_\mu, \xi, \eta) \quad \text{in } \mathfrak{o},$$

$$(13) \quad a_n + b_n \varphi(x_n, \xi, \eta) \quad \text{in } \mathfrak{h}_n.$$

Jede ist in einem Gebiete definiert, das die Kurve (8) enthält, und besitzt in diesem Gebiet stetige partielle Ableitungen nach ξ und η . Auf der Kurve (8) haben die Funktionen die Werte

$$a_\mu + b_\mu U_\mu \quad \text{bzw.} \quad a_n + b_n V_n,$$

die wegen (10) übereinstimmen. Da $\psi(\xi, \eta)$ oberhalb der Kurve (8) mit der Funktion (13) und unterhalb mit der Funktion (12) übereinstimmt (vorausgesetzt, daß man sich nicht zu weit von der Kurve entfernt), ist die Existenz und zugleich die Stetigkeit von $\psi'_\eta(\xi, \eta)$ auch in den Punkten der Kurve (8) bewiesen, wenn ich zeigen kann, daß die Ableitungen der Funktionen (12) und (13) nach η in den Punkten von (8) übereinstimmen.

In der Tat ist für irgendeinen Punkt ξ, η in \mathfrak{h}_n auf der Kurve (8), weil auch der Punkt x_n, V_n auf dieser Kurve liegt und wegen (9)

$$\begin{aligned} \frac{\partial}{\partial \eta} (a_\mu + b_\mu \varphi(x_\mu, \xi, \eta)) &= b_\mu e^{\int_{\xi}^{x_\mu} f'_\varphi(x, \varphi(x, \xi, \eta)) dx} \\ &= b_\mu \cdot \varphi'_\eta(x_\mu, x_n, V_n) \frac{e^{\int_{\xi}^{x_\mu} f'_\varphi(x, \varphi(x, \xi, \eta)) dx}}{e^{\int_{x_n}^{x_\mu} f'_\varphi(x, \varphi(x, x_n, V_n)) dx}} \\ &= b_n e^{\int_{\xi}^{x_n} f'_\varphi(x, \varphi(x, \xi, \eta)) dx}, \end{aligned}$$

also

$$(14) \quad \frac{\partial}{\partial \eta} (a_\mu + b_\mu \varphi(x_\mu, \xi, \eta)) = \frac{\partial}{\partial \eta} (a_n + b_n \varphi(x_n, \xi, \eta))$$

für die Punkte ξ, η von (8). Wegen $b_n > 0$ ist also insbesondere auch (2) erfüllt.

Endlich ist bekanntlich

$$(15) \quad \begin{cases} \frac{\partial}{\partial \xi}(a_\mu + b_\mu \varphi(x_\mu, \xi, \eta)) = -b_\mu f(\xi, \eta) \varphi'_\eta(x_\mu, \xi, \eta), \\ \frac{\partial}{\partial \xi}(a_\mu + b_\mu \varphi(x_n, \xi, \eta)) = -b_n f(\xi, \eta) \varphi'_\eta(x_n, \xi, \eta). \end{cases}$$

Da auf (8) beide Ausdrücke wegen (14) denselben Wert haben, existieren auch die partiellen Ableitungen nach ξ in ganz \mathfrak{h}_n und sind stetig. Schließlich folgt aus jeder der Gleichungen (15) die Gültigkeit von (3) auch auf (8), also in ganz \mathfrak{h}_n .

§ 4.

Zusätze.

1. Verzichtet man auf die Differenzierbarkeit von $f(x, y)$, so gilt noch folgender

Satz 2. *In dem offenen Gebiete \mathfrak{G} möge $f(x, y)$ stetig sein und die Lipschitzbedingung erfüllen. Dann gibt es für jedes offene und einfach zusammenhängende Gebiet g , das total in \mathfrak{G} liegt und in dem $f(x, y)$ beschränkt ist, eine Funktion $\psi(x, y)$ mit den Eigenschaften a) und b) des Hauptsatzes und*

c*) $\psi(x, y)$ ist stetig in g , und es ist

$$(2^*) \quad \psi(x, y_1) < \psi(x, y_2)$$

für je zwei Punkte x, y_1 und x, y_2 , wenn die ganze Strecke mit diesen Endpunkten zu g gehört und $y_1 < y_2$ ist.

Für den Beweis braucht man nur zu beachten, daß § 2 auch im vorliegenden Falle gilt. Darauf ist schon am Anfang von § 2 hingewiesen. Dann ergibt sich aber der Satz 2 durch die in § 3 angestellten Überlegungen, wenn man dort alles, was sich auf die Ableitungen von ψ bezieht, fortläßt, $b_n = 1$ setzt und Gleichung (9) fortläßt.

Aus dem Satz 2 ergeben sich weiter mit völlig unverändertem Wortlaut die Behauptungen d) bis f) des § 1 auch für den vorliegenden Fall.

2. Der Hauptsatz liefert mit der Relation (3) natürlich die Existenz eines Integrals der partiellen Differentialgleichung^{?)}

$$(16) \quad \frac{\partial z}{\partial x} + f(x, y) \frac{\partial z}{\partial y} = 0,$$

^{?)} Als Integral der partiellen Differentialgleichung für ein Gebiet g wird hier, wie üblich, eine Funktion bezeichnet, die in g nicht nur die Differentialgleichung erfüllt, sondern auch stetige partielle Ableitungen hat.

und zwar in einem von vornherein angebbaren Gebiet (nämlich in jedem einfach zusammenhängenden Gebiet, das total in \mathfrak{G} liegt und in dem $f(x, y)$ beschränkt ist), während diese Existenz bisher nur für ein $\mathfrak{o}(\beta)$ bewiesen war²⁾.

Zugleich ist auch eine völlige Übersicht über die sämtlichen Integrale von (16) vorhanden. Da die üblichen Darstellungen auch in diesem Punkte unbefriedigend sind, benutze ich die Gelegenheit, die Verhältnisse kurz klarzustellen.

Als *Hauptintegral* der partiellen Differentialgleichung (16) für das Gebiet g wird ein Integral bezeichnet, das in keinem Teilgebiet von g konstant ist. Der Hauptsatz liefert in der Funktion $\psi(x, y)$ wegen (2) ein solches Hauptintegral. Und weiter besteht der

Satz 3. *Unter den Voraussetzungen des Hauptsatzes ist die Gesamtheit der Integrale von (16) für das Gebiet g genau die Menge der von einem Hauptintegral abhängigen Funktionen.*

Dabei ist dem Begriff der Abhängigkeit von Funktionen die von den Herren K. Knopp und R. Schmidt⁸⁾ gegebene Definition unterzulegen.

Daß je zwei Integrale $\psi(x, y)$ und $\chi(x, y)$ von (16) abhängig sind, ist klar, da aus den mit diesen Funktionen aufgeschriebenen Differentialgleichungen (16) sofort folgt, daß ihre Funktionaldeterminante verschwindet. Nach einem zuerst von den Herren Knopp und R. Schmidt bewiesenen Satz über Funktionaldeterminanten und Abhängigkeit von Funktionen⁸⁾ sind ψ und χ daher in g abhängig.

Es sei nun umgekehrt $\chi(x, y)$ in g eine von einem Hauptintegral $\psi(x, y)$ abhängige Funktion. Nach dem eben zitierten Satz ist dann

$$\begin{vmatrix} \psi'_x & \psi'_y \\ \chi'_x & \chi'_y \end{vmatrix} \equiv 0 \quad \text{in } g.$$

Für jeden Punkt x, y von g sind daher die Gleichungen

$$(17) \quad \begin{cases} \lambda_1(x, y)\psi'_x + \lambda_2(x, y)\chi'_x = 0 \\ \lambda_1(x, y)\psi'_y + \lambda_2(x, y)\chi'_y = 0 \end{cases}$$

durch Zahlen λ_1, λ_2 lösbar, die in keinem Punkte beide gleichzeitig verschwinden. Addiert man die zweite der Gleichungen (17) nach Multiplikation mit $f(x, y)$ zur ersten und berücksichtigt, daß ψ ein Integral von (16) ist, so ergibt sich

$$(18) \quad \lambda_2(x, y)\{\chi'_x + f(x, y)\chi'_y\} \equiv 0 \quad \text{in } g.$$

⁸⁾ Math. Zeitschr. 25 (1926), S. 373—381. Es sei darauf hingewiesen, daß die Ausdehnung der Ergebnisse dieser Arbeit auf p Funktionen von q Variablen ($q > p$) bisher noch nicht gelungen ist.

Ich habe zu zeigen, daß auch χ ein Integral von (16) ist, d. h. daß in (18) der zweite Faktor in g identisch verschwindet. Wäre dies in irgendeinem Punkt von g nicht der Fall, so gäbe es wegen der Stetigkeit von χ'_x , χ'_y und f einen Kreis \mathfrak{f} in g , so daß

$$\chi'_x + f(x, y)\chi'_y \neq 0 \quad \text{in } \mathfrak{f}$$

wäre. Wegen (18) wäre dann $\lambda_2(x, y) \equiv 0$ in \mathfrak{f} , also $\lambda_1(x, y) \neq 0$ in dem ganzen Kreis \mathfrak{f} . Dann würde aber aus (17) folgen

$$\psi'_x \equiv \psi'_y \equiv 0 \quad \text{in } \mathfrak{f},$$

also wäre ψ gegen die Voraussetzung kein Hauptintegral.

3. Endlich liefert der Hauptsatz, insbesondere die Relation (3) folgendes Ergebnis über die Existenz eines Multiplikators:

Satz 4. Unter den Voraussetzungen des Hauptsatzes besitzt die Differentialgleichung

$$y' = f(x, y)$$

in g einen Multiplikator, nämlich z. B. $\psi'_y(x, y)$.

Das Neue an diesem Resultat ist wiederum, daß das Gebiet, in dem der Multiplikator existiert, von vornherein angebar ist.

Tübingen, 10. 7. 1927.

(Eingegangen 11. 7. 1927.)

Zum Rechnen mit Potenzreihen.

Von

R. Mehmke in Stuttgart.

Die drei Aufgaben, das Produkt und den Quotienten zweier Potenzreihen zu berechnen wie auch eine Potenzreihe umzukehren, werden in den Lehrbüchern als schwierig hingestellt, was die vollständige Durchführung betrifft¹⁾. Deshalb erscheint es nicht überflüssig, Wege zu zeigen, auf denen man jene Aufgaben leicht und sicher lösen kann. Die Verfahren, die ich mitteilen werde, leisten, glaube ich, in bezug auf Ersparnis an Zeit und Mühe das Wünschenswerte und Mögliche. Gleiches gilt von der Aufgabe, die als vierte behandelt werden soll, nämlich zu untersuchen, ob zwischen mehreren Funktionen, die als Potenzreihen dargestellt sind, eine lineare Abhängigkeit besteht, und wenn es der Fall ist, auch die Faktoren, die dabei auftreten, zu ermitteln. Um nicht zu weitläufig zu werden, sehe ich von Betrachtungen über die Konvergenz der vorkommenden unendlichen Potenzreihen ab. Ich darf das um so eher, als die einschlägigen Lehrbücher diesen Punkt genügend behandeln.

1. Produkt zweier Potenzreihen.

Verlangt seien die Größen c_0, c_1, c_2, \dots , wenn $c_0 + c_1 x + c_2 x^2 + \dots = (a_0 + a_1 x + a_2 x^2 + \dots)(b_0 + b_1 x + b_2 x^2 + \dots)$ ist und $a_0, a_1, a_2, \dots; b_0, b_1, b_2, \dots$ gegeben sind. Im Fall unendlicher Potenzreihen soll, wie wir voraussetzen, das gewöhnlich nach Cauchy benannte Produkt²⁾ genommen werden, für welches

$$c_n = a_0 b_n + a_1 b_{n-1} + \dots + a_{n-1} b_1 + a_n b_0$$

¹⁾ So sagt bezüglich der dritten Aufgabe K. Knopp in seinem Buche „Theorie und Anwendung der unendlichen Reihen“, 2. Aufl., Berlin 1924, S. 186: „Die tatsächliche Herstellung der Reihe $(x=)y + b_2 y^2 + \dots$ aus der Reihe $(y=)x + a_2 x^2 + \dots$ ist ... meist mit erheblichen Schwierigkeiten verknüpft“, und ähnlich urteilt er über die ersten beiden Aufgaben.

²⁾ Es entspricht bei dezimal geschriebenen Zahlen dem „blitzbildenden“ Produkt der Inder, vgl. M. Cantor, Vorlesungen über Geschichte der Mathematik, I, 2. Aufl., Leipzig 1894, S. 571. Stelle bei A. L. Cauchy: *Analyse algébrique*, Paris 1821, p. 156, oder in der deutschen Übersetzung von Huzler, Königsberg 1828, S. 116.

ist. Um die c der Reihe nach zu finden, wendet man am besten J. Fouriers Schiebverfahren an³⁾. Man schreibt auf den unteren Rand eines Schiebers von Papier, Pergament oder dergl. die Konstanten des einen Faktors, z. B. b_0, b_1, b_2, \dots , jedoch von rechts nach links, wenn die Konstanten a_0, a_1, a_2, \dots des andern Faktors von links nach rechts geschrieben worden sind, hält den Schieber zuerst mit b_0 über a_0 , was $c_0 = a_0 b_0$ ergibt, hierauf mit b_0 über a_1 , so daß b_1 über a_0 zu stehen kommt und man durch Addieren der beiden Produkte der jetzt übereinander stehenden Konstanten erhält: $c_1 = a_0 b_1 + a_1 b_0$, usw., d. h. bei jedem folgenden Schritt hat man den Schieber um eine Stelle nach rechts zu rücken, und der Vorteil besteht eben darin, daß die Faktoren der Teilprodukte, aus denen das gerade an der Reihe befindliche c sich zusammensetzt, senkrecht übereinander stehen.

Schieber: $\boxed{\dots b_2 \quad b_1 \quad b_0}$

1. Stellung:

$\boxed{\dots b_2 \quad b_1 \quad b_0}$

$a_0 \quad a_1 \quad \dots,$

2. Stellung:

$\boxed{\dots b_2 \quad b_1 \quad b_0}$

$a_0 \quad a_1 \quad a_2 \quad \dots,$

3. Stellung:

$\boxed{\dots b_2 \quad b_1 \quad b_0}$

$a_0 \quad a_1 \quad a_2 \quad a_3 \quad \dots$

Will man von $(a_0 + a_1 x + a_2 x^2 + \dots)$ das Quadrat und die höheren Potenzen mit ganzzahligen Exponenten erhalten, so wird man $b_i = a_i$ setzen und das Verfahren die nötige Anzahl von malen wiederholen. (Beispiele in der nächsten Nummer und in Nr. 6.)

2. Anwendung: Einsetzen einer Potenzreihe in eine andere.

Wenn in den Lehrbüchern verlangt wird, für

$$u = b_0 + b_1 y + b_2 y^2 + \dots$$

eine Potenzreihe in x herzustellen, falls

$$y = a_0 + a_1 x + a_2 x^2 + \dots$$

ist, so wird offenbar angenommen, daß man die Reihen für y^2, y^3, \dots aus derjenigen für y mit Hilfe des binomischen Satzes herleitet. Soll nun z. B.

$$u = \sqrt[3]{8 + x \lg(1 + x)}$$

³⁾ J. Fourier, *Analyse des équations déterminées*, Paris 1831, p. 190, oder in der von A. Loewy besorgten deutschen Ausgabe Ostwalds Klassiker Nr. 127, Leipzig 1902, S. 183. Bei Fourier ist zwar nur von dem Produkt zweier dezimal geschriebenen Zahlen, also Potenzreihen mit $x = 10$, die Rede, die Ausdehnung auf beliebige Potenzreihen lag jedoch äußerst nahe und ich habe, nachdem ich sie schon in den achtziger Jahren des letzten Jahrhunderts bei zahlentheoretischen Rechnungen mit großem Nutzen angewendet hatte, in der *Encyclopädie der mathem. Wiss.* 1, 2, Leipzig 1900–1904, S. 942, Anm. ⁶⁾, empfehlend auf sie hingewiesen, was aber nicht beachtet worden zu sein scheint.

in eine Potenzreihe nach x entwickelt werden⁴⁾, und setzt man

$$u = 2(1 + y)^{\frac{1}{3}}$$

mit

$$y = \frac{1}{8} x \lg(1 + x),$$

dann ist zwar beim Anschreiben der Potenzreihe in y für u der binomische Satz nicht zu entbehren, dagegen ist es beim Ausrechnen von y^2, y^3, \dots bequemer, der Schlußbemerkung von Nr. 1 entsprechend zu verfahren, d. h. wiederholt mit der Reihe für y zu multiplizieren. Man wird hierbei die Rechnung etwa so anordnen:

$$u = 2 + \frac{2}{3}y - \frac{2}{9}y^2 + \frac{10}{81}y^3 - \dots,$$

$$y = \frac{1}{8}x^2 - \frac{1}{16}x^3 + \frac{1}{24}x^4 - \frac{1}{32}x^5 + \frac{1}{40}x^6 + \dots,$$

also Schieber:

...	$\frac{1}{40}$	$-\frac{1}{32}$	$\frac{1}{24}$	$-\frac{1}{16}$	$\frac{1}{8}$
-----	----------------	-----------------	----------------	-----------------	---------------

Faktoren		x^2	x^3	x^4	x^5	x^6	...
$\frac{2}{3}$	y	$\frac{1}{8}$	$-\frac{1}{16}$	$\frac{1}{24}$	$-\frac{1}{32}$	$\frac{1}{40}$...
$-\frac{2}{9}$	y^2			$\frac{1}{64}$	$-\frac{1}{64}$	$\frac{11}{768}$...
$\frac{10}{81}$	y^3					$\frac{1}{512}$...
...

Ergebnis:

$$u = 2 + \frac{1}{12}x^2 - \frac{1}{24}x^3 + \frac{7}{288}x^4 - \frac{5}{288}x^5 + \frac{1423}{103680}x^6 + \dots$$

3. Quotient zweier Potenzreihen.

Auch das Verfahren, den Quotienten zweier Potenzreihen zu berechnen, das ich für das zweckmäßigste halte, ist nicht gerade neu, sondern 1896 von C. Reuschle angegeben⁵⁾, aber von den Verfassern der einschlägigen

⁴⁾ Aufgabe aus: H. v. Mangoldt, Einführung in die höhere Mathematik 2, Leipzig 1912, S. 239.

⁵⁾ C. Reuschle, Abgekürzte algebraische Division bei quadratischem und höherem Divisor, Zeitschr. Math. Phys. 41 (1896), S. 93. Von Reuschles abgekürzter Division, die er selbst als Verallgemeinerung des bekannten (Ruffini-)Horner'schen Verfahrens bei linearem Divisor betrachtet, bildet Fouriers geordnete Division (s. hier Anm. 3) eine Vorstufe; dem Divisor zweiten oder höheren Grades bei Reuschle entspricht bei

Lehrbücher noch nicht berücksichtigt worden⁶⁾. Seien gegeben die Konstanten a_0, a_1, a_2, \dots und b_0, b_1, b_2, \dots in

$$\begin{aligned} f(x) &= a_0 + a_1 x + a_2 x^2 + \dots, \\ g(x) &= b_0 + b_1 x + b_2 x^2 + \dots, \end{aligned}$$

gesucht c_0, c_1, c_2, \dots gemäß der Entwicklung

$$\frac{f(x)}{g(x)} = c_0 + c_1 x + c_2 x^2 + \dots$$

Es empfiehlt sich, dem Nenner des Bruches die Form

$$\varphi(x) = 1 - \beta_1 x - \beta_2 x^2 - \dots$$

zu geben, dadurch, daß man etwa

$$\begin{aligned} \frac{f(x)}{b_0} &= a'_0 + a'_1 x + a'_2 x^2 + \dots, \\ \frac{b_i}{b_0} &= -\beta_i \quad (i = 1, 2, \dots) \end{aligned}$$

setzt. Alsdann wird

$$c_0 = a'_0, \quad c_1 = a'_1 + \beta_1 c_0, \quad c_2 = a'_2 + \beta_1 c_1 + \beta_2 c_0, \quad \dots$$

Wir können deshalb mit Reuschle die Konstanten β_1, β_2, \dots von rechts nach links auf den unteren Rand eines Schiebers schreiben und die c der Reihe nach so ausrechnen: Nachdem wir $c_0 = a'_0$ unter a'_0 wiederholt haben, legen wir den Schieber so auf, daß β_1 über c_0 zu stehen kommt — a'_0 wird hierbei verdeckt —, bilden das Produkt $\beta_1 c_0$, addieren es zu a'_1 und schreiben es für c_1 unter a'_1 . Dann rücken wir den Schieber um eine Stelle nach rechts, wodurch a'_1 und a'_2 verdeckt werden, multiplizieren die jetzt übereinander stehenden Größen β und c miteinander, bilden also $\beta_1 c_1$ und $\beta_2 c_0$, addieren beide Produkte zu a'_2 und schreiben die Summe als c_2 unter a'_2 , usw. Allgemein: Hat man dem Schieber eine solche Lage gegeben, daß von den daraufstehenden Größen β die äußerste rechts, also β_1 , über die zuletzt berechnete Größe c , sagen wir c_{n-1} , zu stehen kommt, wobei a_{n-1}, a_{n-2}, \dots durch den Schieber verdeckt werden, so braucht man bloß ein jedes der schon vorhandenen c mit dem darüber stehenden β zu multiplizieren und die erhaltenen Produkte zu a_n zu addieren, um c_n zu finden.

Fourier der „bezeichnete“ Divisor, wenn er mehr als eine Ziffer bekommen hat, nur daß eben $x = 10$ und nicht beliebig ist. Zugleich macht Fourier bereits, wie Reuschle, vom Verschieben des auf ein besonderes Blatt geschriebenen Divisors Gebrauch. (Von Florian Cajori ist im Bull. Americ. Mathem. Soc. 17 (1911), S. 409 nachgewiesen worden, daß Paolo Ruffini schon 1804, 15 Jahre vor Horner, das bei uns noch immer nach letzterem allein benannte Divisions- und Annäherungsverfahren veröffentlicht hat.)

⁶⁾ Von dem Fall, daß der Divisor eine unendliche Potenzreihe ist, spricht Reuschle allerdings nicht. Sein Verfahren liefert übrigens bei im Endlichen abbrechenden Potenzreihen außer dem in einer ganzen Funktion bestehenden Quotienten auch den Rest.

4. Beispiel: Bernoullische Polynome.

Als das Bernoullische Polynom n -ten Grades der Veränderlichen x wird von manchen⁷⁾ der mit $n!$ multiplizierte Koeffizient von t^n in der Entwicklung von

$$\frac{t e^{xt}}{e^t - 1}$$

genannt, so daß

$$\frac{t e^{xt}}{e^t - 1} = \sum_{n=0}^{\infty} \frac{B_n(x)}{n!} t^n$$

ist. Für $x=0$ ergeben sich die Bernoullischen Zahlen. (Der Übereinstimmung mit der unten angeführten Quelle zuliebe ist die Bezeichnung gegen die vorher benützte geändert, nämlich t statt x geschrieben worden, während x eine neue Veränderliche bezeichnet.) Wir können die „erzeugende Funktion“ auf der linken Seite schreiben:

$$\frac{1 + xt + \frac{x^2}{2}t^2 + \frac{x^3}{6}t^3 + \frac{x^4}{24}t^4 + \dots}{1 + \frac{1}{2}t + \frac{1}{6}t^2 + \frac{1}{24}t^3 + \frac{1}{120}t^4 + \dots} = \frac{f(t)}{g(t)},$$

also ist, wenn entsprechend Nr. 3

$$f(t) = a_0 + a_1 t + a_2 t^2 + \dots, \quad g(t) = \varphi(t) = 1 - \beta_1 t - \beta_2 t^2 - \dots$$

gesetzt wird,

$$a_0 = 1, \quad a_1 = x, \quad a_2 = \frac{1}{2}x^2, \quad \dots; \quad a'_i = a_i; \quad \beta_1 = -\frac{1}{2}, \quad \beta_2 = -\frac{1}{6}, \quad \dots,$$

und die Rechnung gestaltet sich wie folgt:

Schieber: $\dots, -\frac{1}{120}, -\frac{1}{24}, -\frac{1}{6}, -\frac{1}{2}$

1. Stellung:

$-\frac{1}{6}, -\frac{1}{2}$	x	$\frac{x^2}{2}$
	1	x
		$-\frac{1}{2}$

2. Stellung:

$-\frac{1}{24}, -\frac{1}{6}, -\frac{1}{2}$	$\frac{x^2}{2}$	$\frac{x^3}{6}$
	1	x
		$\frac{x^2}{2}$
		$-\frac{x}{2}$
		$+\frac{1}{12}$

⁷⁾ Siehe z. B. R. Courant-D. Hilbert, Methoden der mathem. Physik, I, Berlin 1924, S. 85.

3. Stellung:

...	$-\frac{1}{120}$	$-\frac{1}{24}$	$-\frac{1}{6}$	$-\frac{1}{2}$	$\frac{x^3}{6}$	$\frac{x^4}{24}$
	1	x	$\frac{x^2}{2}$	$\frac{x^3}{6}$	$\frac{x^2}{4}$	$\frac{x}{12}$
		$-\frac{1}{2}$	$-\frac{x}{2}$	$-\frac{x^2}{4}$	$+\frac{x}{12}$	
			$+\frac{1}{12}$	$+\frac{x}{12}$		

Mithin erhält man:

$$B_0(x) = 1, \quad B_1(x) = x - \frac{1}{2}, \quad B_2(x) = x^2 - x + \frac{1}{6},$$

$$B_3(x) = x^3 - \frac{3}{2}x^2 + \frac{1}{2}x, \quad \dots$$

Diese Polynome werden sich kaum auf irgendeine andere Weise ebenso schnell und bequem finden lassen.

5. Umkehrung einer Potenzreihe.

Die nun vorliegende Aufgabe ist bekanntlich, wenn in der Reihe

$$y = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \dots$$

die Größen a_0, a_1, a_2, \dots gegeben sind, x in eine Potenzreihe von y zu entwickeln, d. h. in der Gleichung

$$x = b_0 + b_1 y + b_2 y^2 + b_3 y^3 + \dots$$

die Größen b_0, b_1, b_2, \dots zu berechnen. Man pflegt $a_0 = 0, a_1 = 1$ zu setzen, was $b_0 = 0, b_1 = 1$ zur Folge hat. Durch eine lineare Transformation läßt sich zwar dieser Fall immer herbeiführen, wofern a_1 nicht Null ist, es besteht aber keine Notwendigkeit für diese Umformung. Wohl sind für die b von b_1 bis b_{13} die allgemeinen Ausdrücke bekannt⁸⁾, aber es ist nicht zu empfehlen, daß man die gegebenen Werte der a , mögen diese in Buchstabenausdrücken oder in Zahlen bestehen, in fertige Formeln einsetzt; man rechnet besser von Fall zu Fall die b nach einem geeigneten Verfahren aus. Einem solchen, das ich für besonders zweckmäßig halte, liegt folgender Gedanke zugrunde. Bildet man die Potenzreihen für y^2, y^3, \dots nach Nr. 1, dann ergeben sich lauter Gleichungen, die in bezug auf x, x^2, x^3, \dots linear sind. Eliminiert man also aus den n ersten dieser Gleichungen x^2, x^3, \dots, x^n und vernachlässigt man die Glieder mit x^{n+1} und den höheren Potenzen von x , dann erscheint für x ein in y, y^2, y^3, \dots, y^n

⁸⁾ C. E. van Orstrand, Reversion of power series, Phil. Mag. 19 (1910), p. 366.

linearer Ausdruck, also das bis zum Glied mit y^n reichende Stück der gesuchten Potenzreihe. Auf dieselbe Weise erhält man eine Potenzreihe für x^r , wenn in der gegebenen die Glieder mit x, x^2, \dots, x^{r-1} fehlen. Man kommt schneller vorwärts, wenn man, statt wie üblich, die zu eliminierenden Größen eine nach der andern vorzunehmen, sie paarweise aus den Gleichungen entfernt, wofür ich „beschleunigtes Eliminieren“ sage.

6. Beispiel.

Die gegebene Reihe heiße

$$y = x - \frac{1}{3}x^3 + \frac{1}{5}x^5 - \frac{1}{7}x^7 + \dots$$

Da y eine ungerade Funktion von x ist, muß auch x eine solche von y sein. Wir bilden zuerst nach Nr. 1 die Reihe für y und erhalten:

$$y^3 = x^3 - \frac{2}{3}x^4 + \frac{23}{45}x^6 - \frac{44}{105}x^8 + \dots$$

Diese Reihe mit der ersten multipliziert — wieder nach Nr. 1 — gibt:

$$y^3 = x^3 - x^5 + \frac{14}{15}x^7 + \dots$$

Hieraus erhält man durch wiederholtes Multiplizieren mit der Reihe für y^2 diejenigen für y^5, y^7, \dots , wobei die jetzt auf dem Schieber stehenden Zahlen, die von y^2 herrühren, unverändert bleiben. Nun bilden wir aus den gefundenen Koeffizienten eine Matrix, von der die Zeilen sich auf y, y^3, y^5, \dots beziehen, die Spalten auf x, x^3, x^5, \dots .

	x	x^3	x^5	x^7	y	y^3	y^5	y^7
1	1	$-\frac{1}{3}$	$\frac{1}{5}$	$-\frac{1}{7}$	1			
$\frac{1}{3}$		1	-1	$\frac{14}{15}$		1		
$\frac{2}{15}$			1	$-\frac{5}{3}$			1	
				1				1

Will man aus den ersten drei Zeilen die Elemente, die in den Spalten für x^3 und x^5 stehen, auf einmal entfernen, so muß man sie der Reihe nach mit den Minoren zweiter Ordnung

$$\begin{vmatrix} 1 & -1 \\ 0 & 1 \end{vmatrix} = 1, \quad \begin{vmatrix} 0 & 1 \\ -\frac{1}{3} & \frac{1}{5} \end{vmatrix} = \frac{1}{3}, \quad \begin{vmatrix} -\frac{1}{3} & \frac{1}{5} \\ 1 & -1 \end{vmatrix} = \frac{2}{15}$$

multiplizieren und hierauf addieren, was die neue Zeile gibt:

	x	x^3	x^5	x^7	y	y^3	y^5	y^7
	1	0	0	$-\frac{17}{315}$	1	$\frac{1}{3}$	$\frac{2}{15}$	

Es ist also

$$x - \frac{17}{315}x^7 + \dots = y + \frac{1}{3}y^3 + \frac{2}{15}y^5$$

oder

$$x = y + \frac{1}{3}y^3 + \frac{2}{15}y^5 + \frac{17}{315}x^7 + \dots$$

Damit sind von der gesuchten Potenzreihe für x die ersten drei Glieder gefunden und man sieht zugleich, daß wenn man hier abbricht, der begangene Fehler ungefähr

$$+ \frac{17}{315}x^7$$

betragen wird. Wollte man weiter rechnen, so könnte man zu der neuen Zeile diejenigen für y^7 und y^9 hinzunehmen und aus diesen drei Zeilen die Glieder mit x^7 und x^9 auf einmal entfernen, wieder durch Multiplikation mit den betreffenden Minoren und Addition. Man erhält also immer zwei weitere Glieder der verlangten Potenzreihe auf einmal.

Man könnte ja auch drei oder mehr Glieder auf einmal berechnen, aber zu empfehlen ist es nicht, weil schon im Falle dreier Glieder vier Minoren dritter Ordnung ausgerechnet werden müßten, was nicht sonderlich bequem ist.

7. Lineare Abhängigkeit von Potenzreihen.

Die bekannten Hilfsmittel, m Funktionen f_1, f_2, \dots, f_m auf ihre lineare Abhängigkeit hin zu prüfen, bestehen in der Untersuchung entweder der Wronskischen oder der Casoratischen oder der Gramschen Determinante. In zahlreichen Fällen läßt sich aber das Ziel mit ganz bedeutend weniger Arbeit erreichen, wie sogleich an einem Beispiel gezeigt werden soll. Es handle sich um drei, durch folgende Potenzreihen erklärte Funktionen:

$$f_1 = 1 - 2x^2 + \frac{5}{3}x^4 - \frac{26}{45}x^6 + \frac{65}{315}x^8 + \dots,$$

$$f_2 = 1 - 8x^2 + \frac{32}{3}x^4 - \frac{256}{45}x^6 + \frac{512}{315}x^8 + \dots,$$

$$f_3 = 1 - 2x^2 + \frac{2}{3}x^4 - \frac{4}{45}x^6 + \frac{2}{315}x^8 + \dots$$

Bildet man aus den Koeffizienten eine Matrix und eliminiert man auf einmal die in den Spalten für x^2 und x^4 stehenden Elemente dadurch,

daß man die Zeilen bzw. mit 8, -1 , -4 multipliziert und alsdann addiert, so findet man — siehe die untenstehende Ausrechnung —, daß die Elemente in den Spalten für x^6 und x^8 von selbst verschwinden, während in der Spalte für x^0 sich 3 ergibt. Folglich ist, wenigstens mit der durch das Abbrechen der Potenzreihen bei den Gliedern mit x^8 bedingten Genauigkeit,

$$8f_1 - f_2 - 4f_3 = 3.$$

(Es ist

$$f_1 = (\cos x)^4, \quad f_2 = \cos 4x, \quad f_3 = \cos 2x,$$

und in der Tat gilt die genaue Beziehung

$$8(\cos x)^4 - \cos 4x - 4\cos 2x = 3.)$$

Ausrechnung:

	x^0	x^2	x^4	x^6	x^8	f_1	f_2	f_3
8	1	-2	$\frac{5}{3}$	$-\frac{26}{45}$	$\frac{65}{315}$	1		
-1	1	-8	$\frac{32}{5}$	$-\frac{256}{45}$	$\frac{512}{315}$		1	
-4	1	-2	$\frac{2}{3}$	$-\frac{4}{45}$	$\frac{2}{315}$			1
	3			0	0	8	-1	-4

Nun bedenke man, wie außerordentlich viel mühsamer es gewesen wäre, das Ergebnis etwa mit Hilfe der Gramschen Determinante zu gewinnen. Nach Hinzufügen der Funktion $f_4 = 1$ hätte man sechs Produkte von je zwei Potenzreihen zu bilden, neun Integrale und eine Determinante vierter Ordnung auszuwerten gehabt, allein um die lineare Abhängigkeit der Funktionen f an sich nachzuweisen. Um dann vollends die Faktoren 8, -1 , -4 , -3 zu finden, wären noch einige weitere Determinanten auszurechnen gewesen, ja, wenn man den Forderungen einiger Lehrbücher nachkommen wollte, hätte man sogar vorher die Funktionen f durch ein äquivalentes System paarweise orthogonaler Funktionen ersetzen müssen.

Wenn die Anzahl m der Funktionen f mehr als drei beträgt, so müssen von den zugehörigen Potenzreihen mindestens $(m + 1)$ Glieder gegeben sein und es ist wieder beschleunigtes Eliminieren zu empfehlen, d. h. ein Vorgehen in der Weise, daß man von der Matrix der Koeffizienten bei jedem Schritt drei Zeilen zusammenfaßt und aus ihnen durch lineare Verbindung die Elemente zweier Spalten gleichzeitig entfernt.

Der Verlauf der Grenzkurven zwischen labilen und stabilen Lösungsgebieten der Mathieschen Differentialgleichung.

Von

M. J. O. Strutt in Eindhoven (Niederlande).

Die Mathiesche Differentialgleichung:

$$(1) \quad \frac{d^2 y}{dx^2} + (R + 2h^2 \cos 2x) \cdot y = 0$$

tritt in vielen Anfangs- und Randwertproblemen der mathematischen Physik auf¹⁾.

Bei der zuerst genannten Problemklasse sind die Parameter R und h^2 vorgegeben und es tritt die Frage auf, für welche Werte dieser Parameter eine „stabile“ Lösung vorhanden ist, derart, daß y für alle x bei beliebiger Wahl der Anfangsbedingungen endlich bleibt. Wächst y mit x in diesem Falle über alle Grenzen, so nenne ich die Lösung „labil“. Bleibt y für alle x bei beliebigen Anfangsbedingungen endlich, so existiert eine periodische Lösung der Gleichung (1), wie Floquets Satz lehrt.

Bei Randwertproblemen ist die Fragestellung eine andere. Handelt es sich um ein Eigenwertproblem, so sind beide Parameter R und h^2 aus den Bedingungen des Problems zu ermitteln. Es wird meistens zuerst durch Eindeutigkeitsforderungen R als Funktion von h^2 bestimmt und sodann die Lösung ebenfalls als Funktion von h^2 ermittelt. Die homogenen Randbedingungen ergeben darauf die Eigenwerte.

Liegt ein lineares Randwertproblem vor, so ist h^2 vorgegeben und durch Eindeutigkeitsforderungen wird R als Funktion von h^2 ermittelt.

Die mehrfach erwähnten Eindeutigkeitsforderungen äußern sich in den meisten Fällen dahin, daß y eine Funktion mit der Periode 2π von x sein soll.

¹⁾ P. Humbert, *Fonctions de Lamé et de Mathieu*, Paris, Gauthier-Villars, 1926. — E. T. Whittaker und G. N. Watson, *Modern Analysis* (1920).

Es läßt sich einfach zeigen²⁾, daß die durch diese Forderung festgelegten Beziehungen zwischen R und h^2 , als Kurven in einer Ebene mit R und h^2 als rechtwinklige Koordinaten gedeutet, die Grenzen ergeben zwischen jenen Gebieten dieser Ebene, für die stabile Lösungen existieren, und den Gebieten, denen labile Lösungen der Gleichung (1) entsprechen.

H. Poincaré³⁾ hat bei der Untersuchung dieser Grenzkurven, bei der er von Reihen F. Tisserands⁴⁾ ausging, folgendes gefunden:

a) der Punkt ($\sqrt{R} = 0; h^2 = 0$) bildet einen isolierten Punkt dieser Grenzkurven in einer (\sqrt{R}, h^2)-Ebene;

b) die Punkte ($\sqrt{R} = g, h^2 = 0$), mit $g = 1, 2, 3, 4, 5, \dots$, sind Doppelpunkte der Grenzkurven;

c) in der unter a) erwähnten Ebene berühren sich diese Kurven auf der \sqrt{R} -Achse:

für $\sqrt{R} = 1$ von nullter Ordnung,

für $\sqrt{R} = 2$ von erster Ordnung,

für $\sqrt{R} = 3$ von zweiter Ordnung,

für $\sqrt{R} = 4$ von dritter Ordnung.

Die Ergebnisse b) und c) geben Poincaré Anlaß zu der Vermutung, daß die Kurven sich auf der \sqrt{R} -Achse für $\sqrt{R} = n$ (ganze Zahl) von $(n - 1)$ -ter Ordnung berühren werden, ohne daß sich diese Vermutung jedoch mit Hilfe der Reihen Tisserands bestätigen ließ.

Mein Ziel ist es, zu zeigen, daß der Punkt ($R = 0, h^2 = 0$) in einer (R, h^2)-Ebene kein isolierter Punkt ist, und weiterhin, zu zeigen, daß sich die Richtigkeit von Poincarés Vermutung leicht nachweisen läßt.

Ich benutze hierzu die Rechnungen Mathieus⁵⁾.

Die Mathieuschen Funktionen, welche sich für

$$h^2 \rightarrow 0$$

reduzieren auf

$$\sin g x \quad (g \text{ ganze Zahl}),$$

bezeichne ich mit

$$S e_g(h^2, x),$$

und jene Mathieusche Funktionen, die sich für

$$h^2 \rightarrow 0$$

²⁾ A. W. Young, Proc. Edinb. math. Soc. **33**, 75 (1914–15). — Verf., Ann. d. Physik **84** (1927) S. 485.

³⁾ Méthodes nouvelles de la Mécanique céleste, II, p. 229.

⁴⁾ Bulletin Astronomique **9** (1892), p. 102.

⁵⁾ E. Mathieu, Cours de Physique Mathématique (1873).

reduzieren auf

$$\cos gx \quad (g \text{ ganze Zahl}),$$

mit

$$Ce_g(h^2, x).$$

Ich behaupte:

Satz 1. Die Punkte $(R = n^2, h^2 = 0)$ mit $n = 1, 2, 3, 4, 5, \dots$ sind Doppelpunkte jener Kurven, welche die Gebiete der (R, h^2) -Ebene, für die stabile Lösungen der Mathieuschen Gleichung existieren, trennen von den Gebieten dieser Ebene, für die labile Lösungen existieren.

Es läßt sich zeigen, daß nie für gleiche Werte von R und h^2 gleichzeitig eine Lösung $Se_g(h^2, x)$ und eine Lösung $Ce_g(h^2, x)$ vorhanden sein kann⁶⁾.

Zu jeder Funktion Ce_g und zu jeder Funktion Se_g gehört je eine Kurve

$$R = f(h^2),$$

die eine Grenzkurve in unserem Sinne ist. Da nun aber für

$$h^2 \rightarrow 0$$

die zuletzt angeschriebene Gleichung sich in beiden Fällen reduziert auf

$$R_{h^2 \rightarrow 0} = g^2 \quad (g \text{ ganze Zahl}),$$

müssen von jedem der Punkte

$$(R = g^2, h^2 = 0) \quad (g \text{ ganze Zahl})$$

je zwei Grenzkurven ausgehen, deren eine zur Funktion Se_g und deren andere zur Funktion Ce_g gehört.

Weil aber sämtliche Se_g und Ce_g , mit Ausnahme von Se_0 , nicht identisch verschwindende, reelle Funktionen sind, geht hieraus hervor, daß die zuletzt erwähnten Punkte Doppelpunkte der Grenzkurven bilden, q. e. d.

Weiterhin behaupte ich:

Satz 2. Durch den Punkt $(R = 0, h^2 = 0)$ geht eine einzige reelle Grenzkurve in der (R, h^2) -Ebene.

Dieser Satz ist eine unmittelbare Folge des früheren. Da die Funktion Se_0 identisch verschwindet, die Funktion Ce_0 aber reell ist, geht vom Punkt $(R = 0, h^2 = 0)$ der (R, h^2) -Ebene eine einzige reelle Grenzkurve aus. Es kann dieser Punkt also keinen isolierten Punkt bilden.

Es gilt noch der

Satz 3. Die Grenzkurve durch den Punkt $(R = 0, h^2 = 0)$ der (R, h^2) -Ebene berührt die h^2 -Achse von erster Ordnung; sie verläuft für kleine h^2 im zweiten Quadranten dieser Ebene.

⁶⁾ E. L. Ince, Proc. Camb. Phil. Soc. 21 (1922), p. 117.

Zum Beweis betrachte ich den analytischen Ausdruck dieser Kurve⁷⁾:

$$(2) \quad R = -\frac{h^4}{2} + \frac{7}{128}h^8 - \frac{116}{2304}h^{12} + \dots,$$

aus der beide Behauptungen des Satzes 3 unmittelbar hervorgehen.

Da Poincaré die Verhältnisse in einer (\sqrt{R}, h^2) -Ebene betrachtete, ist klar, daß er die Grenzkurve (2) nicht fand (sie wurde bei ihm imaginär für kleine h^2) und nur feststellen konnte, daß der Punkt $(\sqrt{R} = 0, h^2 = 0)$ in der zuletzt genannten Ebene einen isolierten Punkt der Grenzkurven bildet.

Zum Schluß beweise ich den

Satz 4. *Je zwei Grenzkurven, die von den Doppelpunkten ($R = n^2, h^2 = 0$), mit ganzzahligem n , ausgehen, berühren sich in diesen Punkten von $(n - 1)$ -ter Ordnung.*

Hierzu betrachte ich die Lösungsmethode Mathieus. Entwickelt man C_{e_n} und S_{e_n} nach Potenzen von h^2 , so zeigt sich, daß bei dieser Entwicklung der Koeffizient von h^{2n} einen Nenner erhält, der verschwindet.

Vom Gliede mit h^{2n} an setzt somit die Spezialrechnung ein. Hieraus ergibt sich, daß $R_{C_{e_n}}(h^2)$ mit $R_{S_{e_n}}(h^2)$ bis zum Gliede $h^{2(n-1)}$ zusammenfällt.

Von diesem Gliede an unterscheidet sich $R_{C_{e_n}}(h^2)$ von $R_{S_{e_n}}(h^2)$. Somit fallen die $(n - 1)$ ersten Differentialquotienten von $R_{C_{e_n}}(h^2)$ und $R_{S_{e_n}}(h^2)$ nach h^2 zusammen, während der n -te Differentialquotient beider Ausdrücke nach h^2 verschieden ist. Dann berühren sich aber die Kurven $R_{C_{e_n}}(h^2)$ und $R_{S_{e_n}}(h^2)$ im Punkte $(R = n^2, h^2 = 0)$ der (R, h^2) -Ebene von $(n - 1)$ -ter Ordnung, q. e. d.⁸⁾

Eindhoven, Natuurkundig Laboratorium der N. V. Philips' Gloeilampenfabrieken.

⁷⁾ E. Mathieu, loc. cit., S. 141.

⁸⁾ Anmerkung bei der Korrektur. Die obigen Sätze stellen Verschärfungen allgemeiner Sätze von O. Haupt (Math. Ann. 79, S. 281 (1919)) für den Spezialfall der Mathieschen Gleichung dar.

Zur Theorie der algebraischen Potentialfunktionen des dreidimensionalen Raumes. I.

Von

Stefan Bergmann in Berlin.

Die allgemeine Theorie der algebraischen Funktionen dreier reeller Veränderlicher x, y, z bietet wesentliche Schwierigkeiten.

Es ist daher zweckmäßig, die Untersuchung auf spezielle Klassen dieser Funktionen zu beschränken. Die Analogie mit dem zweidimensionalen Fall legt die Vermutung nahe, daß hierzu die algebraischen Potentialfunktionen durch einfaches Verhalten besonders geeignet sind.

Wir betrachten im folgenden stets komplexe Funktionen reeller Veränderlicher, es ist also jede betrachtete Funktion $F(x, y, z)$

$$(1) \quad F(x, y, z) \equiv V(x, y, z) + iW(x, y, z),$$

wo V und W reelle (*voneinander völlig unabhängige*) Funktionen sind. Im Einklang mit der Funktionentheorie verstehen wir unter einer algebraischen Potentialfunktion eine Funktion $\bar{R}(x, y, z) \equiv R(x, y, z, \zeta)$, wobei

1. $R(x, y, z, \zeta)$ eine rationale Funktion der Veränderlichen x, y, z, ζ ist, zwischen welchen die algebraische Gleichung

$$(2) \quad P(x, y, z, \zeta) = 0$$

besteht, und wobei

2. $\bar{R}(x, y, z)$ die Differentialgleichung

$$(3) \quad \frac{\partial^2 \bar{R}}{\partial x^2} + \frac{\partial^2 \bar{R}}{\partial y^2} + \frac{\partial^2 \bar{R}}{\partial z^2} = 0$$

befriedigt.

In einer früheren Arbeit¹⁾ wurde gezeigt, daß man jeder harmonischen Funktion $F(x, y, z)$ eine analytische Funktion $f(u, \zeta)$ einer komplexen Veränderlichen u

$$\left(u \equiv x + iy \cos t + iz \sin t \equiv x + \frac{iy-z}{2} \frac{1}{\zeta} + \frac{iy+z}{2} \zeta \right), \quad [e^{it} \equiv \zeta]$$

die noch einen komplexen Parameter e^{it} bzw. ζ enthält, zuordnen kann, derart, daß in einem gewissen Raumteile $G_\alpha(a)$ F sich in der Form

$$(4a) \quad F(x, y, z) = \frac{1}{2\pi} \int_0^{2\pi} f(u, e^{it}) dt$$

$$(4b) \quad \equiv \frac{1}{2\pi i} \int_a f \left[\left(x + \frac{iy-z}{2} \frac{1}{\zeta} + \frac{iy+z}{2} \zeta \right), \zeta \right] \frac{d\zeta}{\zeta}$$

darstellen läßt. Hierbei ist α eine ebene Jordankurve in der komplexen evtl. mehrblättrigen ζ -Ebene. Die Gesamtheit derjenigen Funktionen f , mit deren Hilfe man F durch (4) erhalten kann, wurde in der Arbeit I als die zu F zugeordnete Klasse bezeichnet.

Es entsteht nun die Frage, ob in der einer algebraischen Potentialfunktion F zugeordneten Klasse Funktionen f existieren, die man auf eine

¹⁾ „Zur Theorie der ein- und mehrwertigen harmonischen Funktionen des dreidimensionalen Raumes“, Math. Zeitschr. 24 (1926), S. 641—669. Wir werden im folgenden diese Untersuchung als Arbeit I bezeichnen. Es wird die Kenntnis mindestens der ersten drei Kapitel dieser Arbeit vorausgesetzt. Die früher benutzten Bezeichnungen wurden im wesentlichen beibehalten, mit der Ausnahme, daß in vorliegender Arbeit — im Gegensatz zu der früheren — die zugeordneten Polynome unter Umständen auch mit großen Buchstaben bezeichnet wurden. Außerdem wird mit u sowohl der Ausdruck

$$\left[x + \frac{iy-z}{2} e^{-it} + \frac{iy+z}{2} e^{it} \right]$$

wie der aus dem letzten transformierte

$$\left[x + \frac{iy-z}{2} \frac{1}{\zeta} + \frac{iy+z}{2} \zeta \right]$$

bezeichnet. Unter U verstehen wir

$$U \equiv \zeta u \equiv \frac{iy-z}{2} + x\zeta + \frac{iy+z}{2} \zeta^2.$$

Ich bemerke, daß in dem Ausdruck (1) des § 1 der Arbeit I sich ein Druckfehler befindet: es soll

$$(1) \quad x + iy \cos t + iz \sin t = x + e^{it} \left(\frac{iy+z}{2} \right) + e^{-it} \left(\frac{iy-z}{2} \right),$$

und nicht

$$= x + e^{it} \left(\frac{iy-z}{2} \right) + e^{-it} \left(\frac{iy+z}{2} \right)$$

heißen.

einfache Weise *funktionentheoretisch* (betrachtet als Funktion von u) charakterisieren kann. Es zeigt sich in der Tat, daß es zu *jedem algebraischen* $F(x, y, z)$ *stets ein* $f(u, e^{it})$ *gibt, das die Form*

$$(5) \quad f(u, e^{it}) = \int_0^1 a(u, e^{it}, T) dT$$

hat, wo $a(u, e^{it}, T)$ eine algebraische Funktion in u, e^{it}, T ist. Dementsprechend liegt es nahe, diejenigen harmonischen Funktionen zu untersuchen, die durch die Mittelwertbildung (4a) aus (5) entstehen.

Der Inhalt der vorliegenden Arbeit wird der Untersuchung der Potentialfunktionen mit algebraisch-logarithmischen Zugeordneten gewidmet, d. h. denjenigen Funktionen $F(x, y, z)$, auf die man geführt wird, falls $f(u, e^{it})$ eine algebraisch-logarithmische Funktion von u und e^{it} ist.

In der Darstellung (4b) ist a eine ebene Jordansche Kurve, die in der mehrblättrigen Riemannschen Fläche von ζ geschlossen oder offen sein kann. Der erste Teil unserer Arbeit untersucht diejenigen Potentialfunktionen, die

1. bei geschlossenen a (§ 2 bis § 5),
2. bei offenen a

durch die Integration (4b) aus einem algebraischen f entstehen.

In § 1 formulieren wir unsere Aufgabe und beweisen die Identität:

$$(6) \quad F(r, \cos \vartheta, \varphi) = \frac{1}{2\pi} \int_0^{2\pi} dt \left\{ \int_0^1 \left[\frac{d \left[\xi F \left(\xi, 0, t - i \ln \sqrt{\frac{T}{1-T}} \right) \right]}{d\xi} \right]_{\xi = \frac{2}{i} \sqrt{T(1-T)} u} + \frac{i}{2\sqrt{T(1-T)}} \frac{d \left[F \left(\frac{2}{i} \sqrt{T(1-T)} u, \cos \vartheta, t - i \ln \sqrt{\frac{T}{1-T}} \right) \right]}{d \cos \vartheta} \Big|_{\cos \vartheta = 0} \right] dT \Big\},$$

wo r, ϑ, φ die Polarkoordinaten bedeuten, woraus die Darstellung (5) von f folgt.

In § 2 wird nun folgendes bewiesen. Ist

$$P_2 \left[\left(x\zeta + \frac{iy-z}{2} + \frac{iy+z}{2} \zeta^2 \right), \zeta \right]$$

ein Polynom [n -ten Grades] in $U \left(\equiv x\zeta + \frac{iy-z}{2} + \frac{iy+z}{2} \zeta^2 \right)$ und [($2n + 2m$)-ten Grades] in ζ , so ist die Diskriminante von $P_2(\zeta) = 0$ der Ausdruck

$$D(P_2) \equiv \left[\prod_{\substack{k, l \\ k \neq l}} (\zeta_k - \zeta_l) \right] \quad (k, l = 1, 2, 3, \dots, 2(m+n))$$

(wo ζ_k, ζ_l die Wurzeln der Gleichung $P_2(\zeta) = 0$ sind) eine in x, y, z rationale Funktion, die eventuell identisch verschwinden kann. Bezeichnet man mit

$$P_1 \left[\left(x\zeta + \frac{iy-z}{2} + \frac{iy+z}{2}\zeta^2 \right), \zeta \right]$$

ein Polynom gleicher Art wie $P_2[U, \zeta]$, so wird uns durch

$$(7) \quad F_q = \frac{1}{2\pi i} \int_a \frac{P_1(U, \zeta)}{P_2(U, \zeta)} d\zeta$$

ein Zweig einer algebraischen Funktion

$$(8) \quad \frac{P_1 \left[\left(x\zeta_q + \frac{iy-z}{2} + \frac{iy+z}{2}\zeta_q^2 \right), \zeta_q \right]}{\frac{\partial P_2 \left[\left(x\zeta_q + \frac{iy-z}{2} + \frac{iy+z}{2}\zeta_q^2 \right), \zeta_q \right]}{\partial \zeta_q}} + R(x, y, z)^2$$

gegeben im Falle, wenn $D(P_2) \neq 0$. $R(x, y, z)$ ist eine rationale, durch die Formel (9) des § 2 gegebene Funktion oder Null. Ist $D(P_2)$ identisch Null, so tritt an Stelle von (8) ein komplizierterer Ausdruck; nämlich der Ausdruck (10) des § 2. Variiert man a^3 bzw. integriert man (7) für verschiedene Punkte des x, y, z -Raumes, so erhält man im allgemeinen verschiedene Zweige der Funktion (8). Die Behauptung, daß die F_q Zweige einer algebraischen Funktion sind, erleidet jedoch eine Ausnahme, falls die F_q nur durch die Fortsetzung über nicht reelle x, y, z -Werte ineinander übergeführt werden können, denn, da wir unsere Untersuchung auf reelle Werte von x, y, z beschränken, müssen wir in diesem Falle von verschiedenen Funktionen sprechen. Man übersieht auch leicht, wie sich das Integral

$$(9) \quad \frac{1}{2\pi i} \int_a \frac{P_1}{P_2} d\zeta$$

in dem x, y, z -Raume verhält, und wo (9) Sprünge erleidet. Die Flächen, an denen dies eintritt, nennen wir Zuordnungstrennflächen, die zu a gehören. Denjenigen Bereich $G_q(a)$ des x, y, z -Raumes, wo die Beziehung (4 b) besteht, bezeichnen wir als Zuordnungsraum von F_q , der zu der Kurve a gehört.

Aus der Darstellung (8) erhält man auch Aufschluß über die Lage der singulären Punkte und Linien, deren Gleichungen man explizite an-

²⁾ Das Symbol der partiellen Differentiation (∂) soll hier wie auch stets im folgenden eine Differentiation bedeuten, bei welcher sämtliche unter dem Funktionszeichen auftretenden Veränderlichen als unabhängige [der Relation $P_2(\zeta_k) = 0$ nicht unterworfen] Variablen zu behandeln sind.

³⁾ Es soll darunter verstanden werden, daß die Kurve a in der Euklidischen Ebene sich bewegen bzw. beliebig variieren darf, ohne ihr Geschlecht zu ändern. a muß derart gewählt werden, daß $P_2(U, \zeta)$ auf a nirgends verschwindet.

geben kann. Die Singularitäten unserer Funktionen besitzen eine interessante Eigenschaft. Ist $f \equiv \frac{P_1(u, \zeta)}{P_2(u, \zeta)}$, so kann man eine endliche Anzahl in x, y, z, ζ rationalen Funktionen $f_k(\zeta, x, y, z)$ (die in einfacher Weise mit f zusammenhängen) angeben; eine Funktion F_a (die zu der von uns hier betrachteten Klasse gehört) läßt sich im großen (a fortiori in der Umgebung der Singularitäten) durch die Ausdrücke der Form

$$(10) \quad \int_a f_k d\zeta$$

darstellen, wobei man nur eine Integrationskurve a (z. B. Einheitskreis) zu verwenden braucht, um den gesamten Wertevorrat der Funktion in ihrem Riemannschen Raum auszuschöpfen.

Umgekehrt gehört jede Funktion, die diese letzte Eigenschaft besitzt, zu unserer Klasse.

In § 3 wird auf den Zusammenhang zwischen dem vollständigen Differential (Näheres siehe S. 657) unserer Funktionen und deren Singularitäten eingegangen, den man durch die Deutung des Cauchyschen Residuensatzes erhält. Bevor wir unser Ergebnis formulieren, müssen wir einige Erklärungen vorausschicken. Ist uns wiederum ein Polynom $P_2[u, \zeta]$ gegeben, so bilden wir die Wurzeln $\alpha_1, \dots, \alpha_n$ der Gleichung

$$P_2(\alpha, \zeta) = 0.$$

Die α_k sind Funktionen, die nur von ζ abhängen. Mit $\delta(P_2)$ bezeichnen wir die Diskriminante $[\prod_{k \neq l} (\alpha_k - \alpha_l)]$,

die eine rationale Funktion von ζ ist. $\delta(P_2)$ kann entweder identisch Null sein oder nicht.

Jedem Polynom $P_2[u, \zeta]$ werden wir n geschlossene Regelflächen C_1, C_2, \dots, C_n zuordnen. Jede der Regelflächen C_k wird in bestimmter Weise durch die Bewegung einer Geraden $T_k(\tau)$ erzeugt, wobei der reelle Parameter τ von 0 bis 2π läuft.

Der Einfachheit halber nehmen wir an, daß die Kurve \mathcal{S} mit einigen der Flächen C_k je zwei Durchstoßpunkte besitzt⁴⁾. Diese Punkte $s_1^{(k)}$ und $s_2^{(k)}$ mögen auf den erzeugenden Geraden $T_k(\tau_1^{(k)})$ und $T_k(\tau_2^{(k)})$ liegen (vgl. Fig. 1).

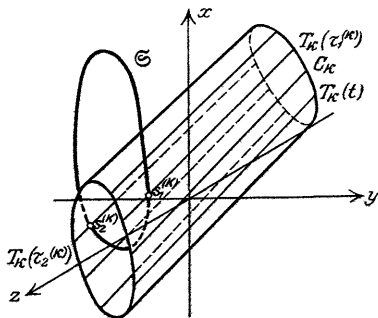
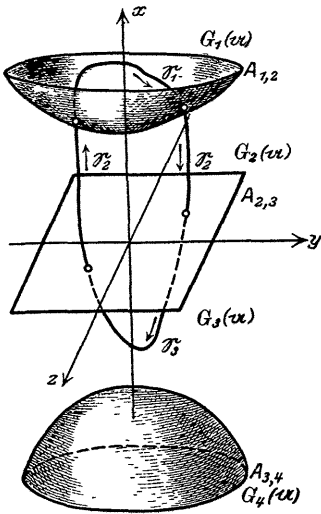


Fig. 1.

⁴⁾ Fallen die beiden Durchstoßpunkte $\tau_1^{(k)}$ und $\tau_2^{(k)}$ zusammen, so muß man sie in einer in § 3 näher angegebenen Weise doppelt zählen.

Sei \mathfrak{S} eine geschlossene Jordansche Raumkurve mit bestimmtem Durchlaufungssinn. \mathfrak{S} wird im allgemeinen in mehrere Zuordnungsräume $G_q(\alpha)$ von (7) eintreten und daher durch die Zuordnungstrennflächen in eine Anzahl von Teilen, die wir mit \mathfrak{S}_q bezeichnen, zerlegt (vgl. Fig. 2). Das Ergebnis des § 3 läßt sich nun folgendermaßen formulieren: Ist F_q ein Zweig einer algebraischen Potentialfunktion, die zu der von uns betrachteten Klasse gehört, so lassen sich zwei andere Funktionen (derselben Klasse) F_q^* und F_q^{**} angeben, derart, daß



$$(11) \quad F_q dx + i F_q^* dy + i F_q^{**} dz$$

ein vollständiges Differential ist, und im Falle, daß $\delta(P_2) \equiv 0$ ist, die Beziehung besteht:

Fig. 2. ($\mathfrak{P}_q \equiv \mathfrak{S}_q$.)

$$(12) \quad \sum_{\mathfrak{S}_q} \int (F_o dx + i F_q^* dy + i F_q^{**} dz) = \sum_{v=0}^{v=n} \int_{\tau_1^{(v)}}^{\tau_2^{(v)}} \frac{P_1(\alpha_v, \zeta)}{\frac{\partial P_2(\alpha_v, \zeta)}{\partial \alpha_v}} d\zeta,$$

$$(13) \quad \mathfrak{S} = \sum \mathfrak{S}_q,$$

Im Falle, daß $\delta(P_2) \equiv 0$ ist, erhalten wir eine ähnliche, etwas kompliziertere Formel.

In § 4⁵⁾ werden diejenigen Potentialfunktionen betrachtet, die aus (4 b) hervorgehen, falls $\frac{f(u, \zeta)}{\zeta}$ eine algebraische Funktion in u und ζ ist, und falls die Kurve α (für den betreffenden Integrationswert x, y, z) sich auf einen Punkt zusammenziehen läßt. Wir erhalten eine Klasse von algebraischen Funktionen, die eine große Ähnlichkeit mit den Potentialfunktionen mit rationalen Zugeordneten besitzen.

In § 5 werden diejenigen Potentialfunktionen untersucht, zu denen man gelangt, falls α sich *nicht* auf einen Punkt zusammenziehen läßt. Wir erhalten transzendente Funktionen, die in einer einfachen (im Text näher auseinandergesetzten) Weise mit den Perioden der Integrale I, II. und III. Gattung zusammenhängen, wobei natürlich die Perioden und die Unendlichkeitsstellen der Integrale Funktionen von x, y, z sind. Bezeichnet man in üblicher Weise mit $\omega_{\alpha\beta}(x, y, z)$, $\omega'_{\alpha\beta}(x, y, z)$, $\eta_{\alpha\beta}(x, y, z)$, $\eta'_{\alpha\beta}(x, y, z)$, $\Omega_{\beta}(x, y, z)$, $\Omega'_{\beta}(x, y, z)$ [$\alpha, \beta = 1, 2, 3, \dots, \rho$; $\rho =$ Ge-

⁵⁾ Die ausführliche Darstellung der §§ 4 und 5 erscheint im Abschnitt II der Arbeit.

schlecht von $f(\zeta)$] diese Perioden, so lassen sich unsere Potentialfunktionen stets durch die algebraischen Funktionen und diese Perioden in geschlossener Form darstellen. Es sind dabei $\omega_{\alpha\beta}(x, y, z)$ und $\omega'_{\alpha\beta}(x, y, z)$ [$\beta = 1, 2, \dots, \rho$] im *allgemeinen* Zweige einer Funktion. Das gleiche gilt für $\eta_{\alpha\beta}(x, y, z)$, $\eta'_{\alpha\beta}(x, y, z)$ und für $\Omega_{\beta}(x, y, z)$, $\Omega'_{\beta}(x, y, z)$. Zu einer zweiten Darstellungsform gelangen wir, falls wir die Thetafunktionen einführen und die Perioden der Integrale zweiter und dritter Gattung durch Thetafunktionen und Perioden erster Gattung ausdrücken. Es sei hervorgehoben, daß dabei sowohl für die Parameter, wie für das Argument die $\omega'_{\alpha\beta}(x, y, z)$ [nach einer Normierung] einzusetzen sind, *die Funktionen von x, y, z sind*.

Unsere Transzendenten besitzen eine wichtige Eigenschaft: Betrachtet man sie als Funktionen *einer* Veränderlichen x (bzw. y oder z), so genügen sie einer gewöhnlichen linearen Differentialgleichung

$$(14) \quad \sum_{\nu=1}^{\nu=h} A_{1\nu}(x, y, z) \frac{d^{\nu} F(x, y, z)}{dx^{\nu}} = A_1(x, y, z)$$

mit in x, y, z algebraischen Koeffizienten. Die Ordnung h der Gleichung (14) hängt in einer einfachen Weise mit dem Geschlecht ρ von f zusammen. Ebenfalls befriedigt F zwei weitere gewöhnliche Dfg. (als Funktion von y und von z).

Die Transzendenten befriedigen somit simultan vier Differentialgleichungen: die Potentialgleichung und drei gewöhnliche lineare Gleichungen von der Form (14).

Die in § 5 erhaltenen Funktionen sind im allgemeinen auf algebraischen Kurven singular. Sie lassen sich in der Umgebung der singulären Linien im allgemeinen nach gewissen algebraischen Funktionen von x, y, z entwickeln, wobei aber außer negativen Potenzen noch Logarithmen dieser algebraischen Funktionen in der Entwicklung auftauchen.

Die Funktionen, zu denen wir bei offenem α gelangen, weisen eine große Ähnlichkeit mit den in § 5 betrachteten Transzendenten auf. Bei gleichen f genügen sie ebenfalls einer linearen Differentialgleichung, deren linke Seite mit derjenigen von (14) übereinstimmt und nur durch ein anderes freies $A_1(x, y, z)$ sich von (14) unterscheidet.

Unsere Betrachtungen lassen sich in verschiedener Weise verallgemeinern.

I. Die Funktionen, die der Wellengleichung

$$(15) \quad \frac{\partial^2 V(x, y, t)}{\partial^2 t} = \frac{\partial^2 V(x, y, t)}{\partial x^2} + \frac{\partial^2 V(x, y, t)}{\partial y^2}$$

genügen, lassen sich fast in derselben Weise wie die Potentialfunktionen behandeln.

Es liegt nahe unsere Potentialfunktionen für komplexe Werte von x, y, z zu betrachten, wobei sich die sämtlichen hier angewandten Schlüsse wiederholen. Betrachtet man dann diese Funktionen in gewissen dreidimensio-

nalen Mannigfaltigkeiten des sechsdimensionalen Raumes, so gelangt man zu Funktionen, die die Wellengleichung befriedigen.

II. Man kann ferner leicht unser Verfahren auf die n -dimensionale Potentialtheorie übertragen, da die n -dimensionale Potentialfunktion im kleinen sich in der Form

$$(16) \int_0^{2\pi} \int_0^{2\pi} \dots \int_0^{2\pi} f[(x_1 + ix_2 \cos t_1 + ix_3 \sin t_1 \cos t_2 + \dots + ix_n \sin t_1 \sin t_2 \dots \sin t_{n-2}) e^{it_1}, e^{it_2}, \dots, e^{it_{n-2}}] dt_1 dt_2 \dots dt_{n-2}$$

darstellt. Diese Verallgemeinerung können wir benutzen, um unsere Betrachtungen auf weitere Typen von Differentialgleichungen zweiter Ordnung zu übertragen: führt man nämlich in einer n -dimensionalen Potentialfunktion neue Koordinaten ein und betrachtet man nur diejenigen Funktionen, die nur von einigen der neuen Veränderlichen abhängen, so genügt die so entstandene Funktionenklasse einer neuen Differentialgleichung. Wesentliche Schwierigkeiten bietet eigentlich die umgekehrte Frage: ob und wie sich die Lösungen einer gegebenen Differentialgleichung als Spezialisierung der n -dimensionalen Potentialfunktionen betrachten lassen.

Ich möchte die Gelegenheit benutzen, den Herren K. Löwner und Felix Pollaczek für mannigfache Ratschläge und Herrn E. Rothe für die Hilfe bei der Redaktion der Arbeit meinen Dank auszusprechen.

§ 1.

Über die Zugeordnete einer dreidimensionalen Potentialfunktion.

In der Arbeit I wurde gezeigt, daß eine im Nullpunkte reguläre dreidimensionale harmonische Funktion

$$F_q \equiv V(x, y, z) + iW(x, y, z)$$

sich im kleinen in der Form

$$(1) \quad F_q = \frac{1}{2\pi} \int_0^{2\pi} f(u, e^{it}) dt$$

$$(2) \quad u \equiv x + iy \cos t + iz \sin t$$

darstellen läßt, wobei $f(u, e^{it})$ als Funktion von u betrachtet in der Umgebung des Nullpunktes für sämtliche reelle t des Intervalles 0 bis 2π regulär ist.

f ist umgekehrt bis auf eine Nullfunktion (vgl. Arbeit I, S. 646⁶⁾)

⁶⁾ In der Formel (6) § 1 der Arbeit I ist ein Druckfehler; es soll natürlich ($|k| > s$) und nicht ($|k| < s$) heißen. In diesen Paragraphen lassen wir den Index q bei F fort. Dieser Index (vgl. Arbeit I S. 649 und 650) ist deswegen angebracht, weil man bei festem f je nach der Wahl von x, y, z im *allgemeinen* zu verschiedenen Potentialfunktionen F_q gelangt.

eindeutig durch F bestimmt. Dadurch aber, daß wir eine beliebige Nullfunktion zu f zufügen dürfen, entsteht eine große Willkür in der Wahl der Zugeordneten. Die wichtige Frage, die sich jetzt aufdrängt, ist die folgende: Wenn F zu einer bestimmten Klasse (z. B. algebraischer) Funktionen gehört, ist es dann möglich, f so zu wählen, daß man f — betrachtet als Funktion der Veränderlichen u — einfach charakterisieren kann? Diese Frage läßt keine eindeutige Antwort zu; wir werden auch im weiteren Teile der Arbeit sie durch systematische Untersuchung aller Zugeordneten zu beantworten versuchen. Für die vorläufigen Betrachtungen wird uns genügen, wenn wir eine solche einfache Zugeordnete aufstellen, zu der wir auf Grund des folgenden Satzes gelangen:

Ist $F(r, \cos \vartheta, \varphi)$ eine in der Umgebung des Nullpunktes reguläre harmonische Funktion, so gilt in der Umgebung des Nullpunktes die Beziehung:

$$(3) \quad F(r, \cos \vartheta, \varphi) = \frac{1}{2\pi} \int_0^{2\pi} dt \left\{ \int_0^1 \left[\frac{d \left[\xi F \left(\xi, 0, t - i \ln \sqrt{\frac{T}{1-T}} \right) \right]}{d \xi} \right]_{\xi = \frac{2}{i} \sqrt{T(1-T)} u} \right. \\ \left. - \frac{i}{2 \sqrt{T(1-T)}} \frac{d \left[F \left(\frac{2}{i} \sqrt{T(1-T)} u, \cos \vartheta, t - i \ln \sqrt{\frac{T}{1-T}} \right) \right]}{d \cos \vartheta} \right]_{\cos \vartheta = 0} dT \right\},$$

r, ϑ, φ bedeuten dabei die Polarkoordinaten des Raumes, u den Ausdruck (2) in Polarkoordinaten, d. h. $r [\cos \vartheta + i \sin \vartheta \cos(\varphi - t)]$.

Beweis. F ist in der Umgebung des Nullpunktes regulär, es läßt sich somit (ebenda) in der Form:

$$(4) \quad F(r, \cos \vartheta, \varphi) = \sum_n \sum_\nu^I A_{n\nu} r^n P_n^{(\nu)}(\cos \vartheta) e^{i\nu\varphi} + \sum_n \sum_\nu^{II} B_{n\nu} r^n P_n^{(\nu)}(\cos \vartheta) e^{i\nu\varphi}$$

darstellen. $A_{n\nu}$ und $B_{n\nu}$ bedeuten Konstanten. $P_n^{(\nu)}(\cos \vartheta)$ sind die zugeordneten Kugelfunktionen⁷⁾, $\sum_n \sum_\nu^I$ bedeutet die Summation von $\nu = -n$

⁷⁾ Vgl. dazu die Enzyklopädie der Mathematischen Wissenschaften, Verl. B. G. Teubner, Leipzig (1904—1916), 2, I. Teil, II. Hälfte, A. 10. Wangerin, Theorie der Kugelfunktionen und der verwandten Funktionen, insbesondere der Laméschen und Besselschen, Kap. III, § 8, S. 708 oder Heine, Handbuch der Kugelfunktionen, Verl. G. Reimer, Berlin (1878), II. Aufl., IV. Kap. Das letzte Werk wird im folgenden kurz als Heine zitiert.

Es sei hier darauf hingewiesen, daß unsere Normierung der zugeordneten Kugelfunktionen nicht ganz mit der Heineschen übereinstimmt. Unsere $P_n^{(\nu)}(\cos \vartheta)$ sind $\frac{P_n^{(\nu)}(\cos \vartheta)}{i^\nu}$ gleich, wo $P_n^{(\nu)}(\cos \vartheta)$ die zugeordneten Kugelfunktionen in der Heineschen Darstellung bedeuten.

bis $\nu = n$ und von $n = 0$ bis $n = \infty$ über alle diejenigen n und ν , wo $n - \nu$ gerade ist, $\sum_n \sum_\nu^{II} -$ über alle n und ν , wo $n - \nu$ ungerade ist.

Nach Heine (Seite 211, Formeln (35a) und (35b)) ist:

$$r^n P_n^{(\nu)}(\cos \vartheta) e^{i\nu\varphi} = \frac{1}{2\pi} 2^n \frac{(n+\nu)! (n-\nu)!}{(2n)!} \int_0^{2\pi} \frac{u^n e^{i\nu t} dt}{i^\nu} \quad (|\nu| \leq n).$$

Die rechte Seite von (4) läßt sich somit in der Form:

$$(5) \quad \frac{1}{2\pi} \int_0^{2\pi} \left\{ \sum_n \sum_\nu^I 2^n \frac{(n+\nu)! (n-\nu)!}{(2n)!} A_{n\nu} \frac{u^n e^{i\nu t}}{i^\nu} + \sum_n \sum_\nu^{II} 2^n \frac{(n+\nu)! (n-\nu)!}{(2n)!} B_{n\nu} \frac{u^n e^{i\nu t}}{i^\nu} \right\} dt$$

darstellen. Zwecks weiterer Umformung zerspalten wir in (5) die Ausdrücke $\frac{(n+\nu)! (n-\nu)!}{(2n)!}$ in zwei Faktoren und schreiben (5) in der Form:

$$(6) \quad \frac{1}{2\pi} \int_0^{2\pi} \left\{ \sum_n \sum_\nu^I \frac{i^n}{i^\nu} \cdot \frac{1.3 \dots (n+\nu-1). 1.3 \dots (n-\nu-1)}{1.3 \dots (2n-1)} \cdot A_{n\nu} 2^n \frac{\left(\frac{n+\nu}{2}\right)! \left(\frac{n-\nu}{2}\right)!}{n!} e^{i\nu t} \right. \\ \left. + \sum_n \sum_\nu^{II} \frac{i^{n-1}}{i^\nu} \cdot \frac{1.3 \dots (n+\nu). 1.3 \dots (n-\nu)}{1.3 \dots (2n-1)} \cdot B_{n\nu} 2^n \frac{\left(\frac{n+\nu-1}{2}\right)! \left(\frac{n-\nu-1}{2}\right)!}{n!} - e^{i\nu t} \right\} dt.$$

Wie aus der Theorie der Betafunktionen folgt, ist

$$(7a) \quad \frac{\left(\frac{n+\nu}{2}\right)! \left(\frac{n-\nu}{2}\right)!}{n!} = (n+1) \int_0^1 T^{\frac{n+\nu}{2}} (1-T)^{\frac{n-\nu}{2}} dT \\ = (n+1) \int_0^1 [T(1-T)]^{\frac{n}{2}} \left[\frac{T}{1-T}\right]^{\frac{\nu}{2}} dT$$

$$(7b) \quad \frac{\left(\frac{n+\nu-1}{2}\right)! \left(\frac{n-\nu-1}{2}\right)!}{n!} = \int_0^1 T^{\frac{n+\nu-1}{2}} (1-T)^{\frac{n-\nu-1}{2}} dT \\ = \int_0^1 \left[\frac{1}{T(1-T)}\right]^{\frac{1}{2}} [T(1-T)]^{\frac{n}{2}} \left[\frac{T}{1-T}\right]^{\frac{\nu}{2}} dT$$

gleich⁸⁾. Wir ersetzen in (6) $\frac{\left(\frac{n+\nu}{2}\right)! \left(\frac{n-\nu}{2}\right)!}{n!}$ und $\frac{\left(\frac{n+\nu-1}{2}\right)! \left(\frac{n-\nu-1}{2}\right)!}{n!}$

⁸⁾ Vgl. Enzyklopädie der Mathematischen Wissenschaften, B. G. Teubner, Leipzig (1899—1916) 2, I. Teil, I. Hälfte. A. 3. G. Brunel, Bestimmte Integrale, § 15, S. 176, Zeile 6 und 12 von oben. Wir ersetzen die in der Enzyklopädie benutzten Größen p und q in (7a) durch $\frac{n+\nu}{2}$ und $\frac{n-\nu}{2}$, und in (7b) durch $\frac{n+\nu-1}{2}$ und $\frac{n-\nu-1}{2}$.

durch die in (7a) und (7b) erhaltenen Integrale und vertauschen in der so entstandenen Potenzreihe die Summation mit der Integration nach T . Da ferner nach Heine S. 207 (Zeile 17 und 19 von oben):

$$(8a) P_n^{(\nu)}(0) = \frac{i^n \cdot 1.3 \dots (n + \nu - 1) \cdot 1.3 \dots (n - \nu - 1)}{i^\nu \cdot 1.3.5 \dots (2n - 1)} \quad \text{bei geraden } n - \nu,$$

$$(8b) P_n^{(\nu)}(0) = 0 \quad \text{bei ungeraden } n - \nu,$$

$$(8c) \frac{dP_n^{(\nu)}(\cos \vartheta)}{d \cos \vartheta} \Big|_{\cos \vartheta=0} = 0 \quad \text{bei geraden } n - \nu,$$

$$(8d) \frac{dP_n^{(\nu)}(\cos \vartheta)}{d \cos \vartheta} \Big|_{\cos \vartheta=0} = \frac{i^{n-1} \cdot 1.3 \dots (n + \nu) \cdot 1.3 \dots (n - \nu)}{i^\nu \cdot 1.3.5 \dots (2n - 1)} \quad \text{bei ungeraden } n - \nu$$

ist, so folgt, daß man (6) auch in der Form

$$(9) \frac{1}{2\pi} \int_0^{2\pi} dt \left\{ \int_0^1 \left[\sum_n \sum_\nu^I A_{n\nu} (n + 1) \left(\frac{2}{i} u \sqrt{T(1-T)}\right)^n P_n^{(\nu)}(0) \left(e^{it} \sqrt{\frac{T}{1-T}}\right)^\nu + \frac{i}{2\sqrt{T(1-T)}} \sum_n \sum_\nu^{II} B_{n\nu} \left(\frac{2}{i} u \sqrt{T(1-T)}\right)^n \frac{dP_n^{(\nu)}(\cos \vartheta)}{d \cos \vartheta} \Big|_{\cos \vartheta=0} \left(e^{it} \sqrt{\frac{T}{1-T}}\right)^\nu \right] dT \right\}$$

darstellen kann. Vergleicht man (9) mit der Darstellung (4) und berücksichtigt man die Relation (8b) und (8c), so folgt:

$$(10) \sum_n \sum_\nu^I A_{n\nu} (n + 1) \left(\frac{2}{i} u \sqrt{T(1-T)}\right)^n P_n^{(\nu)}(0) e^{i\nu \left(t - i \ln \sqrt{\frac{T}{1-T}}\right)} = \frac{d \left[\xi F \left(\xi, 0, t - i \ln \sqrt{\frac{T}{1-T}} \right) \right]}{d\xi} \quad \xi = \frac{2}{i} u \sqrt{T(1-T)}$$

$$(11) \sum_n \sum_\nu^{II} B_{n\nu} \left(\frac{2}{i} u \sqrt{T(1-T)}\right)^n \frac{dP_n^{(\nu)}(\cos \vartheta)}{d \cos \vartheta} \Big|_{\cos \vartheta=0} e^{i\nu \left(t - i \ln \sqrt{\frac{T}{1-T}}\right)} = \frac{d \left[F \left(\frac{2}{i} u \sqrt{T(1-T)}, \cos \vartheta, t - i \ln \sqrt{\frac{T}{1-T}} \right) \right]}{d \cos \vartheta} \Big|_{\cos \vartheta=0}$$

Aus (9), (10) und (11) folgt nun die Relation (3). Ist F eine algebraische Funktion in x, y, z , so tritt in $\frac{d \left[\xi F \left(\xi, 0, t - i \ln \sqrt{\frac{T}{1-T}} \right) \right]}{d\xi} \quad \xi = \frac{2}{i} u \sqrt{T(1-T)}$

und in $\frac{d \left[F \left(\frac{2}{i} u \sqrt{T(1-T)}, \cos \vartheta, t - i \ln \sqrt{\frac{T}{1-T}} \right) \right]}{d \cos \vartheta} \Big|_{\cos \vartheta=0}$ nicht die Funktion

$\varphi_1 \equiv t - i \ln \sqrt{\frac{T}{T-1}}$ selbst ein, sondern nur die Funktionen $\sin \nu \varphi_1$ und $\cos \nu \varphi_1$. Die beiden Ausdrücke sind demnach algebraische Funktionen von u , e^{it} und T .

Man kann natürlich auch zu anderen Zugeordneten ähnlicher Art, wie die abgeleitete, gelangen, wenn man z. B. eine andere Integraldarstellung der Betafunktion benutzt ⁹⁾.

Zusatz. Eine um die x -Achse rotationssymmetrische Funktion $F(x, y, z)$, die in der Umgebung des Nullpunktes regulär ist, läßt sich in der Form

$$F(x, y, z) = \frac{1}{2\pi} \int_0^{2\pi} F(u, 0, 0) dt$$

darstellen.

Ein anderer Weg (auf den hier kurz eingegangen wird) führt zu einer etwas komplizierten Zugeordneten. Sei F in einem Parallelepipedon mit den Seitenlängen $2A_1, 2A_2, 2A_3$ und mit dem Koordinatenanfangspunkt als Mittelpunkt regulär. $4\pi F$ läßt sich offenbar im Parallelepipedon in der Form

$$\begin{aligned}
 & 4\pi F(x, y, z) \\
 = & \int_{a_1=-A_1}^{a_1=A_1} \int_{a_2=-A_2}^{a_2=A_2} \left[\frac{\frac{\partial [F(A_1, a_2, a_3)]}{\partial A_1}}{\sqrt{(x-A_1)^2 + (y-a_2)^2 + (z-a_3)^2}} - \frac{\frac{\partial [F(-A_1, a_2, a_3)]}{\partial A_1}}{\sqrt{(x+A_1)^2 + (y-a_2)^2 + (z-a_3)^2}} \right] da_2 da_3 \\
 + & \int_{a_1=-A_1}^{a_1=A_1} \int_{a_2=-A_2}^{a_2=A_2} \left[\frac{\frac{\partial [F(a_1, a_2, A_3)]}{\partial A_3}}{\sqrt{(x-a_1)^2 + (y-a_2)^2 + (z-A_3)^2}} - \frac{\frac{\partial [F(a_1, a_2, -A_3)]}{\partial A_3}}{\sqrt{(x-a_1)^2 + (y-a_2)^2 + (z+A_3)^2}} \right] da_1 da_3 \\
 + & \int_{a_1=-A_1}^{a_1=A_1} \int_{a_2=-A_2}^{a_2=A_2} \left[\frac{\frac{\partial [F(a_1, A_2, a_3)]}{\partial A_2}}{\sqrt{(x-a_1)^2 + (y-A_2)^2 + (z-a_3)^2}} - \frac{\frac{\partial [F(a_1, -A_2, a_3)]}{\partial A_2}}{\sqrt{(x-a_1)^2 + (y+A_2)^2 + (z-a_3)^2}} \right] da_1 da_3 \\
 - & \int_{a_1=-A_1}^{a_1=A_1} \int_{a_2=-A_2}^{a_2=A_2} \left[\frac{(F(A_1, a_2, a_3))(x-A_1)}{(\sqrt{(x-A_1)^2 + (y-a_2)^2 + (z-a_3)^2})^3} - \frac{(F(-A_1, a_2, a_3))(x+A_1)}{(\sqrt{(x+A_1)^2 + (y-a_2)^2 + (z-a_3)^2})^3} \right] da_2 da_3 \\
 - & \int_{a_1=-A_1}^{a_1=A_1} \int_{a_2=-A_2}^{a_2=A_2} \left[\frac{(F(a_1, a_2, A_3))(z-A_3)}{(\sqrt{(x-a_1)^2 + (y-a_2)^2 + (z-A_3)^2})^3} - \frac{(F(a_1, a_2, -A_3))(z+A_3)}{(\sqrt{(x-a_1)^2 + (y-a_2)^2 + (z-A_3)^2})^3} \right] da_1 da_3 \\
 - & \int_{a_1=-A_1}^{a_1=A_1} \int_{a_2=-A_2}^{a_2=A_2} \left[\frac{(F(a_1, A_2, a_3))(y-A_2)}{(\sqrt{(x-a_1)^2 + (y-A_2)^2 + (z-a_3)^2})^3} - \frac{(F(a_1, -A_2, a_3))(y+A_2)}{(\sqrt{(x-a_1)^2 + (y+A_2)^2 + (z-a_3)^2})^3} \right] da_1 da_3
 \end{aligned}
 \tag{12}$$

⁹⁾ Es entsteht die wichtige und interessante Frage, ob es nicht möglich ist eine Zugeordnete f von noch einfacherer Bauart zu konstruieren, z. B. ob man zu einem algebraischen F stets ein algebraisch-logarithmisches f herstellen kann. Die algebraischen Zugeordneten f können uns *nicht* die Gesamtheit der algebraischen Potentialfunktionen liefern. Wie aus § 2 und § 4 folgt, sind die Potentialfunktionen mit alge-

darstellen. Nach § 3, S. 653 der Arbeit I können wir (12) auch in der Form

$$\begin{aligned}
 & \frac{1}{2\pi} \int_0^{2\pi} dt \left\{ \int_{a_2=-A_2}^{a_2=A_2} \int_{a_3=-A_3}^{a_3=A_3} \left[\frac{\frac{\partial [F(A_1, a_2, a_3)]}{\partial A_3}}{u - (A_1 + ia_2 \cos t + ia_3 \sin t)} - \frac{\frac{\partial [F(-A_1, a_2, a_3)]}{\partial A_1}}{u - (-A_1 + ia_2 \cos t + ia_3 \sin t)} \right] da_2 da_3 \right. \\
 & + \int_{a_1=-A_1}^{a_1=A_1} \int_{a_2=-A_2}^{a_2=A_2} \left[\frac{\frac{\partial [F(a_1, a_2, A_3)]}{\partial A_3}}{u - (a_1 + ia_2 \cos t + iA_3 \sin t)} - \frac{\frac{\partial [F(a_1, a_2, -A_3)]}{\partial A_3}}{u - (a_1 + ia_2 \cos t - iA_3 \sin t)} \right] da_1 da_2 \\
 & - \int_{a_1=-A_1}^{a_1=A_1} \int_{a_3=-A_3}^{a_3=A_3} \left[\frac{\frac{\partial [F(a_1, A_2, a_3)]}{\partial A_2}}{u - (a_1 + iA_2 \cos t + ia_3 \sin t)} - \frac{\frac{\partial [F(a_1, -A_2, a_3)]}{\partial A_2}}{u - (a_1 - iA_2 \cos t + ia_3 \sin t)} \right] da_1 da_3 \\
 & - \int_{a_2=-A_2}^{a_2=A_2} \int_{a_3=-A_3}^{a_3=A_3} \left[\frac{F(A_1, a_2, a_3)}{(u - (A_1 + ia_2 \cos t + ia_3 \sin t))^2} - \frac{F(-A_1, a_2, a_3)}{(u - (-A_1 + ia_2 \cos t + ia_3 \sin t))^2} \right] da_2 da_3 \\
 & - \int_{a_1=-A_1}^{a_1=A_1} \int_{a_2=-A_2}^{a_2=A_2} \left[\frac{(F(a_1, a_2, A_3)) \sin t}{(u - (a_1 + ia_2 \cos t + iA_3 \sin t))^2} - \frac{(F(a_1, a_2, -A_3)) \sin t}{(u - (a_1 + ia_2 \cos t - iA_3 \sin t))^2} \right] da_1 da_2 \\
 & \left. - \int_{a_1=-A_1}^{a_1=A_1} \int_{a_3=-A_3}^{a_3=A_3} \left[\frac{(F(a_1, A_2, a_3)) \cos t}{(u - (a_1 + iA_2 \cos t + ia_3 \sin t))^2} - \frac{(F(a_1, -A_2, a_3)) \cos t}{(u - (a_1 - iA_2 \cos t + ia_3 \sin t))^2} \right] da_1 da_3 \right\}
 \end{aligned}$$

schreiben, woraus wir zu einer Zugeordneten der algebraischen Funktion F gelangen, die sich als ein doppeltes Integral einer in u, e^{it} und zwei weitere Parameter (a_k) algebraischer Funktion darstellen läßt.

Es soll noch auf eine einfache Überlegung eingegangen werden, die zeigt, daß die Potentialfunktionen mit algebraisch-logarithmischer Zugeordneten eine einfache geometrisch-physikalische Charakterisierung zulassen. In § 3, S. 653 der Arbeit I wurden Punktpole studiert. Stellen wir uns im dreidimensionalen x, y, z -Raume eine offene bzw. geschlossene algebraische Kurve \mathfrak{C} vor, deren Koordinaten durch eine Parameterdarstellung $x = \varphi_1(\tau), y = \varphi_2(\tau), z = \varphi_3(\tau)$ gegeben sind. $\varphi_1, \varphi_2, \varphi_3$ seien algebraische Funktionen von τ ; τ durchlaufe die reellen Werte von $\tau = a$ bis $\tau = b$ ($a < b$), falls die Kurve \mathfrak{C} offen ist, bzw. von $\tau = 0$ bis $\tau = 2\pi$, falls die Kurve \mathfrak{C} geschlossen ist. In jedem Punkt der Kurve sei ein Punktpol (Punktquelle) vorhanden, der die Gestalt

braischen Zugeordneten, die im Endlichen singular sind, im 6-dimensionalen Raume $\Re(x), \Im(x), \Re(y), \Im(y), \Re(z), \Im(z)$ stets mehrdeutig [\Re bedeutet den reellen, \Im den imaginären Teil der eingeklammerten Größe]. Es gibt aber im Endlichen singuläre Potentialfunktionen, die auch bei analytischer Fortsetzung ins Komplexe von x, y, z rational bleiben $\left[\text{z. B. } \frac{x}{x^2 + y^2} \right]$.

$$(14) \left\{ \sum_{s=n}^{s=n} \sum_{k=-(s-1)}^{k=(s-1)} \varphi_{sk}(\tau) \frac{\Phi_{sk}[x-\varphi_1(\tau), y-\varphi_2(\tau), z-\varphi_3(\tau)]}{[\sqrt{[x-\varphi_1(\tau)]^2 + [y-\varphi_2(\tau)]^2 + [z-\varphi_3(\tau)]^2}]^{2s+1}} \right\} d\tau$$

hat, wobei n eine endliche von τ unabhängige ganze Zahl darstellt und alle $\varphi_{sk}(\tau)$ algebraische Funktionen von τ sind. [$\Phi_{sk}(x, y, z) \equiv \text{konst. } r^n P_s^{(k)}(\cos \vartheta) e^{ik\varphi}$ sind die komplexen Kugelfunktionen¹⁰⁾.]

Unter der Voraussetzung, daß die Integration einen Sinn hat, bilden wir das Integral von (14) längs der Kurve \mathfrak{S} ; wir bekommen dann (für einen Punkt x, y, z , der nicht auf der Kurve \mathfrak{S} liegt) für den auf diese Weise zustande gekommenen Linienpol die Darstellung

$$(15) \int_{\tau=a}^{\tau=b} \left\{ \sum_{s=0}^{s=n} \sum_{k=-(s-1)}^{k=s-1} \varphi_{sk}(\tau) \frac{\Phi_{sk}[x-\varphi_1(\tau), y-\varphi_2(\tau), z-\varphi_3(\tau)]}{[\sqrt{[x-\varphi_1(\tau)]^2 + [y-\varphi_2(\tau)]^2 + [z-\varphi_3(\tau)]^2}]^{2s+1}} \right\} d\tau.$$

Nach dem Ergebnis des § 3 der Arbeit I läßt sich für jeden Wert x , der größer als das Maximum von $\varphi_1(\tau)$ [$a \leq \tau \leq b$ bzw. $0 \leq \tau \leq 2\pi$] ist, die Funktion (15) in der Form

$$(16) \int_{\tau=a}^{\tau=b} \left\{ \sum_{s=0}^{s=n} \sum_{k=-(s-1)}^{k=s-1} \frac{1}{2\pi} \int_0^{2\pi} \frac{\varphi_{sk}(\tau) e^{ikt}}{[u - (\varphi_1(\tau) + i\varphi_2(\tau) \cos t + i\varphi_3(\tau) \sin t)]^s} dt \right\} d\tau$$

schreiben. Da man für jeden Wert x, y, z [mit $x > \text{Max. von } \varphi_1(\tau)$] die Integrationsfolge vertauschen darf, so ist (16) ebenfalls in der Form

$$(17) \frac{1}{2\pi} \int_0^{2\pi} dt \left\{ \int_{\tau=a}^{\tau=b} \left[\sum_{s=0}^{s=n} \sum_{k=-(s-1)}^{k=s-1} \frac{\varphi_{sk}(\tau) e^{ikt}}{[u - (\varphi_1(\tau) + i\varphi_2(\tau) \cos t + i\varphi_3(\tau) \sin t)]^s} \right] d\tau \right\}$$

darstellbar. Ist nun $\sum_{s=0}^{s=n} \sum_{k=-(s-1)}^{k=s-1} \frac{\varphi_{sk}(\tau) e^{ikt}}{[u - (\varphi_1(\tau) + i\varphi_2(\tau) \cos t + i\varphi_3(\tau) \sin t)]^s}$ eine

Funktion, die sich mit Hilfe der algebraisch-logarithmischen Funktionen ausintegrieren läßt, so gelangen wir ebenfalls zu Potentialfunktionen mit algebraisch-logarithmischer Zugeordneten.

Die Formel (1) läßt schließlich eine geometrische Deutung zu. Bezeichnet man mit $f^*[(x + iy \cos t + iz \sin t), e^{it}]$ die Funktion $f(u, e^{it})$, um anzudeuten, daß man sie im x, y, z -Raume für ein konstantes t betrachtet, so bedeutet der Übergang von f zu f^* eine Abbildung der komplexen Ebene auf den dreidimensionalen Raum, wobei jedem Punkte der komplexen u -Ebene eine Gerade des x, y, z -Raumes entspricht. Da die Funktion F_q durch die Mittelwertbildung aus f^* erhalten wird, so können wir in

¹⁰⁾ Wegen hier eingeführten Bezeichnungen für die Kugelfunktionen vgl. Fußnote ⁵⁾ S. 645 der Arbeit I.

gewissem Sinne von einer Pseudo-Abbildung im kleinen des zweidimensionalen auf den dreidimensionalen Raum sprechen.

Wir wollen uns an dieser Stelle mit den beiden Andeutungen auf die Beziehungen von geometrischem Charakter begnügen, da wir an einer anderen Stelle die Potentialfunktionen mit algebraisch-logarithmischer Zugeordneten einer genaueren Untersuchung in dieser Hinsicht unterziehen werden.

Setzen wir $F_q(x, y, z)$ in dem Raum (x, y, z) analytisch fort, so bleibt die Beziehung (1) so lange bestehen, als $f(u, e^{iu})$ noch für alle Werte t des Integrationsintervalles regulär bleibt.

Führt man die neue Variable

$$(18) \quad \zeta = e^{it}$$

ein, so kann man f bei festem x, y, z als eine Funktion der einen komplexen Veränderlichen ζ betrachten. Die in (1) angegebene Integration bedeutet dann die Integration längs des Einheitskreises. Wie man sofort sieht, ist (1) ein spezieller Fall der folgenden allgemeineren Darstellung der harmonischen Funktionen

$$(19) \quad F_q = \frac{1}{2\pi i} \int_a f \left[\left(\frac{iy-z}{2} \frac{1}{\zeta} + x + \frac{iy+z}{2} \zeta \right), \zeta \right] \frac{d\zeta}{\zeta},$$

wobei die Integration über die geschlossene oder offene Kurve a der (evtl. mehrdeutigen) ζ -Ebene erstreckt ist. In Übereinstimmung mit Früherem wird das Gebiet des x, y, z -Raumes, wo die Beziehung (3) richtig bleibt, als Zuordnungsgebiet $G_q(a)$ von F_q bezeichnet; $G_q(a)$ ist natürlich von dem Integrationsweg a abhängig. Fällt der Weg a speziell mit der Peripherie des Einheitskreises zusammen, so werden wir ihn mit e und dementsprechend das Gebiet des x, y, z -Raumes, wo die Beziehung (1) gilt mit $G_q(e)$ bezeichnen (vgl. Fig. 2).

Jede Funktion einer komplexen Veränderlichen $f(u, e^{iu})$ gibt Veranlassung zur Bildung einer Reihe von Potentialfunktionen $F_1, F_2, F_3, F_4, \dots$, indem man für jeden regulären Punkt von $f(u, e^{iu})$ durch die in (1) bzw. (3) Mittelwertbildung F_q bildet (vgl. dazu S. 649 und 650 der Arbeit I). In der Arbeit I wurde die Vermutung ausgesprochen, daß zwischen den auf diese Weise erhaltenen Potentialfunktionen F_q gewisse einfache Beziehungen existieren. Wie ferner in der Arbeit I (S. 646) ausführlich besprochen ist, läßt sich die durch (1) ausgedrückte Beziehung zwischen f und F_q durch passende Normierung der f , im kleinen ein-eindeutig umkehrbar machen. Daher liegt es nahe, den durch (1) bzw. (3) ausgedrückten Zusammenhang zwischen den dreidimensionalen Potentialfunktionen F_q und den Funktionen f einer komplexen Veränderlichen dazu in der Weise benutzen, daß man die bekannte Einteilung der letzteren in

rationale, algebraische, ..., deren Integrale usw. auch als Ausgangspunkt für die Klassifizierung der ersten wählt. Es entsteht natürlich die Frage, ob man die auf diese Weise erhaltenen Klassen von Potentialfunktionen auf irgendeine andere Art definieren kann, ohne dabei die Integraldarstellung (1) bzw. (19) zu Hilfe nehmen zu müssen.

Die harmonischen Funktionen F_q , die aus einem ganzen rationalen f , das nur Pole von der Form

$$\frac{e^{ikt}}{[u + A + iB \cos t + iC \sin t]^q}$$

$[A, B, C, q, k$ reell, k, q ganz] besitzt, hergeleitet sind, wurden in der Arbeit I erledigt. Dort wurde gezeigt, daß diese F_q bei $|k| < q$ nur Punktpole endlicher Ordnung (vgl. S. 653) und bei $|k| \geq q$ nur Linienpole¹¹⁾ (vgl. S. 654) besitzen können und sich sonst im ganzen x, y, z -Raume regulär verhalten. In § 2 werden wir zu der allgemeinen Untersuchung von Potentialfunktionen mit rationalen Zugeordneten übergehen.

§ 2.

Algebraische Potentialfunktionen mit rationalen Zugeordneten: ihre Darstellung und Singularitäten.

In dem vorliegenden Paragraph werden wir Funktion f von der Form

$$(1) \quad f \equiv \frac{p_1(u, e^{it})}{p_2(u, e^{it})} e^{it[\varrho + m' + n' - m - n + 1]}$$

betrachten, wo

$$(2) \quad p_1(u, e^{it}) \equiv \sum_{k=0}^{k=n'} u^k \left(\sum_{v=-(m'+n'-k)}^{v=(m'+n'-k)} B_v^{(n'-k)} e^{ivt} \right)$$

und

$$(3) \quad p_2(u, e^{it}) \equiv \sum_{k=0}^{k=n} u^k \left(\sum_{v=-(m+n-k)}^{v=(m+n-k)} A_v^{(n-k)} e^{ivt} \right)$$

Polynome in u und $e^{\pm it}$ sind. p_1 und p_2 sollen keinen gemeinschaftlichen Teiler haben; ferner soll $p_2 \cdot e^{i(m+n)t}$ keinen Teiler von der Form $(e^{it} - C)$ besitzen¹²⁾, wo C eine von x, y, z unabhängige von 0 verschiedene Konstante ist. q ist eine ganze Zahl, die wir zwecks Vereinfachung weiterer Formeln schon jetzt einführen.

¹¹⁾ Bei der Angabe eines Beispielles für die Linienpole ist in der Arbeit I ein

Druckfehler; es soll sein $\frac{e^{i\varphi} \operatorname{tg} \frac{k}{2} \frac{\vartheta}{2}}{R}$ und nicht $\frac{e^{i\varphi} \operatorname{tg} \frac{k}{2} \frac{\vartheta}{2}}{R}$. Ferner S. 654 Zeile 2

von unten und S. 655 Zeile 2 von oben soll es $s = -1$ und nicht $s = 1$ heißen.

¹²⁾ Diese Voraussetzung ist — wie wir im folgenden zeigen werden — unwesentlich.

Führen wir die Transformation

$$(4) \quad \zeta = e^{it}$$

in $p_2(u, e^{it})e^{i(m+n)t}$ aus, so bekommen wir ein Polynom $P_2(\zeta)$, $2(m+n)$ -ten Grades in ζ , dessen Koeffizienten Polynome in x, y, z sind. $P_2(\zeta)$ läßt sich in der Form

$$(5) \quad P_2(\zeta) = \zeta_0 \prod_{k=1}^{k=2(m+n)} (\zeta - \zeta_k)$$

schreiben, wo ζ_k [$k = 1, 2, 3, \dots, 2(m+n)$] algebraische Funktionen in x, y, z sind; ζ_0 ist in x, y, z rational. Die Diskriminante

$$(6) \quad D(P_2) \equiv \prod_{s \neq k} (\zeta_s - \zeta_k) \quad \left(\begin{array}{l} s = 1, 2, 3, \dots, 2m + 2n \\ k = 1, 2, 3, \dots, 2m + 2n \end{array} \right)$$

wird eine rationale Funktion in x, y, z , die evtl. identisch verschwinden kann. Das Ergebnis unserer ersten Betrachtung lautet:

I. Fall: $D(P_2)$ verschwindet nicht identisch. Ist f die durch Formel (4) definierte Funktion und F_q die auf Grund der Integralbeziehung (3) aus f (bei entsprechender Wahl des Integrationsweges α) hergeleitete Potentialfunktion, so ist F_q gleich

$$(7) \quad F_q = \frac{p_1 \left[\left(x + \frac{iy-z}{2} \frac{1}{\zeta_q} + \frac{iy+z}{2} \zeta_q \right), \zeta_q \right] \zeta_q^{2m'+n'}}{\zeta_0 \frac{\partial}{\partial \zeta_q} \left[p_2 \left(\left(x + \frac{iy-z}{2} \frac{1}{\zeta_q} + \frac{iy+z}{2} \zeta_q \right), \zeta_q \right) \zeta_q^{m+n} \right]} + R(x, y, z),$$

wo $R(x, y, z)$ eine in x, y, z rationale Funktion bedeutet und ζ_q eine Wurzel der Gleichung

$$(8) \quad p_2 \left[\left(x + \frac{iy-z}{2} \frac{1}{\zeta_q} + \frac{iy+z}{2} \zeta_q \right), \zeta_q \right] \zeta_q^{m+n} = 0$$

($q = 1, 2, 3, \dots, 2m + 2n$)

ist. $R(x, y, z)$ ist für $\varrho \geq 0$ stets gleich Null, und für $\varrho < 0$ gleich

$P_2(0)$	0	0	0	$P_1(0)$
$\left[\frac{\partial P_2(\zeta)}{\partial \zeta} \right]_{\zeta=0}$	$P_2(0)$	0	0	$\left[\frac{\partial P_1(\zeta)}{\partial \zeta} \right]_{\zeta=0}$
$\left[\frac{\partial^2 P_2(\zeta)}{\partial \zeta^2} \right]_{\zeta=0}$	$1! \binom{2}{1} \left[\frac{\partial P_2(\zeta)}{\partial \zeta} \right]_{\zeta=0}$	$2! \binom{2}{2} P_2(0)$	0	$\left[\frac{\partial^2 P_1(\zeta)}{\partial \zeta^2} \right]_{\zeta=0}$
$\left[\frac{\partial^{(\varrho -1)} P_2(\zeta)}{\partial \zeta^{(\varrho -1)}} \right]_{\zeta=0}$	$1! \binom{ \varrho -1}{1} \left[\frac{\partial^{(\varrho -2)} P_2(\zeta)}{\partial \zeta^{(\varrho -2)}} \right]_{\zeta=0}$	$2! \binom{ \varrho -1}{2} \left[\frac{\partial^{(\varrho -3)} P_2(\zeta)}{\partial \zeta^{(\varrho -3)}} \right]_{\zeta=0}$	0	$\left[\frac{\partial^{(\varrho -1)} P_1(\zeta)}{\partial \zeta^{(\varrho -1)}} \right]_{\zeta=0}$
$1 \cdot 1! \cdot 2! \dots (\varrho -1)! [P_2(0)]^{ \varrho }$				

$\left[\frac{\partial^s P_2(\zeta)}{\partial \zeta^s} \right]_{\zeta=0}$ ist dabei

	$A_{-(m+n-k)}^{(n-k)}$	$A_{-(m+n-k-1)}^{(n-k)}$	$A_{-(m+n-k-2)}^{(n-k)}$	$A_{-(m+n-k-3)}^{(n-k)}$
	$-s \left(\frac{iy-z}{2} \right)$	$(k-s+1)x$	$(2k-s+2) \left(\frac{iy+z}{2} \right)$	0
	0	$(-s+1) \left(\frac{iy-z}{2} \right)$	$(k-s+2)x$	$(2k-s+3) \left(\frac{iy+z}{2} \right)$
$\sum_{k=0}^{k=n}$	0	0	$(-s+2) \left(\frac{iy-z}{2} \right)$	$(k-s+3)x$

	0	0	0	0
	0	0	0	0

$\left[\frac{\partial^s P_1(\zeta)}{\partial \zeta^s} \right]_{\zeta=0}$ erhält man aus den eben angegebenen Ausdruck, wenn man in diesem die $A_{-(m+n-k-l)}^{(n-k)}$ durch $B_{-(m'+n'-k-l)}^{(n'-k)}$ ersetzt und die Summation nicht von $k=0$ bis $k=n$, sondern bis $k=n'$ erstreckt.

II. Fall. $D(P_2)$ verschwindet identisch.

Es muß dann

$$P_2(\zeta) = [P_3(\zeta)]^{\beta_3} [P_4(\zeta)]^{\beta_4} \dots [P_s(\zeta)]^{\beta_s}$$

sein, wo P_3, P_4, \dots, P_s verschiedene Polynome von derselben Gestalt wie P_2 sind und mindestens ein β_k größer wie 1 ist. Wir bezeichnen mit $P^{(k)}(\zeta)$ [$k=3, 4, \dots, s$] das Polynom $\frac{P_2(\zeta)}{[P_k(\zeta)]^{\beta_k}}$ bzw. $\frac{P_2(\zeta) \zeta^{|\epsilon|}}{[P_k(\zeta)]^{\beta_k}}$, falls ϵ negativ ist.

Unter den gleichen Bedingungen wie im Fall I ist dann F_q gleich

$P^{(q)}(\zeta_q)$	0	0	$P_1(\zeta_q) \zeta_q^{\epsilon^*}$
$\frac{\partial P^{(q)}(\zeta_q)}{\partial \zeta_q}$	$\binom{1}{1} P^{(q)}(\zeta_q)$	0	$\frac{\partial [P_1(\zeta_q) \zeta_q^{\epsilon^*}]}{\partial \zeta_q}$
$\frac{\partial^2 P^{(q)}(\zeta_q)}{\partial \zeta_q^2}$	$\binom{2}{1} \frac{\partial P^{(q)}(\zeta_q)}{\partial \zeta_q}$	$2! \binom{2}{2} P^{(q)}(\zeta_q)$	$\frac{\partial^2 [P_1(\zeta_q) \zeta_q^{\epsilon^*}]}{\partial \zeta_q^2}$
$\frac{\partial^{(\beta_q-1)} P^{(q)}(\zeta_q)}{\partial \zeta_q^{\beta_q-1}}$	$\binom{\beta_q-1}{1} \frac{\partial^{(\beta_q-2)} P^{(q)}(\zeta_q)}{\partial \zeta_q^{\beta_q-2}}$	$2! \binom{\beta_q-1}{2} \frac{\partial^{(\beta_q-3)} P^{(q)}(\zeta_q)}{\partial \zeta_q^{\beta_q-3}}$	$\frac{\partial^{(\beta_q-1)} [P_1(\zeta_q) \zeta_q^{\epsilon^*}]}{\partial \zeta_q^{\beta_q-1}}$

10) $F_q = \frac{\dots}{1 \cdot 1! \cdot 2! \cdot \dots \cdot (\beta_q - 1)! [P^{(q)}(\zeta_q)]^{\beta_q}} +$

$+ R(x, y, z),^{13)}$

¹³⁾ Es wäre natürlich möglich, wie in (9) die $P(\zeta_q)$ durch die $A_k^{(n)}$ auszudrücken; die so entstehende Formel ist aber sehr lang; wir verzichten daher auf ihre Wiedergabe.

$$\begin{array}{ccc|c}
 A_{-(m+n-k-4)}^{(n-k)} & \cdot & \cdot & A_{-(m+n-k-s+1)}^{(n-k)} & A_{-(m+n-k-s)}^{(n-k)} \\
 0 & \cdot & \cdot & 0 & 0 \\
 0 & \cdot & \cdot & 0 & 0 \\
 2k-s+4 \left(\frac{iy+z}{2}\right) & \cdot & \cdot & 0 & 0 & \left(\frac{iy-z}{2}\right)^{k+s} \\
 \cdot & \cdot & \cdot & \cdot & \cdot & \\
 \cdot & \cdot & \cdot & \cdot & \cdot & \\
 0 & \cdot & \cdot & (k-1)x & 2k\left(\frac{iy+z}{2}\right) \\
 0 & \cdot & \cdot & -\left(\frac{iy-z}{2}\right) & kx
 \end{array}$$

wobei ζ_q voneinander verschiedene Wurzeln der Gleichung (8) sind, $R(x, y, z)$ die frühere Bedeutung hat und $\varrho^* = \varrho$, wenn $\varrho > 0$, und $\varrho^* = 0$, wenn $\varrho \leq 0$ ist.

Man sieht somit in beiden Fällen, daß alle Funktionen F_q , die man aus (4) durch die Integration (1) bzw. (3) bei beliebigen x, y, z erhält, Zweige ein und derselben algebraischen Funktion sind. Auf eine Ausnahme von dieser letzten Behauptung haben wir schon in der Vorrede (vgl. S. 632) hingewiesen.

Ist α speziell e , so liefert uns F_q eine Summe von Ausdrücken der Form (7) bzw. (10), wobei die Summation über diejenigen ζ_q zu nehmen ist, deren Betrag für das betreffende (x, y, z) (für die man die Integration durchführt) kleiner als 1 ist.

Beweis. In dem Ausdruck (1)

$$\frac{p_1 \left[\left(x + \frac{iy-z}{2} e^{-it} + \frac{iy+z}{2} e^{it} \right), e^{it} \right] e^{it(m'+n'+\varrho+1)}}{p_2 \left[\left(x + \frac{iy-z}{2} e^{-it} + \frac{iy+z}{2} e^{it} \right), e^{it} \right] e^{it(m+n)}}$$

führen wir die Transformation (4) aus. Wir erhalten dann einen Ausdruck

$$(11) \quad \frac{P_1(\zeta)}{P_2(\zeta)} \zeta^{e+1},$$

wobei P_1 und P_2 Polynome in ζ sind, deren Koeffizienten Funktionen in x, y, z sind.

Wir wollen nunmehr zeigen, daß (11) (abgesehen von dem ausgeschlossenen Fall, daß $P_2(\zeta)$ einen Faktor von der Form $(\zeta - C)$ hat) für gewisse x, y, z stets integrierbar ist.

Wir werden uns zunächst auf den Fall beschränken, daß $u = e$ ist, und zeigen: Gibt es zu jedem x, y, z einen Punkt des Einheitskreises e , so daß $P_2(\zeta)$ verschwindet, so besitzt $P_2(\zeta)$ notwendigerweise einen Faktor von der Form $(\zeta - C)$ [$|C| = 1$]. $P_2(\zeta)$ ist gleich

$$\sum_{k=0}^{k=n} \left(x\zeta + \frac{iy-z}{2} + \frac{iy+z}{2} \zeta^2 \right)^k \left(\sum_{\nu=0}^{\nu=2(m+n-k)} \zeta^\nu A_{\nu-(m+n-k)}^{(n-k)} \right).$$

Verschwindet nun $P_2(\zeta)$ für jeden Wert x, y, z , so wird der Koeffizient der höchsten Potenz von $P_2(\zeta)$, also der Ausdruck $\sum_{\nu=0}^{\nu=2m} \zeta^\nu A_{\nu-m}^{(0)}$, einen Faktor der Form $(\zeta - C)$ besitzen, wo C ein Punkt des Integrationsintervalles ist. Wäre das nicht der Fall, so müßte eine untere positive Schranke M_0 existieren, derart daß $\left| \sum_{\nu=0}^{\nu=2m} A_{\nu-m}^{(0)} \zeta^\nu \right|$ größer als M_0 wäre für jeden Punkt ζ des Integrationsweges. Dann wäre aber $P_2(\zeta)$ in der Umgebung eines Punktes $x = x_1, y = 0, z = 0$ (wo

$$x_1 > \frac{\sum_{\nu=0}^{\nu=2(m+1)} |A_{\nu-(m+1)}^{(1)}| + \sum_{\nu=0}^{\nu=2(m+2)} |A_{\nu-(m+2)}^{(2)}| + \dots + \sum_{\nu=0}^{\nu=2(m+k)} |A_{\nu-(m+k)}^{(n)}|}{M_0} > 1$$

ist) bestimmt von Null verschieden, was unserer Annahme widerspricht.

Nehmen wir nun an, daß die Gleichung $\sum_{\nu=0}^{\nu=2m} A_{\nu-m}^{(0)} \zeta^\nu = 0$ eine Anzahl von Wurzeln besitzt, die auf dem Integrationswege liegen, etwa $\zeta^{(1)}, \zeta^{(2)}, \dots, \zeta^{(r)}$. Ist $P_2(\zeta)$ für jedes x, y, z irgendwo auf dem Integrationswege gleich 0, so muß auch der Koeffizient von u^{n-1} , d. h. $\sum_{\nu=0}^{\nu=2m+2} A_{\nu-(m+1)}^{(1)} \zeta^\nu$ mindestens einen gemeinschaftlichen Teiler $\zeta - \zeta^{(q)}$ mit $\sum_{\nu=0}^{\nu=2m} A_{\nu-m}^{(0)} \zeta^\nu$ haben. Wäre es nicht der Fall, so hätten die Zahlen $\left| \sum_{\nu=0}^{\nu=2m+2} A_{\nu-(m+1)}^{(1)} \zeta^{(q)\nu} \right|$ [$q = 1, 2, 3, \dots, r$] eine kleinste unter ihnen, die wir mit M_1 bezeichnen, wobei $M_1 > 0$ wäre. Es wäre dann $P_2(\zeta)$ in der Umgebung eines Punktes $x = x_2, y = z = 0$ gewiß nicht Null, wenn

$$|x_2| > |x_1| > \frac{\sum_{\nu=0}^{\nu=2(m+2)} |A_{\nu-(m+2)}^{(2)}| + \sum_{\nu=0}^{\nu=2(m+3)} |A_{\nu-(m+3)}^{(3)}| + \dots + \sum_{\nu=0}^{\nu=2(m+n)} |A_{\nu-(m+n)}^{(n)}|}{M_1}$$

ist; und dies widerspricht der Annahme. Genau auf dieselbe Weise kann man zeigen, daß alle Koeffizienten $\sum_{\nu=0}^{\nu=2(m+k)} A_{\nu-(m+k)}^{(k)} \zeta^\nu$ mindestens einen gemeinsamen Faktor von der Form $(\zeta - C)$ besitzen müssen. Durch eine

einfache Veränderung des Beweisganges (auf die wir nicht eingehen) läßt sich zeigen, daß unsere Behauptung richtig bleibt, wenn der Integrationsweg α eine geschlossene, ganz im Endlichen gelegene Jordansche Kurve ist.

Im ersten Falle müssen wir zwei Unterfälle unterscheiden, und zwar

1. $2(n + m) > (2n' + 2m' + \varrho)$,
2. $2(n + m) \leq (2n' + 2m' + \varrho)$.

1. Im ersten Unterfalle ist:

$$(12) \quad \frac{P_1(\zeta) \zeta^\varrho}{P_2(\zeta)} = \sum_{q=1}^{q=2(n+m)} \frac{\zeta_q^\varrho P_1(\zeta_q)}{\frac{\partial P_2(\zeta_q)}{\partial \zeta_q}} \frac{1}{\zeta - \zeta_q} + R^*(\zeta; x, y, z),$$

wo $R^*(\zeta; x, y, z) = 0$ ist, wenn $\varrho \geq 0$, $D(P_2) \neq 0$, ist, und

$$(13) \quad R^*(\zeta; x, y, z) = \frac{P_1(0)}{P_2(0)} \frac{1}{\zeta^{|\varrho|}} + \frac{|P_2(0) P_1(0)|}{1.1!} \frac{1}{\zeta^{|\varrho|-1}} + \dots + R(x, y, z) \frac{1}{\zeta},$$

wenn $\varrho < 0$ ist. ζ_q sind dabei die früher angegebenen Wurzeln der Gleichung $P_2(\zeta) = 0$.

Führen wir nun die Integration

$$(14) \quad \frac{1}{2\pi i} \int_{\alpha} \frac{P_1(\zeta) \zeta^\varrho}{P_2(\zeta)} d\zeta$$

(für einen bestimmten Wert x, y, z) durch, so erhalten wir das Integrationsresultat

$$(15) \quad \sum \frac{\zeta_q^\varrho P_1(\zeta_q)}{\frac{\partial P_2(\zeta_q)}{\partial \zeta_q}},$$

wobei die Summation über sämtliche in α befindliche Wurzeln ζ_q zu nehmen ist. Liegt der Nullpunkt im Innern von α , so kommt zu (15) noch $R(x, y, z)$ hinzu.

2. $(2n' + 2m' + \varrho) \geq (2n + m)$.

Ist $\varrho \geq 0$, so läßt sich in diesem Falle

$$(16) \quad \frac{P_1(\zeta) \zeta^\varrho}{P_2(\zeta)} = h(\zeta) + \frac{[P_1(\zeta) \zeta^\varrho - h(\zeta) P_2(\zeta)]}{P_2(\zeta)}$$

setzen, wo $h(\zeta)$ eine ganze rationale Funktion in ζ ist, und der Grad von $[P_1(\zeta) \zeta^\varrho - h(\zeta) P_2(\zeta)]$ kleiner wie $2(n + m)$ ist.

Ist $\varrho < 0$, so ist

$$(17) \quad \frac{P_1(\zeta)}{\zeta^{|\varrho|} P_2(\zeta)} = h(\zeta) + \frac{P_1(\zeta) - \zeta^{|\varrho|} h(\zeta) P_2(\zeta)}{\zeta^{|\varrho|} P_2(\zeta)}$$

Integriert man (16) bzw. (17) längs eines Weges α , so erhält man dieselbe Formel (15), wie im ersten Unterfalle.

II. Fall: $D(P_2) \equiv 0$.

Wir müssen wiederum zwei Unterfälle unterscheiden:

1. $2(n + m) > (2n' + 2m' + \varrho)$,
2. $2(n + m) \leq (2n' + 2m' + \varrho)$.

Im ersten Unterfalle ist:

$$\begin{aligned}
 \frac{P_1(\zeta)\zeta^\varrho}{P_2(\zeta)} &= \sum_{q=3}^{q=\beta} \left\{ \frac{P_1(\zeta_q)\zeta_q^{\varrho^*}}{P^{(q)}(\zeta_q)} \frac{1}{(\zeta - \zeta_q)^{\beta_q}} \right. \\
 &+ \frac{\left[\begin{array}{cc} P^{(q)}(\zeta_q) & P_1(\zeta_q)\zeta_q^{\varrho^*} \\ \frac{\partial P^{(q)}(\zeta_q)}{\partial \zeta_q} & \frac{\partial [P_1(\zeta_q)\zeta_q^{\varrho^*}]}{\partial \zeta_q} \end{array} \right]}{[P^{(q)}(\zeta_q)]^2} \frac{1}{(\zeta - \zeta_q)^{\beta_q - 1}} \\
 &+ \frac{\left[\begin{array}{ccc} P^{(q)}(\zeta_q) & 0 & P_1(\zeta_q)\zeta_q^{\varrho^*} \\ \frac{\partial P^{(q)}(\zeta_q)}{\partial \zeta_q} & P^{(q)}(\zeta_q) & \frac{\partial [P_1(\zeta_q)\zeta_q^{\varrho^*}]}{\partial \zeta_q} \\ \frac{\partial^2 P^{(q)}(\zeta_q)}{\partial \zeta_q^2} & \frac{\partial P^{(q)}(\zeta_q)}{\partial \zeta_q} & \frac{\partial^2 [P_1(\zeta_q)\zeta_q^{\varrho^*}]}{\partial \zeta_q} \end{array} \right]}{1 \cdot 1! \cdot 2! [P^{(q)}(\zeta_q)]^3} \frac{1}{(\zeta - \zeta_q)^{\beta_q - 2}} + \dots \\
 &+ \left. \frac{\left[\begin{array}{ccc} P^{(q)}(\zeta_q) & 0 & 0 & P_1(\zeta_q)\zeta_q^{\varrho^*} \\ \frac{\partial P^{(q)}(\zeta_q)}{\partial \zeta_q} & \binom{1}{1} P^{(q)}(\zeta_q) & 0 & \frac{\partial [P_1(\zeta_q)\zeta_q^{\varrho^*}]}{\partial \zeta_q} \\ \frac{\partial^2 P^{(q)}(\zeta_q)}{\partial \zeta_q^2} & \binom{2}{1} \frac{\partial P^{(q)}(\zeta_q)}{\partial \zeta_q} & 2! \binom{2}{2} P^{(q)}(\zeta_q) & \frac{\partial^2 [P_1(\zeta_q)\zeta_q^{\varrho^*}]}{\partial \zeta_q} \\ \frac{\partial^{(\beta_q - 1)} P^{(q)}(\zeta_q)}{\partial \zeta_q^{\beta_q - 1}} & \binom{\beta_q - 1}{1} \frac{\partial^{(\beta_q - 2)} P^{(q)}(\zeta_q)}{\partial \zeta_q^{\beta_q - 2}} & 2! \binom{\beta_q - 1}{2} \frac{\partial^{(\beta_q - 3)} P^{(q)}(\zeta_q)}{\partial \zeta_q^{\beta_q - 3}} & \frac{\partial^{(\beta_q - 1)} [P_1(\zeta_q)\zeta_q^{\varrho^*}]}{\partial \zeta_q^{\beta_q - 1}} \end{array} \right]}{1 \cdot 1! \cdot 2! \cdot (\beta_q - 1)! [P^{(q)}(\zeta_q)]^{\beta_q}} \frac{1}{(\zeta - \zeta_q)} \right\} \\
 &+ R^*(\zeta, x, y, z),
 \end{aligned}
 \tag{18}$$

wobei die Summation über sämtliche voneinander verschiedene Wurzeln der Gleichung (8) zu nehmen ist.

Führen wir die Integration von (14) durch, so erhalten wir als Endresultat eine Summe von Ausdrücken (10), wobei die Summation wieder über diejenige ζ_q zu nehmen ist, die in α liegen.

Auf den zweiten Unterfall wollen wir nicht ausführlicher eingehen, da die Rechnung genau wie in früheren Fällen verläuft.

Wir wollen zeigen, daß man zu jedem Punkt x_0, y_0, z_0 des Regularitätsbereiches eine Kurve α angeben kann derart, daß in der Umgebung von x_0, y_0, z_0 das Integral (3) einen vorgegebenen Zweig F_q liefert. Wir schließen die Punkte x_0, y_0, z_0 für die im Falle I: $D(P_2) = 0$ und im Falle II: $D(P_3, P_4, \dots, P_s) = 0$ ist, aus unserer Betrachtung aus, da sie, wie wir am Schluß des § 2 zeigen werden, im allgemeinen singulär sind. Betrachten wir unsere Funktion f^{++} ¹⁴⁾ im fünfdimensionalen Raume mit den Koordinaten: $x, y, z, \Re(\zeta), \Im(\zeta)$.

Für den Punkt x_0, y_0, z_0 werden die Werte unserer Funktion f^{++} die gesamte komplexe Ebene durchlaufen, wobei den Polen von $f^+(x_0, y_0, z_0, \zeta)$ [also Nullstellen von P_2] $(2n + 2m)$ Punkte der ζ -Ebene entsprechen (die wir in der Zeichnung 3 markieren). Läßt man x_0, y_0, z_0 variieren, so beschreiben die Nullpunkte von P_2 gewisse Raumstücke im fünfdimensionalen Raume. (In der Zeichnung sind sie als Kurven angedeutet.)

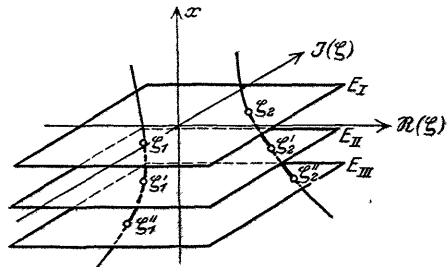


Fig. 3.

Da $D(P_2) \neq 0$ [bzw. $D(P_3, P_4, \dots, P_s) \neq 0$] an der Stelle x_0, y_0, z_0 ist, und die Nullstellen eines Polynomes stetige Funktionen der Koeffizienten sind; so werden diejenigen Teile der auf diese Weise erhaltenen $(2n + 2m)$ Räume keine gemeinsamen Punkte haben, wenn man sich auf eine genügend kleine Umgebung im x, y, z -Raume des Punktes x_0, y_0, z_0 beschränkt.

Hiermit ist die behauptete Darstellung von F_q in der Umgebung von x_0, y_0, z_0 bewiesen.

Aus unseren Betrachtungen sieht man auch, daß die früher gemachte Voraussetzung, daß $P_2(\zeta)$ keinen Faktor von der Form $(\zeta - C)$ hat, nicht wesentlich ist. Es darf natürlich, falls $P_2(\zeta)$ einen solchen Faktor besitzt, der Wert C auf dem Integrationswege α nicht liegen. Wir werden aber die gemachte Voraussetzung dennoch beibehalten.

In den Formeln (7) und (10) haben wir eine allgemeine Darstellung für die harmonischen Funktionen mit rationaler Zugeordneten gewonnen:

¹⁴⁾ Das Zeichen ++ bei f bedeutet, daß wir sie jetzt als Funktion von $x, y, z, \Re(\zeta)$ und $\Im(\zeta)$ betrachten.

Sie ist durch diese Formeln in ihrem ganzen Regularitätsbereiche definiert. Offenbar ist sie überall harmonisch.

Sei ζ_q eine Wurzel der Gleichung (8). Wir wollen feststellen, in welchem Raumteile $G_q(\alpha)$ des (x, y, z) -Raumes bei der Integration längs der (geschlossenen) Kurve α wir denjenigen Zweig F_q bekommen, welcher der Wurzel ζ_q entspricht.

$a_1(t', t'') + i a_2(t', t'')$ sei das Innere der Kurve α . Durch das Gleichungssystem

$$(19) \quad \Re(\zeta_q) = a_1(t', t''), \quad \Im(\zeta_q) = a_2(t', t'') \quad (t', t'' \text{ reell, } t'^2 + t''^2 \leq 1)$$

(wo \Re den reellen, \Im den imaginären Teil der eingeklammerten Wurzel ζ_q

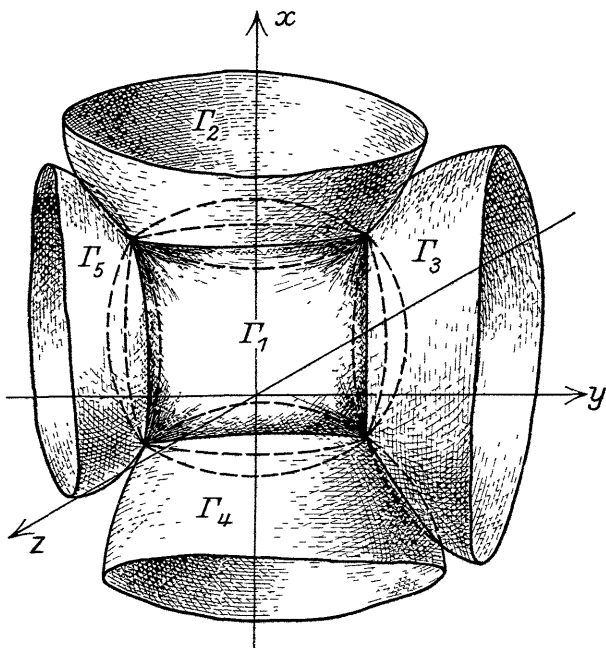


Fig. 4.

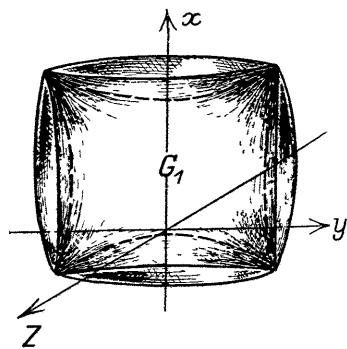


Fig. 5.

bedeutet), ist uns ein Raumteil $G_q(\alpha)$ des x, y, z -Raumes gegeben. Dabei kann es natürlich vorkommen, daß $G_q(\alpha)$ von den in gleicher Weise für die übrigen Wurzeln ζ_p ($p \neq q$) erhaltenen Raumteile $G_p(\alpha)$ teilweise überdeckt wird. In dem von diesen Überdeckungen freien Teile von $G_q(\alpha)$, welcher der auf der Seite 643 definierte Zuordnungsraum $G_q(\alpha)$ ist, wird die Integration (3) den ζ_q entsprechenden Zweig ergeben. Ist ein Punkt des Raumes von mehreren $G_p(\alpha)$ überdeckt, so wird die Integration (3) uns eine Summe der Ausdrücke (7) bzw. (10) für die betreffende ζ_p geben.

Wie früher gezeigt wurde, kann man a stets so wählen, daß $G_q(a)$ ein von Null verschiedenes Volumen hat. Ist $a = e$, so lautet die Ungleichung für $\Gamma_q(e)$

$$(20) \quad [\Re(\zeta_q)]^2 + [\Im(\zeta_q)]^2 < 1.$$

Ist k ein-eindeutig q zugeordnet, derart, daß wenn q sämtliche ganze Zahlen von 1 bis $2(n+m)$ durchläuft, k dieselbe Zahlenfolge in einer anderen Reihenfolge durchläuft, so ist im Falle I

$$(21) \quad f_n = \sum_{q=1}^{q=2(n+m)} \frac{\zeta_k^q P_1(\zeta_k)}{\frac{\partial P_2(\zeta_k)}{\partial \zeta_k}} \frac{1}{\zeta - \zeta_q} + R^*(\zeta; x, y, z)$$

eine rationale Funktion in x, y, z .

Setzt man f_k in (3) an Stelle von f , so werden sich die Räume $\Gamma_q(a)$ untereinander vertauschen. Durch die Ausdrücke von der Form

$$(22) \quad \frac{1}{2\pi i} \int_a f_k d\zeta$$

wird der gesamte Wertevorrat der Funktion (7) in ihrem Riemannschen Raume ausgeschöpft.

Ein entsprechendes Ergebnis bekommt man für die Funktionen von Form (10) [d. h. im zweiten Falle, wenn $D(P_2) \equiv 0$ ist].

Wir wollen nunmehr noch auf zwei von (7) bzw. (10) verschiedenen Darstellungen für unsere Potentialfunktionen eingehen.

I. Setzt man (im Falle $D(P_2) \not\equiv 0$)

$$(23) \quad \frac{P_1(\zeta_q) \zeta_q^e}{\frac{\partial P_2(\zeta_q)}{\partial \zeta_q}} = V + iW,$$

wo V und W noch undefinierte Konstanten sind, und eliminiert man aus (23) und (8) ζ_q , so erhält man eine rationale Funktion in x, y, z, V und W . Trennt man ferner den reellen und imaginären Teil der erhaltenen Gleichung und eliminiert man einmal V , das andere Mal W , so bekommt man schließlich zwei Gleichungen

$$(23^*) \quad G_1(x, y, z, V) = 0, \quad G_2(x, y, z, W) = 0,$$

die uns die Niveauflächen des reellen bzw. des imaginären Teiles unserer Funktion F_q ergeben, an denen der konstante Wert V bzw. W angenommen wird.

II. Die Darstellung (3) erlaubt uns bequem die ursprünglich durch (1) definierte Funktion außerhalb ihres Zuordnungsgebietes $G_q(e)$ analytisch fortzusetzen.

Durch geeignete Wahl der Kurve α_s kann man F_q in jedem Punkte des x, y, z -Raumes in der Form

$$(24) \quad F_q = \frac{1}{2\pi i} \int_{\alpha_s} \frac{P_1(\zeta) \zeta^e}{P_2(\zeta)} d\zeta$$

darstellen, wobei jetzt im Gegensatz zur Darstellung (22) der Index s bei α_s die Integration längs verschiedener Kurven der komplexen Ebene bedeutet. Durch die Formel (7) bzw. (10) ist die Frage der Singularitäten der von uns betrachteten Funktionen auf die Frage der Singularitäten der algebraischen Funktionen (22) zurückgeführt. Die Lage derselben läßt sich in jedem einzelnen Falle durch (etwas mühselige) rechnerische Arbeit erledigen.

Es ist aber zweckmäßiger, die Singularitätenuntersuchung als speziellen Fall gewisser Pseudo-Abbildungen des zweidimensionalen Raumes auf den dreidimensionalen (wovon früher die Rede war) zu betrachten. *Die Singularitäten unserer Funktionen sind dadurch ausgezeichnet, daß sie aus den Polen durch diese Pseudo-Abbildungen hervorgegangen sind.* Auf diese geometrischen Untersuchungen werden wir in einer anderen Arbeit eingehen.

Es ist bemerkenswert, daß man mit Hilfe der Integrale (24) (in welchen als Integranden lediglich rationale Funktionen auftreten) die ganze Singularitätenumgebung ausschöpfen kann.

Wir können aus unseren Betrachtungen noch etwas Weiteres schließen. Die Singularitäten unserer Funktionen werden offenbar [im Falle $D(P_2) \neq 0$] durch das Gleichungssystem

$$(25) \quad \frac{\partial P_2(\zeta)}{\partial \zeta} = 0, \quad P_2(\zeta) = 0$$

gegeben. Eliminieren wir ζ aus (25) und trennen den reellen und imaginären Teil, so erhalten wir zwei Gleichungen:

$$(26) \quad K_1(x, y, z) = 0, \quad K_2(x, y, z) = 0$$

für die Lage der Singularitäten. Diese Gleichungen können nicht abhängig sein, weil sonst unsere algebraische Potentialfunktion auf einer Fläche singularär wäre, was unmöglich ist. (26) stellt deswegen eine Raumkurve oder einen bzw. mehrere Punkte dar.

Wir haben beim Beweis des ersten Satzes die Voraussetzung gemacht, daß für keinen in Betracht kommenden Wert von x, y, z im Falle I: $D(P_2)$, und im Falle II: $D(P_3, P_4, \dots, P_s)$ verschwindet. Diese Annahme ist aber vollständig berechtigt, da *die Stellen des x, y, z -Raumes, wo $D(P_2)$ [bzw. $D(P_3, P_4, \dots, P_s)$] verschwindet, im allgemeinen mit der singularären Stelle eines Zweiges der Funktion (7) [bzw. (10)] identisch ist,*

was man durch eine einfache Überlegung einsehen kann. Wir beschränken uns im weiteren auf den Fall I, da die Übertragung dieser Betrachtung auf den Fall II keine Schwierigkeiten bietet. Bekanntlich ist im Falle I:

$$(27) \quad \frac{\partial P_2(\zeta_q)}{\partial \zeta_q} \equiv \prod_s (\zeta_q - \zeta_s)$$

$$[s = 1, 2, \dots, (q - 1), (q + 1), \dots, (2n + 2m)].$$

Ist nun ein ζ_s gleich ζ_q , so wird der entsprechende Nenner in (7) Null. Hat der Zähler einen von Null verschiedenen Wert, so ist der Punkt singulär.

Markieren wir uns für eine bestimmte reguläre Stelle x, y, z die komplexe ζ -Ebene (vgl. Fig. 3), so werden den $(2m + 2n)$ -Werten von ζ_q $(2m + 2n)$ verschiedene Punkte der ζ -Ebene entsprechen. Denken wir die ζ -Ebene in dem fünfdimensionalen Raum

$$(28) \quad X_1 = x, X_2 = y, X_3 = z, X_4 = \Re(\zeta), X_5 = \Im(\zeta)$$

gelegen, so werden wir durch jedes Gleichungssystem

$$(29) \quad X_1 = x, X_2 = y, X_3 = z, X_4 = \Re(\zeta_q) \equiv g_1(x, y, z);$$

$$X_5 = \Im(\zeta_q) \equiv g_2(x, y, z)$$

eine dreidimensionale Mannigfaltigkeit erhalten. Die Schnitte dieser $(2m + 2n)$ Mannigfaltigkeiten stellen Kurven oder Punkte im fünfdimensionalen Raume dar. Denkt man sie im dreidimensionalen x, y, z -Raume liegend, so sind sie mit den singulären Stellen unserer Funktion identisch.

§ 3.

Über das vollständige Differential der algebraischen Funktion mit rationalen Zugeordneten.

In der zweidimensionalen Potentialtheorie spielt das vollständige Differential eine hervorragende Rolle. Ist $G(x, y)$ eine Potentialfunktion des zweidimensionalen Raumes und G^* die konjugierte (d. h. bestehen zwischen G und G^* die Cauchy-Riemannschen Differentialgleichungen), so ist

$$(1) \quad G dx + i G^* dy$$

ein vollständiges Differential und

$$(2) \quad \int_{\Gamma} (G dx + i G^* dy),$$

genommen über eine geschlossene Kurve Γ , einen nur von den im Innern von Γ eingeschlossenen Singularitäten abhängigen Wert besitzt.

Der Cauchysche Residuensatz läßt sich aber auch für unsere dreidimensionalen Potentialfunktionen deuten.

Ist F_q ein Zweig einer dreidimensionalen harmonischen Funktion, so lassen sich zwei andere Funktionen F_q^* und F_q^{**} angeben, die mit F_q durch das Gleichungssystem

$$(3) \quad \frac{\partial F_q}{\partial y} = i \frac{\partial F_q^*}{\partial x}, \quad \frac{\partial F_q}{\partial z} = i \frac{\partial F_q^{**}}{\partial z}, \quad \frac{\partial F_q^*}{\partial z} = \frac{\partial F_q^{**}}{\partial y}$$

verbunden sind, und demnach

$$(4) \quad F_q dx + i F_q^* dy + i F_q^{**} dz$$

ein vollständiges Differential ist.

Die Funktion F_q^* ist durch F_q bis auf eine Funktion von y und z allein bestimmt.

Unter diesen Funktionen zeichnen wir eine aus. Ist

$$(5) \quad F_q = \frac{1}{2\pi i} \int_a f d\zeta,$$

so sei

$$(6) \quad F_q^* = \frac{1}{2\pi i} \int_a f \frac{\zeta + \frac{1}{\zeta}}{2} d\zeta, \quad F_q^{**} = \frac{1}{2\pi i} \int_a f \frac{\zeta - \frac{1}{\zeta}}{2} d\zeta.$$

Wir kehren nunmehr zu unseren algebraischen Funktionen mit rationalen Zugeordneten zurück und spezialisieren die Integrationskurve a zu e . In

$$(7) \quad F_q = \frac{1}{2\pi} \int_0^{2\pi} \frac{p_1}{p_2} dt, \quad F_q^* = \frac{1}{2\pi} \int_0^{2\pi} \frac{p_1}{p_2} \cos t dt, \quad F_q^{**} = \frac{1}{2\pi} \int_0^{2\pi} \frac{p_1}{p_2} \sin t dt$$

erhalten wir — wie in § 2 gezeigt wurde — Zweige dreier algebraischen Funktionen F , F^* und F^{**} . Ist uns F_q in der Form (7) bzw. (10) (des § 1) gegeben, so ist es möglich, die Funktionen F_q^* und F_q^{**} in entsprechender Weise aufzuschreiben

Wie früher seien $G_1(e)$, $G_2(e)$, ..., $G_s(e)$ die Zuordnungsräume von F_1 , F_2 , ..., F_s , so daß also in $G_q(e)$ die Beziehung (7) gilt.

Sei \mathfrak{S} eine geschlossene Jordansche Raumkurve mit bestimmtem Durchlaufungssinn. Wir bezeichnen allgemein mit $\mathfrak{S}_q(e)$ denjenigen Teil von \mathfrak{S} [vgl. Fig. 2, S. 634], der in dem Zuordnungsraum $G_q(e)$ liegt [$\mathfrak{S}_q(e)$ kann natürlich auch aus mehreren getrennten Stücken bestehen].

Bilden wir

$$(8) \quad \sum_{v=1}^{v=k} \int_{\mathfrak{S}_v(e)} (F_v dx + i F_v^* dy + i F_v^{**} dz),$$

so ist der Ausdruck (8)

$$(9) \quad \frac{1}{2\pi} \int_{\mathfrak{S}} \int_0^{2\pi} \frac{p_1}{p_2} (dt dx + i \cos t dt dy + i \sin t dt dz)$$

gleich. Wie wir noch nachträglich zeigen werden, können wir die Integrationsfolge in (9) vertauschen. Der Ausdruck (8) ist somit weiter gleich

$$(10) \quad \frac{1}{2\pi} \int_0^{2\pi} \left(\int_{\mathfrak{S}(t)} \frac{p_1}{p_2} du \right) dt,$$

wo $\mathfrak{S}(t)$ diejenige Kurve in der u -Ebene, die \mathfrak{S} bei dem betreffenden t entspricht, ist. (Vgl. Fig. 2 und 6.)

Wir müssen nunmehr feststellen, welche Singularitäten von $\frac{p_1}{p_2}$ für ein bestimmtes t die Kurve $\mathfrak{S}(t)$ in ihrem Innern enthält. Sei wie früher

$$(11) \quad p_2(u, e^{it}) \equiv \sum_{k=0}^{k=n} u^k \left(\sum_{r=-(m+k)}^{r=(m+k)} A_r^{(n-k)} e^{ivt} \right)$$

und seien mit $\alpha_1, \alpha_2, \dots, \alpha_n$ die Wurzel der Gleichung

$$(12) \quad \sum_{k=0}^{k=n} \alpha^k \left(\sum_{r=-(m+k)}^{r=(m+k)} A_r^{(n-k)} e^{ivt} \right) = 0$$

bezeichnet, so sind die α_k Funktionen von t allein (also von x, y, z unabhängig).

Wir bilden nun die Diskriminante von (12)

$$(13) \quad \delta(P_2) \equiv \prod_{k \neq l} (\alpha_k - \alpha_l) \quad (k, l = 1, 2, 3, 4, \dots, n).$$

$\delta(P_2)$ kann [wie früher $D(P_2)$] nicht identisch Null sein oder identisch verschwinden.

Die Gleichungen

$$(14) \quad x = \Re[\alpha_k(t)], \quad y \cos t + z \sin t = \Im[\alpha_k(t)]$$

stellen offenbar bei festem t eine Gerade $T_k(t)$ im x, y, z -Raume dar, die für das betreffende t der Nullstelle α_k von p_2 in der u -Ebene entspricht. Variiert man t in den Grenzen von 0 bis 2π , so beschreibt $T_k(t)$ eine Regelfläche $C_k(e)$ (vgl. Fig. 1, S. 633). Auf diese Weise erhalten wir zu jedem Polynom $p_2(u, e^{it})$ n Regelflächen $C_1(e), C_2(e), \dots, C_n(e)$. Die Kurve \mathfrak{S} kann eine Teilschar der Geraden $T_k(t)$ umfassen¹⁵⁾, und zwar umfasse sie diejenigen Geraden der Regelfläche $C_k(e)$, die den Werten t des Intervalles $\tau_1^{(k)} \leqq t \leqq \tau_2^{(k)}$ entsprechen. Ist p_2 gleich dem Ausdruck (11), so ergibt im Falle $\delta(P_2) \neq 0$ die unmittelbare Rechnung

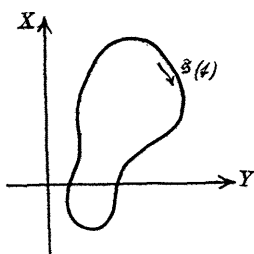


Fig. 6.

Die Kurve $\mathfrak{S}(t)$ entspricht der Kurve \mathfrak{S} der Fig. 2.

¹⁵⁾ Wir sagen: \mathfrak{S} umfaßt eine Geradenschar, wenn man \mathfrak{S} nicht auf einen Punkt zusammenziehen kann, ohne dabei jede Gerade der Schar zu schneiden. (Wie aus dem Beweis leicht zu ersehen ist, unterliegt \mathfrak{S} noch einer Beschränkung: \mathfrak{S} darf mit C_k nur endlichviele Punkte gemeinsam haben.)

$$(15) \quad \sum_q \int_{\mathfrak{C}_q} (F_q dx + i F_q^* dy + i F_q^{**} dz) = i \sum_{\nu=0}^{\nu=n} \int_{\tau_1^{(\nu)}}^{\tau_3^{(\nu)}} \frac{p_1(\alpha_\nu, e^{it})}{\frac{\partial p_2(\alpha_\nu, e^{it})}{\partial \alpha_\nu}} dt, .$$

$$\mathfrak{C} = \sum_q \mathfrak{C}_q.$$

Wenn $\delta(p_2)$ für einen speziellen reellen $\tau_3^{(\nu)}$ -Wert, der in dem Intervall $\tau_1^{(\nu)}$ bis $\tau_3^{(\nu)}$ liegt, verschwindet, so verstehen wir unter der rechten Seite von (15) den Grenzwert

$$(16) \quad \lim_{\varepsilon \rightarrow 0} \sum_{\nu=0}^{\nu=n} \left(\int_{\tau_1^{(\nu)}}^{\tau_3^{(\nu)} - \varepsilon} \dots + \int_{\tau_3^{(\nu)} + \varepsilon}^{\tau_3^{(\nu)}} \dots \right).$$

Im Falle $\delta(p_2) \equiv 0$ erhalten wir eine ähnliche Formel wie (15) mit dem Unterschied, daß auf der rechten Seite die Summation über verschiedene Wurzeln der Gleichung (12) zu nehmen ist und an Stelle der unter dem Integral stehenden Ausdrücke gewisse, ähnlich wie im § 2 gebaute, Determinanten auftreten. Wir gehen auf diese rein formale Rechnung nicht ein.

Wie sofort einzusehen ist, läßt sich unser Satz mit geringfügigen Modifikationen auch für den Fall beweisen, daß α eine beliebige Jordansche Kurve (nicht speziell ϵ) ist. Unser Satz erlaubt außerdem verschiedene Modifikationen und Kombinationen, die insbesondere interessant sind, wenn man außer f noch definierte Funktionen f_k (vgl. S. 633) heranzieht.

In vielen Fällen lassen sich die Riemannschen Räume der algebraischen Funktion so zusammenheften, daß die in der Formel (15) auftretenden Zweige F_q durch analytische Fortsetzung in gleicher Reihenfolge auseinander hervorgehen, wie die Stücke \mathfrak{C}_q aufeinanderfolgen. Im Riemannschen (mehrdeutigen) Raume erscheint dann $\sum \int_{\mathfrak{C}_q} \dots$ als ein Integral, welches über eine stetige nicht notwendig geschlossene Kurve erstreckt ist.

In Arbeit I (S. 657) wurde ein Spezialfall unseres Satzes besprochen, welcher sich auf Punkt-, Linien- und Kreispole bezieht. (Die Kreispole sind natürlich in dem x, y, z -Raume Verzweigungsstellen der Funktion.) Da die Darstellung dort sehr knapp gehalten ist, so sei (um evtl. Mißverständnissen vorzubeugen) etwas ausführlicher auf diesen interessanten Spezialfall eingegangen. Wir können uns auf den Fall beschränken, daß die Funktion nur eine Singularität besitzt, wobei der Pol im Falle des Punktpoles mit dem Nullpunkt zusammenfällt, im Falle des Linien-

poles eine vom Nullpunkte ausgehende Halbgerade ist und im Falle des Kreispoles ein Kreis vom Radius 1 mit Nullpunkt als Mittelpunkt (vgl. dazu S. 653—655 bzw. Fig. 3, 4, 5, 6 der Arbeit I). Es kommt dann nur eine Ebene C_1 vor. In den beiden ersten Fällen wird sie die zweifach überdeckte Ebene $x = 0$. Die erzeugenden Geraden $T_1(t)$ liegen in dieser Ebene und gehen durch den Nullpunkt. Als Parameter t können wir den

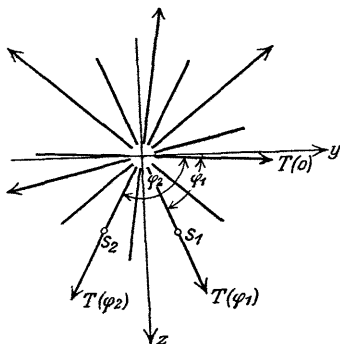


Fig. 7.

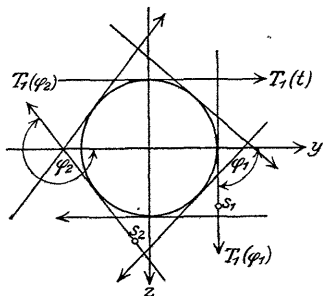


Fig. 8.

Winkel von $T(\varphi)$ mit der y -Achse nehmen¹⁶⁾. Stellen wir uns vor, daß die Kurve \mathfrak{S} in den zwei Punkten, die den Winkel φ_1 und φ_2 entsprechen, C_1 schneidet. Die rechte Seite der Formel (15) lautet dann

$$(17) \quad \int_{\varphi_1}^{\varphi_2} [\text{Residuum von } f] d\varphi + \int_{\varphi_1+\pi}^{\varphi_2+\pi} [\text{Residuum von } f] d\varphi.$$

Im Falle des Kreispoles (mit der Verzweigung: Peripherie des Einheitskreises in der yz -Ebene) wird C_1 das Äußere dieses Kreises. Die Erzeugenden T sind Tangenten dieses Kreises. Als Parameter t kann man den Winkel, den diese Geraden mit der y -Achse bilden, benutzen. Entsprechen zwei Durchstoßpunkten die Winkel φ_1 und φ_2 , so wird die rechte Seite von (15)

$$(18) \quad \int_{\varphi_1}^{\varphi_2} [\text{Residuum von } f] d\varphi \quad ^{17)}$$

gleich sein.

¹⁶⁾ Diese Wahl des Parameters φ stimmt nicht ganz mit der in der Arbeit I vorgenommenen überein. Die Formeln (17) und (18) unterscheiden sich dabei formell etwas von den entsprechenden in der Arbeit I.

¹⁷⁾ Schneidet die Kurve \mathfrak{S} die Fläche C_1 nur in einem Punkte (was z. B. in dem Falle stattfindet, wenn die Kurve \mathfrak{S} die Ebene $x = 0$ das zweite Mal im Innern des Einheitskreises durchstößt), so ist natürlich der Winkel φ_1 dem Winkel φ_2 ungleich.

Über die Realität von Nullstellen ganzer transzendenter Funktionen.

Von

Nikolaj Tschebotareff in Odessa.

I.

Wie bekannt, besteht die notwendige und hinreichende Bedingung für die Realität (bzw. Positivität) von Nullstellen eines reellen Polynoms $f(z) = a_0 + a_1 z + \dots + a_n z^n$ ($a_n \neq 0$) darin, daß die quadratische Form

$$\sum_{i,j}^{0, \dots, n-1} s_{m+i+j} x_i x_j$$

positiv definit sein muß, wobei $-s_k$ die Koeffizienten der Entwicklung der logarithmischen Ableitung von $f(z)$ sind:

$$\frac{f'(z)}{f(z)} = -s_1 - s_2 z - \dots,$$

während m eine beliebige gerade (bzw. ungerade) Zahl ist. (Man nimmt gewöhnlich $m = 0$ und als s_k die Koeffizienten der Entwicklung von $\frac{f'(z)}{f(z)}$ nach fallenden Potenzen von z ; doch liegt dies keineswegs in der Natur der Frage).

Herr Grommer hat in seiner grundlegenden Arbeit¹⁾ diese Bedingung auf Nullstellen von ganzen transzendenten Funktionen erweitert. Er hat nämlich folgendes gezeigt:

A. Für jede reelle ganze transzendente Funktion vom Typus

$$(1) \quad f(z) = e^{g(z)} \prod_{\nu=1}^{\infty} \left(1 - \frac{z}{\alpha_{\nu}}\right)^{A_{\nu}} e^{g_{\nu}(z)},$$

¹⁾ J. Grommer, Über ganze transzendente Funktionen usw., J. f. M. **144** (1914), S. 114.

wo $g(z) = \gamma_0 + \dots + \gamma_m z^m$ ($\gamma_m \geq 0$), $g_\nu(z) = \gamma_{0,\nu} + \dots + \gamma_{m-1,\nu} z^{m-1}$, α_ν reell (bzw. positiv), A_ν ganz positiv, m gerade (bzw. ungerade) ist, sind die quadratischen Formen

$$(2) \quad \sum_{\mu, \nu}^{0, \dots, p-1} s_{m+\mu+\nu} x_\mu x_\nu \quad (p=1, 2, \dots, \text{ad inf.})$$

positiv-definit, wobei $-s_k$ die Koeffizienten der Entwicklung

$$(3) \quad \varphi(z) = \frac{f'(z)}{f(z)} = -s_1 - s_2 z - \dots$$

sind.

B. Sind umgekehrt alle Formen (2) positiv-definit, so kann man behaupten, daß die reelle ganze transzendente Funktion $f(z)$ den Typus (1) hat und dabei α_ν reell bei geradem m und reell positiv bei ungeradem m sind.

Herr Grommer löst in dieser Arbeit auch eine allgemeinere Aufgabe. Er betrachtet reelle meromorphe Funktionen $\varphi(z)$ mit der Bedingung, daß die entsprechenden quadratischen Formen (2) positiv-definit sind. Dann kommt er zum folgenden Ergebnis:

C. Die notwendige und hinreichende Bedingung dafür, daß die Formen (2) positiv-definit sind, besteht darin, daß die entsprechende reelle meromorphe Funktion folgende Mittag-Lefflersche Entwicklung zuläßt:

$$(4) \quad \varphi(z) = - \sum_{\nu=1}^{\infty} \left(\frac{A_\nu}{z - \alpha_\nu} + A_\nu \beta_\nu + A_\nu \beta_\nu^2 z + \dots + A_\nu \beta_\nu^{n-1} z^{m-2} \right) + \alpha_1 + \alpha_2 z + \dots + \alpha_m z^{m-1},$$

wo α_ν reell, $\beta_\nu = \frac{1}{\alpha_\nu}$, $\alpha_m \geq 0$, $\text{sign } A_\nu = \text{sign } \alpha_\nu^m$. Ist dabei von vornherein bekannt, daß A_ν positiv sind (wie dies z. B. bei B. der Fall ist), so kann man bei ungeradem m behaupten, daß die α_ν auch alle positiv sind.

Herr Grommer gelangt zu diesen Ergebnissen mit Hilfe der Kettenbruchentwicklung der Funktion (3). Dabei benutzt er weitgehende mengentheoretische Hilfsmittel.

Herr Kritikos²⁾ kommt zu den Ergebnissen A. und B. auf viel elementarerem Wege, indem er nur Kettenbrüche und den Hadamardschen Satz über ganze transzendente Funktionen anwendet.

Im ersten Abschnitt der vorliegenden Abhandlung^{2a)} zeige ich den Zusammenhang der *Signatur* der quadratischen Form (2) mit der Anzahl der

²⁾ N. Kritikos, Über ganze transzendente Funktionen usw., *Math. Annalen* 81, S. 97.

^{2a)} Vgl. auch meine frühere Arbeit: Über die Realität von Wurzeln ganzer transzendenter Gleichungen (russisch), *Journ. des Forsch.-Inst. Odessa* 1, Nr. 1 (1923).

komplexen (bzw. negativen) Nullstellen von $f(z)$, die unter den p absolut kleinsten Nullstellen vorkommen. Nimmt man dann für jeden gewählten Index p einen genügend großen Index m , so gibt der genannte Zusammenhang das Kriterium für die Realität (bzw. Positivität) von Nullstellen, welches für *alle* reellen ganzen transzendenten Funktionen (d. h. auch unendlicher Ordnung) gültig ist. An der Spitze dieser Untersuchung steht die asymptotische Formel

$$(5) \quad s_m = \frac{A_1}{\alpha_1^m} + \frac{A_2}{\alpha_2^m} + \dots + \frac{A_k}{\alpha_k^m} + \frac{\varepsilon_{m,k}}{\alpha_k^m},$$

wo $\alpha_1, \alpha_2, \dots, \alpha_k$ Nullstellen von $f(z)$ derart sind, daß die Ungleichungen

$$|\alpha_1| \leq |\alpha_2| \leq \dots \leq |\alpha_k| < |\alpha_l| \quad (l > k)$$

gelten und $\varepsilon_{m,k}$ für genügend großes m beliebig klein wird. Diese ganz elementare Methode schließt sich eng an die Bernoulli-Graeffesche Methode zur Berechnung der Wurzeln an, die von Herrn Pólya³⁾ auf transzendente Funktionen erweitert wurde. Diese Methode erlaubt, die Hurwitzsche Aufgabe auch für den Fall ganzer transzendenter Funktionen zu lösen, wie ich dies mit einigen Beschränkungen durchführe (Herr Grommer hat dies für den Fall ganzer transzendenter Funktionen von der Höhe Null getan).

Im zweiten Abschnitt beweise ich den verallgemeinerten Satz C., indem ich statt meromorpher alle reellen eindeutigen analytischen Funktionen $\varphi(z)$ betrachte, die durch die Mittag-Lefflersche verallgemeinerte Partialbruchentwicklung darstellbar sind. Hier benutze ich ein neues Kriterium für die Realität von Nullstellen, welches auch an und für sich bemerkenswert ist:

Damit die Nullstellen von $f(z)$ alle reell seien, ist notwendig und hinreichend, daß der Ausdruck

$$J\left(\frac{f'(z)}{f(z)}\right) : J(z),$$

wo das Symbol $J(\dots)$ den imaginären Teil bezeichnet, sein Vorzeichen auf der ganzen Ebene behält.

Dieses Kriterium gilt für reelle Polynome und ganze transzendente Funktionen von der Höhe Null und Eins. Es ist mit der Positivität von Formen (2) für $m=2$ äquivalent. Um dies zu beweisen, bediene ich mich eines modifizierten Borel-Lindelöfschen Summationsverfahrens^{4,5)}.

³⁾ G. Pólya, Über das Graeffesche Verfahren, Zeitschr. f. M. u. Ph. 63, S. 275.

⁴⁾ É. Borel, Leçons sur les séries divergentes. Paris 1901, p. 164–168.

⁵⁾ E. Lindelöf, Sur l'application de la théorie des résidus etc., J. d. M. (5) 9 (1903), p. 213.

Diese Zusammenstellung beider Kriterien ist mit dem Carathéodory-Toeplitz'schen Ideenkreis⁶⁾ verwandt.

Ich spreche meinen Dank den Herren Krein, Ostrowski, Krawtchouk und Grandjot aus, die mir auf verschiedene Weise behilflich waren.

§ 1.

Es sei $f(z)$ eine beliebige *reelle ganze transzendente Funktion*

$$(1) \quad f(z) = a_0 + a_1 z + a_2 z^2 + \dots \quad (a_0 \neq 0)$$

(d. h. mit reellen Koeffizienten ihrer Maclaurinschen Entwicklung). (Ist $f(z)$ nicht reell, so brauchen wir nur an Stelle von $f(z)$ die Funktion $f(z) \cdot \bar{f}(z)$ zu betrachten, wobei $\bar{f}(z)$ die mit $f(z)$ konjugiert imaginäre Funktion ist.) Bezeichnen wir mit $-s_k$ die Koeffizienten der Maclaurinschen Entwicklung ihrer logarithmischen Ableitung:

$$(2) \quad \frac{f'(z)}{f(z)} = -s_1 - s_2 z - s_3 z^2 - \dots,$$

so kann man die s_k mit Hilfe der Newtonschen rekurrenten Formeln erhalten:

$$(3) \quad \begin{aligned} a_1 s_1 + a_1 &= 0, \\ a_0 s_2 + a_1 s_1 + 2a_2 &= 0, \\ &\dots \end{aligned}$$

Es seien ferner $\alpha_1, \alpha_2, \dots$ alle verschiedenen Nullstellen der Funktion $f(z)$ von der Multiplizität bzw. A_1, A_2, \dots nach der Größe ihrer absoluten Beträge geordnet:

$$(4) \quad |\alpha_1| \leq |\alpha_2| \leq \dots$$

Ist k eine Nummer, für welche $|\alpha_k| < |\alpha_{k+1}|$ gilt (solche Nummern wollen wir *Hauptnummern* nennen), so gilt folgende asymptotische Formel:

$$(5) \quad s_m = \frac{A_1}{\alpha_1^m} + \frac{A_2}{\alpha_2^m} + \dots + \frac{A_k}{\alpha_k^m} + \frac{\varepsilon_{m,k}}{\alpha_k^m},$$

wobei $\varepsilon_{m,k}$ für genügend große Werte von m beliebig klein wird. Denn $\frac{f'(z)}{f(z)}$ hat α_ν ($\nu = 1, 2, \dots$) als einfache Pole mit Residuen bzw. A_ν ($\nu = 1, 2, \dots$) und bleibt in anderen Stellen regulär. Daher ist der Konvergenzradius der Funktion

$$\begin{aligned} \frac{f'(z)}{f(z)} &= \frac{A_1}{z - \alpha_1} + \frac{A_2}{z - \alpha_2} + \dots + \frac{A_k}{z - \alpha_k} \\ &= \sum_{m=1}^{\infty} \left(-s_m + \frac{A_1}{\alpha_1^m} + \frac{A_2}{\alpha_2^m} + \dots + \frac{A_k}{\alpha_k^m} \right) z^{m-1} \end{aligned}$$

⁶⁾ G. Herglotz, Über Potenzreihen mit positivem reellen Teil im Einheitskreis. Ber. Lpz. 63 (1911), S. 501. Dort ist auch die vollständige Literatur angeführt.

gleich $|\alpha_{k+1}|$, d. h. größer als $|\alpha_k|$. Bezeichnen wir ihren Entwicklungskoeffizienten bei z^{m-1} mit $\frac{\varepsilon_{m,k}}{\alpha_k^m}$, so konvergiert die Reihe $\sum_{m=1}^{\infty} \frac{\varepsilon_{m,k}}{\alpha_k^m} \alpha_k^m$, woraus folgt $\lim_{m \rightarrow \infty} \varepsilon_{m,k} = 0$, w. z. b. w.

Wir setzen in der Formel (5) $\alpha_i = \frac{1}{\beta_i}$:

$$(5') \quad s_m = A_1 \beta_1^m + A_2 \beta_2^m + \dots + A_k \beta_k^m + \beta_k^m \cdot \varepsilon_{m,k}$$

und betrachten die quadratische Form

$$(6) \quad F(x_0, x_1, \dots, x_{p-1}) = \sum_{\mu, \nu}^{0, \dots, p-1} s_{m+\mu+\nu} x_{\mu} x_{\nu}$$

$$= \sum_{k=1}^p A_k \beta_k^m (x_0 + \beta_k x_1 + \dots + \beta_k^{p-1} x_{p-1})^2 + \beta_p^{m \cdot 0, \dots, p-1} \sum_{\mu, \nu} \varepsilon_{m+\mu+\nu, p} x_{\mu} x_{\nu},$$

wo p eine Hauptnummer ist.

Kommen in der Reihe $\beta_1, \beta_2, \dots, \beta_p$ Paare konjugiert imaginärer Größen vor, so verfahren wir folgendermaßen. Es sei $\beta = \rho(\cos \varphi + i \sin \varphi)$, $\bar{\beta} = \rho(\cos \varphi - i \sin \varphi)$ ein solches Paar. Dann nimmt das entsprechende Paar Glieder in der ersten Summe von (6) folgende Gestalt an:

$$A \beta^m (x_0 + \beta x_1 + \dots + \beta^{p-1} x_{p-1})^2 + A \bar{\beta}^m (x_0 + \bar{\beta} x_1 + \dots + \bar{\beta}^{p-1} x_{p-1})^2$$

$$= A \rho^m (\cos m \varphi + i \sin m \varphi) (u + i v)^2 + A \rho^m (\cos m \varphi - i \sin m \varphi) (u - i v)^2$$

$$= 2 A \rho^m \cos m \varphi (u^2 - v^2) - 4 A \rho^m \sin m \varphi u v,$$

wobei die neuen Bezeichnungen folgende Bedeutung haben:

$$y = x_0 + \beta x_1 + \dots + \beta^{p-1} x_{p-1}, \quad \bar{y} = x_0 + \bar{\beta} x_1 + \dots + \bar{\beta}^{p-1} x_{p-1},$$

$$u = \frac{y + \bar{y}}{2}, \quad v = \frac{y - \bar{y}}{2i}.$$

Nun wollen wir zwei Fälle unterscheiden:

$$\text{I. } |\cos m \varphi| \geq \frac{1}{\sqrt{2}}, \quad |\operatorname{tg} m \varphi| \leq 1, \quad |\sec m \varphi| \leq \sqrt{2};$$

$$\text{II. } |\sin m \varphi| > \frac{1}{\sqrt{2}}, \quad |\operatorname{cotg} m \varphi| < 1, \quad |\operatorname{cosec} m \varphi| < \sqrt{2}.$$

Man sieht leicht, daß immer einer dieser Fälle eintreten muß.

Im ersten Falle führen wir folgende Transformation aus:

$$2 A \rho^m \cdot \cos m \varphi \cdot (u^2 - v^2) - 4 A \rho^m \cdot \sin m \varphi u v$$

$$= 2 A \rho^m \cdot \cos m \varphi \cdot (u - \operatorname{tg} m \varphi \cdot v)^2 - 2 A \rho^m \cdot \sec m \varphi \cdot v^2.$$

Im zweiten Falle setzen wir: $u = w + t$, $v = w - t$; $uv = w^2 - t^2$,
 $u^2 - v^2 = 4 w t$:

$$\begin{aligned}
& 2A \varrho^m \cos m\varphi (u^2 - v^2) - 4A \varrho^m \sin m\varphi \cdot u \cdot v \\
&= -4A \varrho^m \sin m\varphi (w^2 - t^2) + 8A \varrho^m \cos m\varphi \cdot w \cdot t \\
&= -4A \varrho^m \sin m\varphi (w - \cotg m\varphi \cdot t)^2 + 4A \varrho^m \operatorname{cosec} m\varphi \cdot t^2.
\end{aligned}$$

Wir erhalten also in beiden Fällen zwei unabhängige Quadrate mit entgegengesetzten Vorzeichen. Division ihrer Koeffizienten durch ϱ^m ergibt Größen, deren absoluter Betrag zwischen zwei von Null verschiedenen und von m unabhängigen Grenzen liegt.

Nun führen wir in der Formel (6) die Substitution

$$y_k = x_0 + \beta_k x_1 + \dots + \beta_k^{p-1} x_{p-1} \quad (k = 1, 2, \dots, p)$$

aus, deren Koeffizienten von m unabhängig sind und deren Determinante von Null verschieden ist. Ferner gehen wir im Falle imaginärer Größen β entweder zu u, v oder zu w, t über. Die Koeffizienten dieser Substitution sind beschränkt und haben von m unabhängige Determinante. Dann nimmt die Form folgende Gestalt an:

$$\begin{aligned}
(7) \quad F = & M_1 |\beta_1|^m \cdot z_1^2 \pm M_2 |\beta_2|^m z_2^m \pm \dots \\
& \pm M_p |\beta_p|^m z_p^2 + |\beta_p|^m \sum_{\mu, \nu}^{0, \dots, p-1} \eta_{\mu+\nu}^{(m)} \cdot z_\mu z_\nu.
\end{aligned}$$

Hier liegen M_k ($k = 1, 2, \dots, p$) zwischen zwei positiven von m unabhängigen Grenzen; $\eta_{\mu+\nu}^{(m)}$ werden für genügend großes m beliebig klein. Wenn wir daher auf die Form die Gaußsche Methode der Zerlegung in Quadrate anwenden, so müssen die sich dadurch ergebenden Koeffizienten bei Quadraten keine Vorzeichenänderung erleiden, wenn nur m genügend groß gewählt ist. Die Anzahl negativer Vorzeichen in der ersten Summe von (7) ist aber gleich d bei geradem m und gleich $d + d'$ bei ungeradem m , wo d die Anzahl Paare konjugiert imaginärer und $d + d'$ die Anzahl negativer Größen in der Reihe $\beta_1, \beta_2, \dots, \beta_p$ bedeuten. Demnach hat die ganze Form (7) dieselbe Zahl negativer Glieder.

Führen wir die Bezeichnung

$$(8) \quad D_m^p = \begin{vmatrix} s_m & & \dots & s_{m+p-1} \\ \dots & \dots & \dots & \dots \\ s_{m+p-1} & \dots & s_{m+2p-2} \end{vmatrix}$$

ein, so wird die Kette von Hauptdiagonalminoren der Form (6) folgendermaßen dargestellt:

$$1, D_m^1, D_m^2, \dots, D_m^p.$$

Die Zusammenstellung mit dem bekannten Satz der Formentheorie ergibt den

Satz. Sind unter den ersten p Nullstellen einer reellen ganzen transzendenten Funktion $f(z)$ d Paare konjugiert imaginär und d' negativ, und ist dabei $|\alpha_p| < |\alpha_{p+1}|$, so bildet die Kette

$$(9) \quad 1, D_m^1, D_m^2, \dots, D_m^p$$

bei genügend großem m d Vorzeichenwechsel, wenn m gerade ist, und $d + d'$, wenn m ungerade ist⁷⁾.

§ 2.

Bisher setzten wir voraus, daß m „genügend groß“ ist. Ist aber die Höhe der ganzen transzendenten Funktion $f(z)$ endlich, so kann man beweisen, daß die Positivität (bzw. Realität) von allen Nullstellen der Funktion $f(z)$ die Positivität von allen Varianten D_m^p für jeden festgewählten (bzw. festgewählten geraden) Wert von m nach sich zieht, der nicht kleiner als die Höhe von $f(z)$ ist. Dies folgt aus der Tatsache, daß in diesem Falle die Reihen $\sum_{k=1}^{\infty} A_k \beta_k^m$ konvergieren und ihre Summen genau gleich s_m sind:

$$(10) \quad s_m = \sum_{k=1}^{\infty} A_k \beta_k^m \quad \text{für } m \geq m_0.$$

Wenn wir demnach den Variablen x_0, x_1, \dots, x_{p-1} ganz beliebige Zahlenwerte zuschreiben, so kann die Form (6) in folgender Gestalt dargestellt werden:

$$F(x_0, x_1, \dots, x_{p-1}) = \sum_{k=1}^{\infty} A_k \beta_k^m (x_0 + \beta_k x_1 + \dots + \beta_k^{p-1} x_{p-1})^2 \quad \text{für } m \geq m_0.$$

Diese Reihe muß entweder bei lauter positiven β_k oder bei lauter reellen β_k und geradem m konvergieren und hat dann immer einen positiven Wert. Also ist die Form (6) positiv-definit, w. z. b. w.

Man kann umgekehrt beweisen, daß die Positivität aller Varianten D_m^p bei einem festgewählten (bzw. festgewählten geraden) Wert von m_0 die Positivität (bzw. Realität) aller Nullstellen von $f(z)$ nach sich zieht. Um dies zu beweisen, nehmen wir das Gegenteil an, d. h. daß bei dieser Voraussetzung nicht alle Nullstellen positiv (bzw. reell) sind. Dann gibt es solche Werte m_1 und p , für welche alle (bzw. alle geraden) Werte $m \geq m_1$ die Ungleichung $D_{m_1}^p < 0$ ergeben. (Wären etwa $D_{m_1}^p = D_{m_1}^{p+1} = \dots = 0$, so mußte das erste nicht verschwindende Glied dieser Kette negativ sein;

⁷⁾ Herr Krawtchouk hat mich aufmerksam gemacht, daß in manchen Fällen in der Kette (9) einige Vorzeichenwechsel verborgen sein können. Dies tritt nämlich ein, wenn im Innern dieser Kette mehrere Glieder verschwinden. Dann kann man in der Matrix $\|s_{\mu+\nu}\|$ ein anderes System von Hauptdiagonalminoren hervorheben, was stets zu dem gewünschten Resultat führt (siehe Frobenius, J. f. M. 114).

wären dagegen diese Varianten sämtlich gleich Null, so wäre $f(z)$ Polynom mit lauter reellen Nullstellen.) Wir können dabei stets einen solchen Wert von m wählen, daß die Differenz $m - m_0$ gerade wird: $m - m_0 = 2k$. Nun betrachten wir die Matrix

$$(11) \begin{matrix} \delta_{m_0}, & \delta_{m_0+1}, & \dots, & \delta_{m_0+k+p-1} \\ \delta_{m_0+1}, & \delta_{m_0+2}, & \dots, & \delta_{m_0+k+p} \\ \dots & \dots & \dots & \dots \\ \delta_{m_0+k+p-1}, & \delta_{m_0+k+p}, & \dots, & \delta_{m+2p-2} \end{matrix} \quad (m_0+2k+2p-2 = m+2p-2)$$

und die ihr entsprechende quadratische Form

$$(12) \quad \sum_{\mu, \nu}^{0, \dots, p-1} \delta_{m_0+\mu+\nu} x_\mu x_\nu.$$

Diese letztere kann nicht positiv-definit sein, da die Kette ihrer Hauptdiagonalminoren (von rechts unten nach links oben):

$$(13) \quad 1, D_{m+2p-2}^1, D_{m+2p-4}^2, \dots, D_m^p, \dots, D_{m_0}^{p+k}$$

wenigstens einen Vorzeichenwechsel hat. Daraus folgt, daß auch die Kette von Hauptdiagonalminoren (von links oben nach rechts unten):

$$(14) \quad 1, D_{m_0}^1, D_{m_0}^2, \dots, D_{m_0}^{p+k}$$

Vorzeichenwechsel haben muß, d. h. wenigstens eine der Varianten (14) negativ sein muß. Dies widerspricht aber unserer Voraussetzung, w. z. b. w.

§ 3.

Die dargelegte Methode erlaubt uns, die Hurwitzsche Aufgabe für den Fall einer reellen ganzen transzendenten Funktion zu lösen, d. h. die Bedingungen aufzustellen dafür, daß ihre Nullstellen positive (oder negative) Realteile haben.

Wir betrachten hier nur den Fall, wenn verschiedene und nicht konjugiert imaginäre Nullstellen von $f(z)$ auch verschiedene absolute Beträge haben. Außerdem soll $f(z)$ keine rein imaginären Nullstellen haben.

Wir ziehen folgende Varianten in Betracht:

$$(15) \quad \Delta_m^p = \begin{vmatrix} \delta_m, & \delta_{m+1}, & \dots, & \delta_{m+p-1} \\ \delta_{m+2}, & \delta_{m+3}, & \dots, & \delta_{m+p+1} \\ \dots & \dots & \dots & \dots \\ \delta_{m+2p-2}, & \delta_{m+2p-1}, & \dots, & \delta_{m+3p-3} \end{vmatrix}.$$

Wenn wir hier für $\delta_{m+\mu+\nu}$ ihre asymptotischen Ausdrücke (5') einsetzen, sehen wir leicht, daß die Varianten $\Delta_m^p: D_m^p$ im Falle einer Hauptnummer p folgende Gestalt annehmen:

$$\begin{aligned}
& \sum_{k=1}^p A_k \beta_k^m + \beta_p^m \cdot \varepsilon_{m,p}, \dots, \sum_{k=1}^p A_k \beta_k^{m+p-1} + \beta_p^{m+p-1} \cdot \varepsilon_{m+p-1,p} \\
& \dots \\
& \sum_{k=1}^p A_k \beta_k^{m+2p-2} + \beta_p^{m+2p-2} \cdot \varepsilon_{m+2p-2,p}, \dots, \sum_{k=1}^p A_k \beta_k^{m+3p-3} + \beta_p^{m+3p-3} \cdot \varepsilon_{m+3p-3,p} \\
& \dots \\
& \sum_{k=1}^p A_k \beta_k^m + \beta_p^m \cdot \varepsilon_{m,p}, \dots, \sum_{k=1}^p A_k \beta_k^{m+p-1} + \beta_p^{m+p-1} \cdot \varepsilon_{m+p-1,p} \\
& \dots \\
& \sum_{k=1}^p A_k \beta_k^{m+p-1} + \beta_p^{m+p-1} \cdot \varepsilon_{m+p-1,p}, \dots, \sum_{k=1}^p A_k \beta_k^{m+2p-2} + \beta_p^{m+2p-2} \cdot \varepsilon_{m+2p-2,p} \\
& \dots \\
& = \frac{A_1 \cdot A_2 \cdot \dots \cdot A_p \cdot \beta_1^m \cdot \beta_2^m \cdot \dots \cdot \beta_p^m \begin{vmatrix} 1, \beta_1^2, \dots, \beta_1^{2p-2} \\ \dots \\ 1, \beta_p^2, \dots, \beta_p^{2p-2} \end{vmatrix} \begin{vmatrix} 1, \beta_1, \dots, \beta_1^{p-1} \\ \dots \\ 1, \beta_p, \dots, \beta_p^{p-1} \end{vmatrix}}{A_1 \cdot A_2 \cdot \dots \cdot A_p \cdot \beta_1^m \cdot \beta_2^m \cdot \dots \cdot \beta_p^m \begin{vmatrix} 1, \beta_1, \dots, \beta_1^{p-1} \\ \dots \\ 1, \beta_p, \dots, \beta_p^{p-1} \end{vmatrix}^2} + \eta_{m,p} \\
& = \frac{\prod_{\lambda < \mu \leq p} (\beta_\lambda^2 - \beta_\mu^2)}{\prod_{\lambda < \mu \leq p} (\beta_\lambda - \beta_\mu)} + \eta_{m,p} = \prod_{\lambda < \mu \leq p} (\beta_\lambda + \beta_\mu) + \eta_{m,p},
\end{aligned}$$

wobei $\eta_{m,p}$ mit wachsendem m beliebig klein wird.

Wir beachten, daß das Vorzeichen des Realteiles von β_λ ($\lambda = 1, 2, \dots$) mit dem Vorzeichen des Realteiles von $\alpha_\lambda = \frac{1}{\beta_\lambda}$ zusammenfällt. Das Vorzeichen aber von $\beta_\lambda + \bar{\beta}_\lambda$, wo $\bar{\beta}_\lambda$ mit β_λ konjugiert imaginär ist, stimmt mit dem Vorzeichen des Realteiles von β_λ überein.

Nun seien β_p und $\beta_{p-1} = \bar{\beta}_p$ konjugiert imaginäre Größen der Reihe β_1, β_2, \dots . Dann ist das Hauptglied der Variante $\frac{A_m^p}{D_m^p} : \frac{A_m^{p-2}}{D_m^{p-2}}$ gleich

$$\frac{\prod_{\lambda < \mu \leq p} (\beta_\lambda + \beta_\mu)}{\prod_{\lambda < \mu \leq p-2} (\beta_\lambda + \beta_\mu)} = (\beta_p + \beta_{p-1}) \cdot \prod_{\lambda \leq p-2} (\beta_\lambda + \beta_p) \cdot \prod_{\lambda \leq p-2} (\beta_\lambda + \beta_{p-1}).$$

Der Faktor $\prod_{\lambda \leq p-2} (\beta_\lambda + \beta_p) \cdot \prod_{\lambda \leq p-2} (\beta_\lambda + \beta_{p-1})$ ist, als Produkt konjugiert imaginärer Größen, positiv, woraus folgt

$$(16) \quad \text{sign} \frac{A_m^p}{D_m^p} = \text{sign} \Re(\beta_p) \cdot \text{sign} \frac{A_m^{p-2}}{D_m^{p-2}} \text{ bei genügend großem } m,$$

wo das Symbol $\Re(\dots)$ den Realteil der eingeklammerten Größe bedeutet. Dieses Ergebnis kann auch so aufgefaßt werden:

Satz. Kommen in der Reihe $\beta_1, \beta_2, \dots, \beta_p$ unter den oben gemachten Voraussetzungen die Paare mit negativen Realteilen vor, so erleidet die Kette der den Hauptnummern aus der Reihe 1, 2, ..., p entsprechenden Varianten

$$1, \frac{\Delta_m^1}{D_m^1}, \frac{\Delta_m^2}{D_m^2}, \dots, \frac{\Delta_m^p}{D_m^p}$$

bei genügend großem m den Vorzeichenwechsel.

Ganz ebenso kann man beweisen, daß der Hauptteil der Varianten $\frac{\Delta_m^{(k)}}{D_m^{(k)}} : \frac{\Delta_m^{(k-2)}}{D_m^{(k-2)}}$ dasselbe Vorzeichen hat wie der Ausdruck

$$\frac{\beta_p^k - \bar{\beta}_p^k}{\beta_p - \bar{\beta}_p},$$

wo wir die Bezeichnung

$$(17) \quad \Delta_m^{(k)} = \begin{vmatrix} s_m & s_{m+1} & \dots & s_{m+p-1} \\ s_{m+k} & s_{m+k+1} & \dots & s_{m+p+k-1} \\ \dots & \dots & \dots & \dots \\ s_{m+k(p-1)} & s_{m+k(p-1)+1} & \dots & s_{m+k(p-1)+p-1} \end{vmatrix}$$

eingeführt haben.

Führen wir noch die Bezeichnung $\beta_p = \rho (\cos \varphi + i \sin \varphi)$ ein, so nimmt dieser Ausdruck folgende Gestalt an:

$$\rho^{k-1} \sin k\varphi : \sin \varphi.$$

Ist diese Größe positiv, so folgt der Winkel φ in einem der folgenden Intervalle: $\pm \left(\frac{\pi}{k}, \frac{2\pi}{k}\right), \pm \left(\frac{3\pi}{k}, \frac{4\pi}{k}\right), \pm \left(\frac{5\pi}{k}, \frac{6\pi}{k}\right), \dots$

II.

In dem folgenden Abschnitt beschäftige ich mich mit folgendem Problem:

Welche Typen umfassen reelle analytische Funktionen

$$(1) \quad \varphi(z) = s_1 + s_2 z + s_3 z^2 + \dots,$$

welche die Eigenschaft besitzen, daß alle quadratischen Formen

$$(2) \quad F(x_0, x_1, \dots, x_{p-1}) = \sum_{\mu, \nu}^{0, \dots, p-1} s_{m+\mu+\nu} x_\mu x_\nu$$

für einen festgewählten Wert von m und jeden Wert von p positiv-definit sind?

Dieses Problem hat Herr Grommer⁸⁾ allgemein gelöst. Insbesondere

⁸⁾ J. Grommer, loc. cit.

hat er gezeigt, daß, wenn $\varphi(z)$ eindeutig und meromorph ist, sie nur einfache reelle Pole α_ν besitzt, deren Vorzeichen mit den Vorzeichen der entsprechenden Residuen $(-A_\nu)$ durch die Relation

$$\text{sign } \alpha_\nu^m = \text{sign } A_\nu$$

verbunden sind; außerdem muß $\varphi(z)$ folgende Mittag-Lefflersche Darstellung zulassen:

$$(3) \quad \varphi(z) = - \sum_{\nu=1}^{\infty} \left(\frac{A_\nu}{z - \alpha_\nu} + A_\nu \beta_\nu + A_\nu \beta_\nu^2 z + \dots + A_\nu \beta_\nu^{m-1} z^{m-2} \right) \\ + a_1 + a_2 z + \dots + a_m z^{m-1},$$

wobei $a_m \geq 0$ ist.

Hier erweitere ich dieses Resultat dadurch, daß ich alle eindeutigen und die Mittag-Lefflersche Darstellung zulassenden analytischen Funktionen betrachte. Es ergibt sich, daß man auf diese Weise keine weiteren Funktionen mit lauter positiv-definiten Formen (2) erhält.

Betrachtet man aber alle im Mittag-Lefflerschen Sterne definierten Funktionen, so kommt man zu neuen Funktionen vom erwähnten Typus. Herr Grommer (loc. cit.) hat gezeigt, daß die allgemeinste Funktion von diesem Typus durch das Integral

$$A \cdot \int_{-\infty}^{+\infty} \frac{d\Phi(u)}{\frac{1}{z} - u} \quad (A > 0)$$

darstellbar ist, wo u reell, $\Phi(u)$ eine reelle monoton wachsende Funktion ≤ 1 und das Integral im Stieltjesschen Sinne zu verstehen ist. Hier ist $m = 2$ genommen. Daß in dieser Klasse nicht nur eindeutige Funktionen vorkommen, zeigt das einfache Beispiel

$$\int_{-1}^{+1} \frac{dx}{\frac{1}{z} - x} = \lg \frac{1+z}{1-z}.$$

Um zu meinem Ergebnis zu kommen, verfähre ich folgendermaßen. Ich betrachte alle möglichen Typen von reellen eindeutigen Funktionen, die nicht vom Typus (3) sind, und wähle jedesmal solche Werte von p und x_ν ($\nu = 0, 1, \dots, p-1$), daß die entsprechende Form (2) einen negativen Wert annimmt.

Dabei benutze ich in den meisten Fällen ein Hilfskriterium, das in folgendem besteht.

Sind alle Formen (2) positiv-definit, so nimmt der Ausdruck

$$(4) \quad \frac{J(F_m(z))}{J(z)}$$

im ganzen Stern, in welchem die Funktion $F_m(z) = s_{m-1} + s_m z + \dots$ definiert ist, nur nicht-negative Werte an. Hier bedeutet das Symbol $J(\dots)$ den komplexen Teil der eingeklammerten Funktion.

Dieser Satz ist auch umkehrbar, wenn wir nur diejenigen Funktionen $F_m(z)$ betrachten, deren imaginärer Teil, falls er positiv wird, größer als $\frac{y}{x^2 + y^2}$ auf der oberen Halbebene werden kann.

Dieser Satz ist dem Carathéodory-Toeplitzschenschen⁹⁾ Satz analog, der das Positivsein des Realteiles einer analytischen Funktion im Einheitskreise mit dem Positivsein einer gewissen Hermiteschen Form in Zusammenhang bringt. Ein wesentlicher Unterschied liegt darin, daß dort nur von den Werten im Innern des Konvergenzkreises die Rede ist.

§ 1.

Wir beweisen den schon in der Einleitung besprochenen Satz 1. *Es sei eine unendliche Folge reeller Zahlen*

$$s_m, s_{m+1}, s_{m+2}, \dots \quad \left(\lim_{n \rightarrow \infty} \sqrt[n]{|s_n|} \text{ nicht} = \infty \right)$$

gegeben. Sind alle quadratischen Formen

$$(5) \quad F(x_0, x_1, \dots, x_{\mu-1}) = \sum_{\mu, \nu}^{0, \dots, \mu-1} s_{m+\mu+\nu} x_\mu x_\nu$$

positiv-definit, so nimmt der Ausdruck

$$(6) \quad \frac{J(\varphi_m(z))}{J(z)},$$

wo $\varphi_m(z)$ die im Mittag-Lefflerschen Stern durch das Element

$$(7) \quad \varphi_m(z) = s_m z + s_{m+1} z^2 + \dots$$

bestimmte analytische Funktion ist, nur nicht negative Werte an.

Beweis. Wir nehmen einen beliebigen Punkt $z = \rho e^{i\varphi}$ im Innern des Sternes und bilden den Bereich A , welcher folgende Eigenschaften besitzt:

1. A liegt ganz innerhalb des Sternes;
2. A umfaßt den Nullpunkt und den Punkt z ;
3. A ist von einer geschlossenen Jordanschen Kurve K begrenzt;
4. jeder aus dem Nullpunkt hervorgehende Strahl schneidet K nur einmal;
5. A ist symmetrisch in bezug auf die reelle Achse.

⁹⁾ G. Herglotz, loc. cit.

Wir nehmen für x_μ folgende zwei Wertssysteme:

$$(8) \quad \text{I) } x_\mu = \frac{\varrho^\mu \cos \mu \varphi}{\mu^{\mu \sigma}} \quad (\mu = 0, 1, 2, \dots),$$

$$(9) \quad \text{II) } x_\mu = \frac{\varrho^\mu \sin \mu \varphi}{\mu^{\mu \sigma}} \quad (\mu = 0, 1, 2, \dots),$$

wo σ eine beliebige reelle positive Größe ist, und stellen die Koeffizienten s_{m+n} in folgender Form dar:

$$(10) \quad s_{m+n} = \frac{1}{2\pi i} \oint \frac{\varphi_m(x)}{x^{2+n}} dx \quad (n = 0, 1, 2, \dots),$$

wobei das Integral längs K genommen ist. Sind alle Formen (5) positiv-definit, so gilt insbesondere:

$$(11) \quad F \left(1, \frac{\varrho \cos \varphi}{1^\sigma}, \frac{\varrho^2 \cos 2\varphi}{2^{2\sigma}}, \dots, \frac{\varrho^{p-1} \cos (p-1)\varphi}{(p-1)^{(p-1)\sigma}} \right) \\ = \frac{1}{2\pi i} \oint \frac{\varphi_m(x)}{x^2} \sum_{\mu, \nu}^{0, \dots, p-1} \frac{\varrho^\mu \cos \mu \varphi}{x^\mu \mu^{\mu \sigma}} \cdot \frac{\varrho^\nu \cos \nu \varphi}{x^\nu \nu^{\nu \sigma}} \cdot dx \\ = \frac{1}{2\pi i} \oint \frac{\varphi_m(x)}{x^2} \left(\sum_{\mu=0}^{p-1} \frac{\varrho^\mu \cos \mu \varphi}{x^\mu \mu^{\mu \sigma}} \right)^2 dx = \frac{1}{2\pi i} \oint \frac{\varphi_m(x)}{x^2} \cdot X^2 \cdot dx \geq 0;$$

$$(12) \quad F \left(0, \frac{\varrho \sin \varphi}{1^\sigma}, \frac{\varrho^2 \sin 2\varphi}{2^{2\sigma}}, \dots, \frac{\varrho^{p-1} \sin (p-1)\varphi}{(p-1)^{(p-1)\sigma}} \right) \\ = \frac{1}{2\pi i} \oint \frac{\varphi_m(x)}{x^2} \left(\sum_{\mu=0}^{p-1} \frac{\varrho^\mu \sin \mu \varphi}{x^\mu \mu^{\mu \sigma}} \right)^2 dx = \frac{1}{2\pi i} \oint \frac{\varphi_m(x)}{x^2} \cdot Y^2 \cdot dx \geq 0;$$

hier führen wir folgende Bezeichnungen ein:

$$(13) \quad X = \sum_{\mu=0}^{p-1} \frac{\varrho^\mu \cos \mu \varphi}{x^\mu \mu^{\mu \sigma}}, \quad Y = \sum_{\mu=0}^{p-1} \frac{\varrho^\mu \sin \mu \varphi}{x^\mu \mu^{\mu \sigma}}.$$

Mit den Bezeichnungen $z = \varrho(\cos \varphi + i \sin \varphi)$, $\bar{z} = \varrho(\cos \varphi - i \sin \varphi)$ folgt aus (13):

$$(14) \quad Z = X + iY = \sum_{\mu=0}^{p-1} \frac{z^\mu}{x^\mu \mu^{\mu \sigma}}, \quad \bar{Z} = X - iY = \sum_{\mu=0}^{p-1} \frac{\bar{z}^\mu}{x^\mu \mu^{\mu \sigma}}.$$

Wir wissen aber¹⁰⁾, daß man entweder im Falle $|z| < |x|$ oder im Falle $\arg z \not\equiv \arg x \pmod{2\pi}$ solche Werte von p und σ finden kann, daß die Werte Z und \bar{Z} sich beliebig wenig von den Größen

¹⁰⁾ Lindelöf, Calcul des résidus, p. 122. Paris 1905.

$$\frac{1}{1-\frac{z}{x}} = \frac{x}{x-z} \quad \text{und bzw.} \quad \frac{1}{1-\frac{\bar{z}}{x}} = \frac{x}{x-\bar{z}}$$

unterscheiden, und daß die Wahl von p und σ für alle Punkte x des Randes K gültig ist, sobald der kleinste Abstand von z (und \bar{z}) bis K nicht kleiner als eine festgewählte Größe ε ist. Da aber

$$X = \frac{z + \bar{z}}{2}, \quad Y = \frac{z - \bar{z}}{2i},$$

so schließen wir durch Grenzübergang, daß

$$(15) \quad \frac{1}{8\pi i} \cdot \oint \frac{\varphi_m(x)}{x^2} \left(\frac{x}{x-z} + \frac{x}{x-\bar{z}} \right)^2 dx \geq 0,$$

$$(16) \quad -\frac{1}{8\pi i} \cdot \oint \frac{\varphi_m(x)}{x^2} \left(\frac{x}{x-z} - \frac{x}{x-\bar{z}} \right)^2 dx \geq 0.$$

Die Addition dieser Formeln gibt uns:

$$(17) \quad \frac{1}{8\pi i} \oint \frac{\varphi_m(x)}{x^2} \cdot \frac{4x^2 \cdot dx}{(x-z)(x-\bar{z})} = \frac{1}{2\pi i(z-\bar{z})} \oint \varphi_m(x) \left(\frac{1}{x-z} - \frac{1}{x-\bar{z}} \right) dx \\ = \frac{\varphi_m(z) - \varphi_m(\bar{z})}{z-\bar{z}} = \frac{J(\varphi_m(z))}{J(z)} \geq 0,$$

w. z. b. w.

Bemerkung 1. Ist die Funktion $\varphi_m(z)$ mehrdeutig, so kann man die Ungleichung (17) für größere Bereiche nachweisen. Wir müssen dann, als „Summationsfaktor“ die Funktion $\frac{1}{n^{\mu\sigma}}$ mit komplexem σ nehmen

$\left(x_\mu = R \left(\frac{\rho^\mu e^{i\mu\varphi}}{\mu^{\mu\sigma}} \right) \right)$ und $x_\mu = J \left(\frac{\rho^\mu e^{i\mu\varphi}}{\mu^{\mu\sigma}} \right)$. So wird der Zweig der Funktion $\varphi_m(z)$ in einem krummlinigen Stern definiert, der mit Hilfe von gewissen logarithmischen Spiralen gebildet ist (Lindelöf, J. de Math., (4) 9 (1903), S. 220). Überhaupt können wir die Gültigkeit der Ungleichung (17) für den ganzen Bereich nachweisen, für den die Darstellung

$$\varphi_m(z) = \lim_{\sigma \rightarrow 0} \sum_{\mu} s_{m+\mu} \varphi(z, \mu, \sigma)$$

möglich ist.

Bemerkung 2. Ist $f(z)$ ein reelles Polynom, so ist die Ungleichung

$$J \left(-\frac{f'(z)}{f(z)} \right) \geq 0$$

eine notwendige und hinreichende Bedingung dafür, daß die Wurzeln von $f(z)$ alle reell sind. Dies folgt aus folgender Formel:

$$(18) \quad \frac{J \left(-\frac{f'(z)}{f(z)} \right)}{J(z)} = \sum_{i=1}^n \frac{(1 - \gamma_i)}{(x - \beta_i)^2 + (y - \gamma_i)^2},$$

wo $\alpha_i = \beta_i + i\gamma_i$ ($i = 1, 2, \dots, n$) Wurzeln von $f(z)$ sind. Sind $\gamma_i = 0$, so ist der Ausdruck (18) positiv. Ist dagegen etwa $\gamma_1 \neq 0$, so setzen wir $x = \beta_1$, $y = \gamma_1 + \varepsilon$. Dann wird das erste Glied der Summe (18) gleich

$$\frac{1}{\varepsilon(\gamma_1 + \varepsilon)}.$$

Nehmen wir $|\varepsilon|$ hinreichend klein und $\text{sign } \varepsilon = -\text{sign } \gamma_1$, so wird dieses Glied negativ und absolut beliebig groß, während die Werte von anderen Gliedern der Summe (18) unter einer Grenze bleiben. Dieses Kriterium gilt auch für mehrfache Wurzeln und für transzendente Funktionen endlicher Höhe. Man kann aus ihm durch Transformation Kriterien dafür herleiten, daß sich alle Wurzeln auf irgendeiner Kurve befinden. (Dies gilt nicht allgemein.)

Dieses Kriterium kann man auch so darstellen:

$$\frac{\partial}{\partial y} (|f(z)|^2) : y \geq 0,$$

was besagt, daß $|f(z)|$ als Funktion von y nur ein Extremum, nämlich das Minimum auf der reellen Achse besitzen muß.

Um den Satz 1 umzukehren, wollen wir über $\varphi_m(z)$ eine beschränkende Voraussetzung machen.

Voraussetzung. Wir betrachten den Körper $\mathfrak{K}(\varphi_m(z))$, d. h. die Gesamtheit aller Funktionen, die aus $\varphi_m(z)$ und allen reellen Polynomen mittels der vier elementaren Operationen gebildet sind. Wir nehmen an, daß es in $\mathfrak{K}(\varphi_m(z))$ keine in $z=0$ reguläre Funktion $\psi(z)$ gibt, deren imaginärer Teil negative Werte annimmt, die aber in ihrem ganzen Definitionsbereich die Ungleichung

$$J(\psi(z)) \geq -\frac{y}{x^2 + y^2}$$

befriedigt.

Die Voraussetzung ist für den Körper aller reellen Funktionen erfüllt, die die Mittag-Lefflersche Darstellung zulassen, wie ich dies später zeigen will. Ich glaube sogar nicht, daß überhaupt Funktionen vom Typus $\psi(z)$ existieren.

Satz 2. Gilt für eine reelle Funktion $\varphi_m(z)$ die Ungleichung

$$J(\varphi_m(z)) : J(z) \geq 0$$

im ganzen Definitionsbereich von $\varphi_m(z)$ und ist $\varphi_m(z)$ in $z=0$ regulär, so sind alle Formen

$$\sum_{\mu, \nu}^{0, \dots, p-1} s_{m+\mu+\nu} x_\mu x_\nu$$

positiv-definit.

Beweis. Wir betrachten folgende Darstellung von $\varphi_m(z)$:

$$\varphi_m(z) = s_{m-1} + s_m z + s_{m+1} z^2 + \dots = s_{m-1} + \frac{s_m}{\frac{1}{z} + a_1 - \psi_1(z)}.$$

Es ist $s_m > 0$. Denn $J(\varphi_m(z)):J(z) = s_m + \chi(x, y)$, wobei $\chi(x, y)$ bei genügend kleinen x, y beliebig klein wird. Wäre $s_m = 0$, so würde

$$J(\varphi_m(z)):J(z) = s_{m+k-1} \varrho^{k-1} \frac{\sin k \varphi}{\sin \varphi} + \chi_1(x, y) \quad (k > 1)$$

bei $\text{sign} \sin k \varphi = -\text{sign} s_{m+k-1}$, $\sin \varphi > 0$ und genügend kleinem ϱ negativ sein. Also ist für $y > 0$:

$$J\left(-\frac{s_m}{\varphi_m(z) - s_{m-1}}\right) = J\left(-\frac{1}{z} - a_1 + \psi_1(z)\right) = J(\psi_1(z)) + \frac{y}{x^2 + y^2} \geq 0.$$

Unsere Voraussetzung erlaubt uns, zu behaupten, daß:

$$J(\psi_1(z)) \geq 0 \quad \text{für } y \geq 0.$$

Führen wir diesen Prozeß für $\psi_1(z)$ aus und fahren wir so fort, so kommen wir schließlich zu folgender (wenn auch formeller) Kettenbruchdarstellung von $\varphi_m(z)$:

$$\varphi_m(z) = \frac{s_m}{\frac{1}{z} + a_1 - \frac{k_1}{\frac{1}{z} + a_2 - \frac{k_2}{\frac{1}{z} + a_3 - \dots}}}$$

wobei alle k_p ($p = 1, 2, \dots$) positiv sind. Es ist aber bekannt (vgl. Grommer, loc. cit., S. 123), daß man k_p so ausdrücken kann:

$$k_p = \frac{D_m^{p+1} \cdot D_m^{p-1}}{(D_m^p)^2} \quad (p = 1, 2, 3, \dots),$$

wo wir die Bezeichnungen

$$D_m^p = \begin{vmatrix} s_m & s_{m+1} & \dots & s_{m+p-1} \\ s_{m+1} & s_{m+2} & \dots & s_{m+p} \\ \dots & \dots & \dots & \dots \\ s_{m+p-1} & s_{m+p} & \dots & s_{m+2p-2} \end{vmatrix}, \quad D_m = 1 \quad (p = 1, 2, 3, \dots)$$

einführen. Daraus folgt, daß alle D_m^p positiv sein müssen, d. h. daß alle Formen (2) positiv-definit sind, w. z. b. w.

§ 2.

Ich will nun zwei Hilfssätze beweisen.

Hilfssatz 1. Für jede reelle ganze transzendente Funktion $f(z)$, die nicht eine lineare Funktion ist, nimmt der Ausdruck $J(f(z)):J(z)$ beliebig große Werte beider Vorzeichen an.

Beweis. Wir nehmen an, daß etwa:

$$J(f(z)): J(z) \leq A.$$

Für $J(z) = y \geq 0$ haben wir also:

$$J(f(z)) \leq Ay \quad (y \geq 0).$$

Dann wird die reelle ganze transzendente Funktion $g(z) = f(z) - Az$ folgender Ungleichung unterworfen:

$$J(g(z)) \leq 0 \quad \text{für } y \geq 0.$$

$g(z)$ kann nirgends außerhalb der reellen Achse verschwinden, da aus $J(g(\alpha + \beta i)) = 0$ ($\beta > 0$) folgen würde: $J(g(\alpha + \beta i + \varepsilon)) > 0$. Auf der reellen Achse kann $g(z)$ nicht mehr als einmal verschwinden, da sonst auch die Ableitung $g'(z)$ auf der reellen Achse verschwände: $g'(\gamma) = 0$, woraus folgen würde:

$$g(z) = g(\gamma) + (z - \gamma)^k \frac{g^{(k)}(\gamma)}{k!} + \dots \quad (g^{(k)}(\gamma) \neq 0, k > 1).$$

Nehmen wir $z - \gamma = \varrho e^{i\varphi}$, $\text{sign}(\sin k\varphi) = \text{sign}(g^{(k)}(\gamma))$, $\sin \varphi > 0$, ϱ genügend klein, so wird:

$$J(g(z)) = \varrho^k \sin k\varphi \frac{g^{(k)}(\gamma)}{k!} + (\varrho^{k+1}) > 0 \quad \text{für } y > 0.$$

Somit läßt $g(z)$ eine der folgenden Darstellungen zu:

$$g(z) = (z - \alpha) e^{h(z)} \quad \text{oder} \quad g(z) = e^{h(z)} \\ (\alpha \text{ reell, } h(z) = u + iv - \text{ganze Funktion, } v(0, 0) = 0).$$

Nehmen wir $z - \alpha = \varrho e^{i\varphi}$, so folgt:

$$J(g(z)) = \varrho e^u \sin(\varphi + v) \quad \text{bzw.} \quad J(g(z)) = e^u \sin v.$$

Unsere Forderung lautet

$$\frac{\sin(\varphi + v)}{\sin \varphi} \leq 0 \quad \text{bzw.} \quad \frac{\sin v}{\sin \varphi} \leq 0.$$

Da aber $\varphi + v$ (bzw. v) eine stetige Funktion von ϱ ist, so muß sie auf jedem Strahle ($\varphi = \text{konst.}$) nicht die Grenzen $-\pi$ und 0 (für $0 < \varphi < \pi$) oder 0 und $+\pi$ (für $-\pi < \varphi < 0$) überschreiten. Somit gilt auf der ganzen Ebene:

$$-2\pi < v + 2\pi,$$

d. h. $v = 0$, $u = \text{konst.}$, $g(z) = (z - \alpha) e^{u_0}$ oder bzw. $g(z) = e^{u_0}$, w. z. b. w.

Diesen Satz hat zum erstenmal Herr Krein bewiesen; Herr Grandjot hat einen anderen Beweis gegeben. Der vorliegende Beweis ist elementarer als die beiden vorigen Beweise.

Hilfssatz 2. Für jede reelle ganze transzendente Funktion $f(z)$, die nicht eine lineare Funktion ist, nimmt der Ausdruck $J(f(z))$ auf der oberen Halbebene beliebig große Werte beider Vorzeichen an.

Beweis. Der Hilfssatz 1. besagt, daß $J(f(z))$ auf der oberen Halbebene irgendwelche positive (und auch negative) Werte annimmt. Wir bezeichnen $J(f(z))$ mit $H(\rho, \varphi)$, wobei $z = \rho e^{i\varphi}$ ist, $H(\rho, \varphi)$ ist eine harmonische Funktion, die auf der reellen Achse verschwindet. Es sei C der Halbkreis vom Radius r , auf dessen Rande $H(\rho, \varphi)$ den Wert $m > 0$ annehme. Wir nehmen einen mit C konzentrischen Halbkreis K vom Radius $R > r$ und betrachten eine harmonische Hilfsfunktion $U(\rho, \varphi)$, die auf K den Wert Eins hat und auf der reellen Achse verschwindet. Es ist:

$$(19) \quad U(\rho, \varphi) = \frac{1}{2\pi} \int_0^\pi \frac{(R^2 - \rho^2) d\psi}{R^2 - 2R \cos \rho(\varphi - \psi) + \rho^2} - \frac{1}{2\pi} \int_\pi^{2\pi} \frac{(R^2 - \rho^2) d\psi}{R^2 - 2R \rho \cos(\varphi - \psi) + \rho^2}. \quad (11)$$

Es gilt auf C folgende Ungleichung:

$$U(r, \varphi) \leq \frac{1}{2\pi} \int_0^\pi \frac{(R^2 - r^2) d\psi}{(R - r)^2} - \frac{1}{2\pi} \int_0^\pi \frac{(R^2 - r^2) d\psi}{(R + r)^2} = \frac{2Rr}{R^2 - r^2}.$$

Ist nun M das Maximum von $H(\rho, \varphi)$ auf K , so nimmt die Funktion $MU(\rho, \varphi) - H(\rho, \varphi)$ auf K und auf der reellen Achse nicht-negative Werte ein. Deshalb muß sie auf C nicht-negativ sein, woraus folgt:

$$M \cdot \frac{2Rr}{R^2 - r^2} - m \geq 0.$$

oder:

$$M \geq \frac{m(R^2 - r^2)}{Rr}.$$

Somit sehen wir, daß M bei genügend großem R beliebig groß wird, w. z. b. w.

§ 3.

Nun sind wir imstande, den folgenden Satz zu beweisen:

Satz 3. Es sei $\varphi(z)$ eine reelle analytische in $z = 0$ reguläre Funktion, die durch ihre Maclaurinsche Entwicklung definiert ist:

$$(20) \quad \varphi(z) = s_1 + s_2 z + s_3 z^2 + \dots,$$

¹¹⁾ Osgood, Lehrbuch der Funktionentheorie 1, S. 640, 3. Satz. Leipzig 1912.

und seien alle quadratischen Formen

$$(21) \quad \sum_{\mu, \nu}^{0, \dots, p-1} \delta_{m+\mu+\nu} x_{\mu} x_{\nu} \quad (m \text{ fixiert, } p = 1, 2, 3, \dots)$$

positiv-definit. Hat $\varphi(z)$ im durch ihren (auch krummlinigen) Stern definierten Bereich isolierte singuläre Stellen, so müssen sie einfache reelle Pole sein. Ist α_{ν} einer dieser Pole und $-A_{\nu}$ das entsprechende Residuum, so muß sein:

$$(22) \quad \text{sign } A_{\nu} = \text{sign } \alpha_{\nu}^m.$$

Läßt $\varphi(z)$ die Mittag-Lefflersche Entwicklung zu, so muß dieselbe die Form haben:

$$(23) \quad \varphi(z) = \sum_{\mu=1}^{\infty} \psi_{\mu} \left(\frac{1}{z - \alpha_{\mu}} \right) + \psi_0(z),$$

wo

$$\psi_0(z) = - \sum_{\nu=1}^{\infty} \left(\frac{A_{\nu}}{z - \alpha_{\nu}} + A_{\nu} \beta_{\nu} + A_{\nu} \beta_{\nu}^2 z + \dots + A_{\nu} \beta_{\nu}^{m-1} z_{\nu}^{m-2} \right) + a_1 + a_2 z + \dots + a_m z^{m-1}$$

ist, wobei $\beta_{\nu} = \frac{1}{\alpha_{\nu}}$, $\text{sign } A_{\nu} = \text{sign } \alpha_{\nu}^m$, $a_m \leq 0$, während

$$\psi_{\mu}(\xi) = \sum_{\nu=1}^{\infty} \left(\frac{A_{\mu, \nu}}{\xi - \alpha_{\mu, \nu}} + A_{\mu, \nu} \beta_{\mu, \nu} \right) - a_{\mu, 1} - a_{\mu, 2} \xi \quad (\mu = 1, 2, 3, \dots)$$

sind, wobei $\beta_{\mu, \nu} = \frac{1}{\alpha_{\mu, \nu}}$, $A_{\mu, \nu} > 0$, $a_{\mu, 2} \geq 0$.

Beweis. Wir nehmen an, die Behauptungen des Satzes seien nicht erfüllt. Dann sind folgende Fälle möglich:

A. Wenn die Voraussetzungen des ersten Teiles erfüllt sind:

1. Fall. $\varphi(z)$ hat eine imaginäre isolierte singuläre Stelle.
2. Fall. $\varphi(z)$ hat eine reelle isolierte singuläre Stelle.
3. Fall. $\varphi(z)$ hat einen reellen mehrfachen Pol.

B. Wenn die Voraussetzungen des zweiten Teiles erfüllt sind:

4. Fall. Der *Konvergenzexponent* der einen Funktion $\psi_{\mu}(\xi)$ ($\mu = 0, 1, 2, 3, \dots$) ist unendlich oder größer als m .

5. Fall. Die Entwicklung von $\psi_{\mu}(\xi)$ enthält eine ganze transzendente Funktion.

6. Fall. Die Entwicklung von $\psi_{\mu}(\xi)$ ($\mu = 1, 2, 3, \dots$) und bzw. von $\psi_0(z)$ enthält ein Polynom vom Grade > 1 oder bzw. $> m - 1$.

7. Fall. Die Entwicklung von $\psi_{\mu}(\xi)$ ($\mu = 1, 2, 3, \dots$) und bzw. von $\psi_0(z)$ enthält ein Polynom mit *positivem* Koeffizienten bei z oder bzw. mit *negativem* Koeffizienten bei z^{m-1} .

Unser Satz wird bewiesen, wenn wir zeigen, daß in jedem dieser Fälle eine der Formen (21) nicht positiv-definit ist. Statt dessen beweisen wir in den meisten Fällen, daß der Ausdruck $J(\varphi_m(z)):J(z)$, wo $\varphi_m(z) = s_m z + s_{m+1} z^2 + \dots$, negative Werte annimmt. Dies genügt, wegen der Sätze 1. und 2., um zu behaupten, daß nicht alle Formen (21) positiv-definit sind.

A. Es gelten die Voraussetzungen des ersten Teiles des Satzes.

1. Fall. Es sei α eine komplexe isolierte singuläre Stelle von $\varphi(z)$, und daher auch von $\varphi_m(z)$. Wir können $\varphi_m(z)$ so darstellen:

$$(24) \quad \varphi_m(z) = \chi(\xi) + \psi(z),$$

wo $\chi(\xi)$ eine ganze transzendente Funktion (oder Polynom) von $\xi = \frac{1}{z-\alpha}$ ist, während $\psi(z)$ sich im Innern eines Kreises K um α mit dem Radius ϱ regulär verhält. Es sei $\alpha = \beta + i\gamma$, wo $\gamma \neq 0$ ist. Dann ist:

$$(25) \quad J(\varphi_m(z)):J(z) = \frac{1}{y} [J(\chi(\xi)) + J(\psi(z))].$$

Nehmen wir $|z - \alpha| < \varrho$, $|\xi| > \frac{1}{\varrho}$. Man kann wegen des Weierstraßschen Satzes ξ im Bereiche $|\xi| > \frac{1}{\varrho}$ so wählen, daß $J(\chi(\xi))$ einen beliebig großen absoluten Wert annimmt, und daß $\text{sign } J(\chi(\xi)) = -\text{sign } \gamma$. Da man dabei $|z - \alpha|$ so klein wählen kann, daß $\text{sign } y = \text{sign } \gamma$ gilt, und $J(\psi(z))$ für $|z - \alpha| < \varrho$ endlich bleibt, so muß der Ausdruck (25) negative Werte annehmen, w. z. b. w.

2. Fall. α ist reell. Dann ist $\chi(\xi)$ eine reelle ganze transzendente Funktion. Man kann wegen des Hilfssatzes 2. z so wählen, daß $y > 0$, $\chi(\xi) < 0$ und $|\chi(\xi)|$ beliebig groß wird. Der Ausdruck muß demnach einen negativen Wert annehmen, w. z. b. w.

3. Fall. α ist reeller n -facher Pol ($n > 1$) von $\varphi(z)$ (und auch von $\varphi_m(z)$). $\chi(\xi)$ ist ein reelles Polynom n -ten Grades:

$$\chi(\xi) = a_0 + a_1 \xi + \dots + a_n \xi^n.$$

Wir nehmen $\xi = \varrho e^{i\varphi}$. Dann ist:

$$J(\chi(\xi)) = \varrho^n \left(a_n \sin n\varphi + \frac{a_{n-1}}{\varrho} \sin(n-1)\varphi + \dots + \frac{a_1}{\varrho^{n-1}} \sin \varphi \right).$$

Man nehme ϱ so groß, daß die Größen $\frac{a_{n-1}}{\varrho}$, $\frac{a_{n-2}}{\varrho^2}$, ..., $\frac{a_1}{\varrho^{n-1}}$ beliebig klein werden, und wähle φ so, daß $\sin \varphi > 0$, $a_n \sin n\varphi < 0$ wird. Dann wird der Wert von $\chi(\xi)$ negativ und absolut beliebig groß. Da $\psi(z)$ endlich bleibt, so folgt, daß $J(\varphi_m(z)):J(z) < 0$ ist, w. z. b. w.

Ist $n = 1$, d. h. ist α einfacher Pol, so gilt:

$$J(\varphi_m(z)):J(z) = -\frac{a_1}{(x-\alpha)^2+y^2} + \frac{1}{y}J(\psi(z)).$$

Damit $J(\varphi_m(z)):J(z) \geq 0$ sei, muß $a_1 < 0$ sein.

B. Wir nehmen an, die Voraussetzungen des zweiten Teiles des Satzes seien erfüllt, d. h. $\varphi(z)$ läßt die Mittag-Lefflersche Entwicklung zu. Dann zeigen uns die früheren Überlegungen, daß die Unstetigkeitsstellen von $\varphi(z)$ lauter reelle einfache Pole sein müssen. Die Mittag-Lefflersche Entwicklung ist andererseits nur dann möglich, wenn die Menge dieser Pole abzählbar ist. Es seien $\bar{\alpha}_0 = \infty, \bar{\alpha}_1, \bar{\alpha}_2, \dots$ ihre Häufungspunkte. Dann kann $\varphi(z)$ so dargestellt werden:

$$(26) \quad \varphi(z) = \sum_{\mu=1}^{\infty} \psi_{\mu}\left(\frac{1}{z-\bar{\alpha}_{\mu}}\right) + \psi_0(z),$$

wobei

$$(27) \quad \psi_{\mu}(\xi) = \sum_{\nu=1}^{\infty} \left(\frac{A_{\mu,\nu}}{\xi-\alpha_{\nu}} + g_{\mu,\nu}(\xi) \right) \quad (\mu = 1, 2, 3, \dots),$$

$$(27a) \quad \psi_0(z) = -\sum_{\nu=1}^{\infty} \left(\frac{A_{\nu}}{z-\alpha_{\nu}} + g_{\nu}(z) \right) + g(z),$$

während $g_{\mu,\nu}(\xi)$ Polynome sind.

Nun sind wir imstande, unsere *Voraussetzung* über das Vorzeichen von $J(\varphi(z))$ für den Fall der die Mittag-Lefflersche Entwicklung zulassenden Funktionen zu beweisen. Wir können offenbar die Punkte $\bar{\alpha}_{\nu}$ so wählen, daß unter ihnen kein Element der zweiten derivierten Menge der Menge (α_{ν}) vorkommt. Dann kann man $\varphi(z)$ stets folgendermaßen darstellen:

$$\varphi(z) = \psi_{\mu}(\xi) + \chi(z),$$

wo $\psi_{\mu}(\xi)$ eine der Funktionen (27) ist, während $\chi(z)$ in $z = \bar{\alpha}_{\mu}$ regulär bleibt. Man muß dabei nicht vergessen, daß $\varphi(z)$ nur einfache reelle Pole haben darf, da sonst $J(\varphi(z))$ beliebig große Werte beider Vorzeichen annimmt, wie wir dies bei Untersuchung der Fälle 1, 2, 3 gesehen haben.

Wir setzen voraus, $J(\varphi(z))$ nehme auf der oberen Halbebene positive Werte an. Dann muß dies für den einen ihrer Bestandteile, etwa für $\psi_{\mu}(\xi)$ ($\mu = 0, 1, 2, 3, \dots$), der Fall sein. Wir nehmen an, es sei $J(\psi_{\mu}(\xi_0)) > 0, J(z_0) > 0, J(\xi_0) < 0$. Wir verfahren so, wie beim Beweis des Hilfssatzes 2., doch umkreisen wir die Pole von $\psi_{\mu}(\xi)$, die sich nur im Unendlichen häufen, mit hinreichend kleinen Halbkreisen. Nimmt $J(\psi_{\mu}(\xi))$ auf einem dieser Halbkreise positive Werte an, so ist alles

bewiesen. Denn dann wächst $J\left(\frac{A_{\mu, \nu}}{\xi - \alpha_\nu}\right) = -\frac{A_{\mu, \nu} v}{(u - \alpha_\nu)^2 + v^2}$ ($\xi = u + vi$) bei abnehmendem $|\xi - \alpha_\nu|$ ins Unendliche, während die übrigbleibenden Glieder endlich bleiben. Daher nehmen wir an, seine Werte auf diesen Halbkreisen seien negativ ($A_{\mu, \nu} < 0$) und $J\left(\frac{A_{\mu, \nu}}{\xi - \alpha_\nu}\right) : v$ hinreichend groß. Da aber die Hilfsfunktion $U(u, v)$ auf diesen Halbkreisen nur positive Werte annimmt¹²⁾, so bleibt die Funktion $M \cdot U - J(\psi_\mu(\xi))$ auf dem ganzen Rande nicht negativ. Danach können wir alle Überlegungen des Beweises des Satzes 2. anwenden.

4. Fall. Wir bemerken, daß der Übergang von $\varphi(z) = s_1 + s_2 z + \dots$ zu $\varphi_m(z) = s_{m-1} + s_m z + \dots$ jedes Residuum bei α_ν um $\frac{1}{\alpha_\nu^{m-2}}$ vergrößert. Häufen sich die Pole um einen endlichen Punkt $\bar{\alpha}$, so sind die Summen $\sum_{\nu=1}^{\infty} \frac{A_\nu}{\alpha_\nu^2}$ und $\sum_{\nu=1}^{\infty} \frac{A_\nu}{\alpha_\nu^m}$ gleichzeitig konvergent oder divergent; häufen sich dagegen die Pole im Unendlichen, so nimmt dabei der Konvergenzexponent um $(m-2)$ ab. Dies erlaubt uns, den Fall zu betrachten, daß in der Entwicklung

$$(28) \quad \varphi_m(z) = \sum_{\mu=1}^{\infty} \psi'_\mu(\xi_\mu) + \psi'_0(z)$$

der Konvergenzexponent einer der Funktionen $\psi'_\mu(\xi_\mu)$ ($\mu = 0, 1, 2, \dots$) größer als 2 ist. Es sei $\psi'_\lambda(\xi_\lambda)$ eine solche Funktion. Wir führen in (28) die Transformation $\xi_\lambda = \frac{1}{z - \bar{\alpha}_\lambda}$ aus:

$$(29) \quad \varphi_m\left(\bar{\alpha}_\lambda + \frac{1}{\xi_\lambda}\right) = \psi'_\lambda(\xi_\lambda) + \chi(\xi_\lambda),$$

wo $\chi(\xi_\lambda)$ im Unendlichen regulär bleibt. Es genügt, zu beweisen, daß der Ausdruck $J\left(\varphi_m\left(\bar{\alpha}_\lambda + \frac{1}{\xi_\lambda}\right)\right) : J(\xi_\lambda) = -[(x - \bar{\alpha}_\lambda)^2 + y^2] \cdot J(\varphi_m(z)) : J(z)$ positive Werte annimmt. Dazu wählen wir eine positive Größe M fest und nehmen einen solchen Radius R , daß bei $|\xi_\lambda| > R$ die Ungleichung gilt:

$$\left| \frac{J(\chi(\xi_\lambda))}{J(\xi_\lambda)} \right| < M.$$

Es sei ferner A der maximale Wert von $|J(\psi'_\lambda(\xi_\lambda)) : J(\xi_\lambda)|$ in $|\xi_\lambda| \leq R$ (wir setzen noch voraus, daß $\psi'_\lambda(\xi_\lambda)$ in $|\xi_\lambda| \leq R$ keinen Pol hat). Wenn wir beweisen, daß $J(\psi'_\lambda(\xi_\lambda) - S \cdot \xi_\lambda) : J(\xi_\lambda)$, wo $S = M + A$ ist, positive Werte annimmt, so ist mithin alles bewiesen. Denn aus

¹²⁾ Wir nehmen an, daß der Halbkreis k nicht Pole auf seinem Rande enthält.

$$J(\psi'_\lambda(\xi_\lambda) - S \cdot \xi_\lambda) : J(\xi_\lambda) > 0 \quad \text{folgt} \quad J(\psi'_\lambda(\xi)) : J(\xi) > M + A,$$

woraus man schließt, daß

$$|\xi| > R, \quad J(\chi(\xi)) : J(\xi) < M, \quad J(\psi'_\lambda(\xi) + \chi(\xi)) : J(\xi) > A > 0$$

ist, w. z. b. w.

Es sei $\psi'_\lambda(\xi) = -t_1 - t_2 \xi - \dots$. Es ist nun zu beweisen, daß $J(-\psi'_\lambda(\xi) + S\xi) : J(\xi)$ negative Werte annimmt. Dazu genügt, wegen des Satzes 2., zu beweisen, daß nicht alle Formen

$$(30) \quad F(x_0, x_1, \dots, x_{p-1}) = Sx_0^2 + \sum_{\mu, \nu}^{0, \dots, p-1} t_{2+\mu+\nu} x_\mu x_\nu \quad (p = 2, 3, \dots)$$

positiv-definit sind.

Ordnen wir die Pole α_ν von $\psi'(\xi)$ nach wachsenden absoluten Beträgen:

$$|\alpha_1| \leq |\alpha_2| \leq |\alpha_3| \leq \dots$$

und setzen $\beta_\nu = \frac{1}{\alpha_\nu}$, so gilt folgende asymptotische Formel (vgl. I, § 1, (5)):

$$(31) \quad t_n = a_1 \beta_1^n + a_2 \beta_2^n + \dots + a_k \beta_k^n + \beta_k^n \cdot \varepsilon_{n,k},$$

wobei $\lim_{n \rightarrow \infty} \varepsilon_{n,k} = 0$ und $|\beta_{k+1}| < |\beta_k|$ ist.

Alle a_k sind positiv, da sonst die Formen $\sum_{\mu, \nu}^{0, p-1} t_{n+\mu+\nu} x_\mu x_\nu$ bei genügend großem n nicht positiv-definit wären. Wir nehmen eine gerade Zahl h so groß, daß in (31) für $n \geq 2 + h$ die Ungleichung $|\varepsilon_{n,k}| < \varepsilon_h$ zutrifft, wo ε_h eine genügend kleine Größe ist; ferner setzen wir $p = qh$ und $x_\mu = 0$, wenn $\mu \not\equiv 0 \pmod{h}$, $x_{\nu h} = y_\nu$. Dann kann man die Form (30) so darstellen:

$$(32) \quad \left(S + t_2 - \sum_{\nu=1}^k a_\nu \beta_\nu^2 \right) y_0^2 + \sum_{\nu=1}^k a_\nu \beta_\nu^2 (y_0 + y_1 \beta_\nu^h + \dots + y_{q-1} \beta_\nu^{h(q-1)})^2 \\ + \sum_{\mu, \nu}^{0, \dots, q-1} \beta_k^{2+(\mu+\nu)h} \varepsilon_{2+(\mu+\nu)h, k} y_\mu y_\nu.$$

(Der Strich bedeutet, daß die Kombination $\mu = 0, \nu = 0$ ausgeschlossen ist.)

Nun nehmen wir für y_ν die den Gleichungen

$$(33) \quad \begin{aligned} y_0 + y_1 \beta_1^h + \dots + y_{q-1} \beta_1^{h(p-1)} &= 0, \\ y_0 + y_1 \beta_2^h + \dots + y_{q-1} \beta_2^{h(p-1)} &= 0, \\ \dots & \\ y_0 + y_1 \beta_{q-1}^h + \dots + y_{q-1} \beta_{q-1}^{h(p-1)} &= 0 \end{aligned}$$

genügenden Werte und dabei $y_0 = 1$. Dann hat das Polynom

$$(34) \quad Y(t) = 1 + y_1 t + y_2 t^2 + \dots + y_{q-1} t^{q-1}$$

die Größen $\beta_1^h, \beta_2^h, \dots, \beta_{q-1}^h$ zu Wurzeln und ist daher so darstellbar:

$$(35) \quad Y(t) = \left(1 - \frac{t}{\beta_1^h}\right) \left(1 - \frac{t}{\beta_2^h}\right) \dots \left(1 - \frac{t}{\beta_{q-1}^h}\right).$$

Da aber $|\beta_\nu| < |\beta_\mu|$ ($\mu = 1, 2, \dots, p-1$; $\nu = p, p+1, \dots, k$) und h gerade ist, so gilt:

$$Y(\beta_\nu^h) < 1 \quad (\nu = p, p+1, \dots, k).$$

Der Ausdruck (32) nimmt daher folgenden Wert an:

$$(36) \quad \left(S + t_2 - \sum_{\nu=1}^k a_\nu \beta_\nu^2\right) + \sum_{\nu=q}^k a_\nu \beta_\nu^2 Y(\beta_\nu^h) + \sum_{\mu, \nu}^{0, \dots, q-1} \beta_k^{2+(\mu+\nu)h} \varepsilon_{2+(\mu+\nu)h, k} y_\mu y_\nu \\ < S + t_2 - \sum_{\nu=1}^{q-1} a_\nu \beta_\nu^2 + \sum_{\mu, \nu}^{0, \dots, q-1} \beta_k^{2+(\mu+\nu)h} \varepsilon_{2+(\mu+\nu)h, k} y_\mu y_\nu.$$

Die Werte y_μ genügen folgenden Ungleichungen:

$$(37) \quad |y_1| = \sum_{\nu=1}^{q-1} \frac{1}{\beta_\nu^h} < \binom{q-1}{1} \cdot \frac{1}{\beta_{q-1}^h}, \\ |y_2| = \sum_{\mu, \nu}^{1, \dots, q-1} \frac{1}{\beta_\mu^h} \frac{1}{\beta_\nu^h} < \binom{q-1}{2} \cdot \frac{1}{\beta_{q-1}^{2h}}, \\ \dots \dots \dots \\ |y_{q-1}| = \frac{1}{\beta_1^h \beta_2^h \dots \beta_{q-1}^h} < \frac{1}{\beta_{q-1}^{(q-1)h}},$$

und darum kann die letzte Summe von (36) folgendermaßen abgeschätzt werden:

$$(38) \quad \left| \sum_{\mu, \nu}^{0, \dots, q-1} \beta_k^{2+(\mu+\nu)h} \varepsilon_{m+(\mu+\nu)h, k} y_\mu y_\nu \right| < \varepsilon_h \beta_k^2 \sum_{\mu, \nu}^{0, \dots, q-1} \binom{q-1}{\mu} \binom{q-1}{\nu} \left(\frac{\beta_k}{\beta_{q-1}}\right)^{h\mu} \cdot \left(\frac{\beta_k}{\beta_{q-1}}\right)^{h\nu} \\ = \varepsilon_h \beta_k^2 \left[1 + \left(\frac{\beta_k}{\beta_{q-1}}\right)^h\right]^{2(q-1)} < \varepsilon_h \beta_k^2 2^{2(q-1)}.$$

Der Faktor ε_h kann für große Werte von h beliebig klein werden, während die übrigen Faktoren nicht von h abhängen. Die Summe (38) kann also beliebig klein werden.

Die übrigbleibenden Glieder $S + t_2 - \sum_{\nu=1}^{q-1} a_\nu \beta_\nu^2$ können für große Werte von q negativ und absolut beliebig groß werden, da $\sum_{\nu=1}^{\infty} a_\nu \beta_\nu^2$, unserer Voraussetzung gemäß, divergiert. Daraus folgt, daß die Form (30) nicht positiv-definit ist, w. z. b. w.¹³⁾

¹³⁾ Im Falle $\lambda = 0$ beweist man ebenso, daß der Ausdruck $J(\psi'_0(z)):J(z)$ negative Werte annehmen kann.

5. Fall. Die obigen Überlegungen zeigen, daß man in $\psi(z)$ bei den Gliedern $\frac{A}{\xi - \alpha}$ bzw. $\frac{A}{z - \alpha}$ Konvergenzpolynome $g_{\mu, \nu}(\xi)$ von nicht höherem als vom 0-ten bzw. $(m - 2)$ -ten Grade nehmen kann. Wir betrachten nun den Fall, daß in der Formel (27a) $g(z)$ eine ganze transzendente Funktion ist. Dasselbe gilt für die Funktion $\varphi_{m+2}(z) = s_{m+1} + s_{m+2}z + \dots$, die somit in folgender Gestalt darstellbar ist:

$$\varphi_{m+2}(z) = - \sum_{\nu=1}^{\infty} \frac{A_{\nu} \beta_{\nu}^m}{z - \alpha_{\nu}} + \bar{g}(z) + \chi(z),$$

wobei $\bar{g}(z)$ eine reelle ganze transzendente Funktion ist, während $\chi(z)$ im Unendlichen regulär bleibt.

Wir führen folgende Bezeichnungen ein:

$$\psi(z) = - \sum_{\nu=1}^{\infty} \frac{A_{\nu} \beta_{\nu}^m}{z - \alpha_{\nu}} + \bar{g}(z), \quad \bar{g}(z) = \sum_{\nu=0}^{\infty} c_{\nu} z^{\nu}, \quad z = \rho e^{i\varphi}.$$

Dann gilt für $J(z) = \rho \sin \varphi > 0$:

$$\begin{aligned} (39) \quad & \frac{1}{R} \int_0^R \frac{J(\psi(\rho e^{i\varphi}))}{J(\rho e^{i\varphi})} \cdot d\rho \\ &= \sum_{\nu=1}^{\infty} \frac{1}{R} \int_0^R \frac{A_{\nu} \beta_{\nu}^m d\rho}{(\rho \cos \varphi - \alpha_{\nu})^2 + \rho^2 \sin^2 \varphi} + \sum_{\nu=1}^{\infty} \frac{c_{\nu}}{R} \int_0^R \rho^{\nu-1} \frac{\sin \nu \varphi}{\sin \varphi} d\rho \\ &= \frac{1}{R \sin \varphi} \sum_{\nu=1}^{\infty} \frac{A_{\nu} \beta_{\nu}^m}{\alpha_{\nu}} \int_{-\operatorname{ctg} \varphi}^{\frac{R - \alpha_{\nu} \cos \varphi}{\alpha_{\nu} \sin \varphi}} \frac{du}{u^2 + 1} + \frac{1}{R \sin \varphi} \cdot \sum_{\nu=1}^{\infty} \frac{c_{\nu}}{\nu} \cdot R^{\nu} \sin \nu \varphi \\ &< \frac{1}{R \sin \varphi} \left[\pi \sum_{\nu=1}^{\infty} |A_{\nu}| \cdot |\beta_{\nu}|^{m+1} + J(h(R e^{i\varphi})) \right] = \frac{J(\pi M R e^{i\varphi} + h(R e^{i\varphi}))}{J(R e^{i\varphi})}, \end{aligned}$$

wobei $h(z) = \sum_{\nu=1}^{\infty} \frac{c_{\nu}}{\nu} \cdot z^{\nu}$, $M = \sum_{\nu=1}^{\infty} |A_{\nu} \beta_{\nu}^{m+1}|$ ist. Dieser Ausdruck kann wegen des Hilfssatzes 1. negative absolut beliebig große Werte annehmen. Da aber

$$\frac{1}{R} \int_0^R \frac{J(\psi(\rho e^{i\varphi}))}{J(\rho e^{i\varphi})} d\rho > \left[\frac{J(\psi(\rho e^{i\varphi}))}{J(\rho e^{i\varphi})} \right]_{\min},$$

so gilt dasselbe für den Ausdruck

$$\frac{J(\psi(\rho e^{i\varphi}))}{J(\rho e^{i\varphi})} = \frac{J(\psi(z))}{J(z)}.$$

Dies kann nur für hinreichend große Werte von $|z|$ eintreten, für welche $\chi(z)$, sowie auch $\frac{J(\chi(z))}{J(z)}$, beschränkt bleibt. Also gibt es Werte von z , für welche gilt:

$$J(\psi(z) + \chi(z)) : J(z) = J(\varphi_{m+2}(z)) : J(z) < 0.$$

Daraus folgt, daß nicht alle Formen $\sum_{\mu, \nu}^{0, \dots, p-1} s_{m+2+\mu+\nu} x_\mu x_\nu$ positiv-definit sind. Wir haben aber gesehen, daß dann auch nicht alle Formen $\sum_{\mu, \nu}^{0, \dots, p-1} s_{m+\mu+\nu} x_\mu x_\nu$ positiv-definit sind, w. z. b. w. (vgl. I, § 2).

6. Fall. $g(z)$ ist ein Polynom vom Grade $> m - 1$. Das entsprechende Polynom $\bar{g}(z)$ von $\varphi_m(z)$ ist dann vom Grade $p > 1$. Dann gilt:

$$\frac{J(\psi_m(z))}{J(z)} = \sum_{\nu=1}^{\infty} \frac{A_\nu \beta_\nu^{m-2}}{(\rho \cos \varphi - a_\nu)^2 + \rho^2 \sin^2 \varphi} + \sum_{\nu=1}^p c_\nu \rho^{\nu-1} \frac{\sin \nu \varphi}{\sin \varphi} + \frac{J(\chi(z))}{J(z)}.$$

Man kann stets φ so wählen, daß gilt: $\sin \varphi \geq \frac{1}{\sqrt{2}}$, $\text{sign} \sin p \varphi = -\text{sign} c_p$. Dann gilt:

$$\frac{J(\psi(z))}{J(z)} < 2 \sum_{\nu=1}^{\infty} A_\nu \beta_\nu^m + \rho^{p-1} \left(a_p \frac{\sin p \varphi}{\sin \varphi} + \dots + \frac{a_1}{\rho^{p-1}} \right) + \frac{J(\chi(z))}{J(z)}.$$

Es ist klar (vgl. den 3. Fall), daß das zweite Glied bei genügend großem ρ negativ und absolut beliebig groß wird, während die übrigen Glieder beschränkt bleiben, w. z. b. w.

7. Fall. $p=1$, $a_1 < 0$. Dann ist:

$$J(\varphi_m(z)) : J(z) = a_1 + \sum_{\nu=1}^{\infty} \frac{A_\nu \beta_\nu^m}{(\rho \beta_\nu - \cos \varphi)^2 + \sin^2 \varphi} + \frac{J(\chi(z))}{J(z)}.$$

Das erste Glied ist negativ, das dritte Glied wird beliebig klein, wenn wir $J(z)$ genügend groß nehmen. Um das zweite Glied abzuschätzen, verfahren wir folgendermaßen. Wir nehmen $\varphi = \frac{\pi}{4}$, $\sin \varphi = \cos \varphi = \frac{1}{\sqrt{2}}$. Ferner wählen wir eine beliebig große Zahl M und nehmen den Index n so groß, daß die Summe

$$2 \sum_{\nu=n+1}^{\infty} A_\nu \beta_\nu^m < \sum_{\nu=1}^{\infty} A_\nu \beta_\nu^m : M^2$$

wird. Setzen wir $\rho = \frac{M + \cos \varphi}{|\beta_n|}$, so ist $|\rho \beta_\nu - \cos \varphi| \geq M$ für $\nu = 1, 2, \dots, n$.

Das zweite Glied zerfällt in zwei Teile:

$$\sum_{\nu=1}^{\infty} \frac{A_\nu \beta_\nu^m}{(\rho \beta_\nu - \cos \varphi)^2 + \sin^2 \varphi} = \sum_{\nu=1}^n \frac{A_\nu \beta_\nu^m}{(\rho \beta_\nu - \cos \varphi)^2 + \sin^2 \varphi} + \sum_{\nu=n+1}^{\infty} \frac{A_\nu \beta_\nu^m}{(\rho \beta_\nu - \cos \varphi)^2 + \sin^2 \varphi}.$$

Der erste Teil ist kleiner als

$$\frac{\sum_{\nu=1}^n A_{\nu} \beta_{\nu}^m}{M^2} < \frac{\sum_{\nu=1}^{\infty} A_{\nu} \beta_{\nu}^m}{M^2}.$$

Der zweite Teil ist kleiner als

$$\frac{1}{\sin^2 \varphi} \sum_{\nu=n+1}^{\infty} A_{\nu} \beta_{\nu}^m = 2 \sum_{\nu=n+1}^{\infty} A_{\nu} \beta_{\nu}^m < \frac{\sum_{\nu=1}^{\infty} A_{\nu} \beta_{\nu}^m}{M^2}.$$

Somit ist das zweite Glied kleiner als

$$\frac{2 \sum_{\nu=1}^{\infty} A_{\nu} \beta_{\nu}^m}{M^2},$$

woraus folgt, daß es mit wachsendem M nach Null strebt, w. z. b. w.

Man beachte, daß a_1 der Koeffizient bei z^{m-1} in $g(z)$ ist. Somit ist der Satz 3. in allen seinen Teilen bewiesen.

Ist von vornherein bekannt, daß alle $A_{\nu} > 0$ sind (wie dies z. B. für logarithmische Derivierte von ganzen transzendenten Funktionen der Fall ist), so schließen wir, daß bei ungeradem m alle Pole α_{ν} positiv sind.

Ist $m=1$ oder $m=2$, so ist leicht zu sehen, daß die Nullstellen von $f(z)$ sämtlich reell sind und daß sie und die Pole sich gegenseitig trennen. In der Tat, in diesem Falle gilt: $J(f(z)) : J(z) \geq 0$, woraus folgt:

$$J\left(-\frac{1}{\bar{f}(z)}\right) : J(z) \geq 0,$$

was uns erlaubt, alle früheren Überlegungen auf $-\frac{1}{\bar{f}(z)}$ anzuwenden. Die Positivität der Residuen in beiden Funktionen zeigt, daß sich ihre Nullstellen und Pole gegenseitig trennen.

(Eingegangen am 17. 12. 1926.)

Zusatz bei der Korrektur. Während des Druckes dieser Arbeit erschienen in C. R. drei Arbeiten von Herrn Krawtchouk, der unabhängig von mir den größeren Teil meiner Ergebnisse erhalten hat.

Über gewisse notwendige Determinantenkriterien für die Fortsetzbarkeit einer Potenzreihe.

Von

G. Pólya in Zürich.

Man betrachte die Potenzreihe¹⁾

$$(1) \quad \frac{a_0}{z} + \frac{a_1}{z^2} + \frac{a_2}{z^3} + \dots + \frac{a_n}{z^{n+1}} + \dots = f(z)$$

und bilde aus ihren Koeffizienten die Determinante

$$(2) \quad A_n^{(k)} = \begin{vmatrix} a_n & a_{n+1} & \dots & a_{n+k-1} \\ a_{n+1} & a_{n+2} & \dots & a_{n+k} \\ \dots & \dots & \dots & \dots \\ a_{n+k-1} & a_{n+k} & \dots & a_{n+2k-2} \end{vmatrix}$$

mit Anfangselement a_n und Kolonnenzahl k ($n = 0, 1, 2, \dots$; $k = 1, 2, 3, \dots$). Es sind verschiedene Zusammenhänge zwischen diesen Determinanten $A_n^{(k)}$ und der analytischen Natur der durch die Reihenentwicklung (1) definierten Funktion $f(z)$ bekannt. Man kann mittels der $A_n^{(k)}$ die notwendige und hinreichende Bedingung dafür ausdrücken, daß $f(z)$ eine rationale Funktion sei²⁾. Man kann ferner, nach Hadamard³⁾, mittels der $A_n^{(k)}$ die notwendige und hinreichende Bedingung dafür ausdrücken, daß $f(z)$ außerhalb eines zum Konvergenzkreis der Reihe (1) konzentrischen Kreises von kleinerem Radius meromorph bleibe. Man kann schließlich mittels der $A_n^{(k)}$ notwendige Bedingungen dafür angeben, daß $f(z)$ über den Konver-

¹⁾ Es ist an und für sich gleichgültig, ob man ein Funktionselement betrachtet, das im Punkt $z = 0$ regulär ist, wie es gewöhnlich geschieht, oder eines, das im Punkt $z = \infty$ regulär ist, wie es hier geschieht. Ich mache den Übergang von z zu $\frac{1}{z}$ jetzt, um ihn nicht später, bei der Betrachtung der Tschebyscheffschen Polynome, an unbequemerer Stelle machen zu müssen.

²⁾ Vgl. z. B. G. Pólya und G. Szegő, Aufgaben und Lehrsätze aus der Analysis (Berlin 1925), Bd. 2, S. 102–105, 304–306.

³⁾ Vgl. J. Hadamard, La série de Taylor et son prolongement analytique (2. Aufl., unter Mitwirkung von M. Mandelbrojt) (Paris 1926), Bd. 1, S. 77–83.

genzkreis der Reihe (1) hinaus fortsetzbar sei. Solche Bedingungen sind in den Untersuchungen von F. Carlson und des Verfassers über Potenzreihen mit ganzzahligen Koeffizienten implizite enthalten ⁴⁾ und sollen hier in vervollständigter Form explizite dargestellt werden.

Um diese Bedingungen richtig formulieren zu können, müssen wir zuerst eine beschränkte und abgeschlossene Punktmenge \mathfrak{A} in der Ebene der komplexen Zahlen betrachten. Es sei $T_n(z) = z^n + \dots$ das zur Punktmenge \mathfrak{A} gehörige *Tschebyscheffsche Polynom* n -ten Grades, d. h. dasjenige Polynom n -ten Grades mit höchstem Koeffizienten 1, das in \mathfrak{A} am wenigsten von 0 abweicht. D. h. das Maximum τ_n von $|T_n(z)|$ in \mathfrak{A} wird von dem Maximum keines anderen Polynoms von demselben Grad und demselben höchsten Koeffizienten in derselben Punktmenge unterboten. Es existiert nach M. Fekete ⁵⁾ der Grenzwert

$$\lim_{n \rightarrow \infty} \sqrt[n]{\tau_n} = \tau$$

und wird als „transfiniten Durchmesser“ von \mathfrak{A} bezeichnet ⁶⁾.

Das Ziel dieser Arbeit ist der Beweis von folgendem Satz:

Es sei \mathfrak{A} eine beschränkte und abgeschlossene Punktmenge, deren Komplementärmenge zusammenhängend ist.

Es sei ϱ der Radius des kleinsten Kreises vom Mittelpunkt $z = 0$, der die abgeschlossene Punktmenge \mathfrak{A} enthält, σ der Radius des kleinsten Kreises von demselben Mittelpunkt, der den perfekten Kern von \mathfrak{A} enthält, und τ der transfinite Durchmesser von \mathfrak{A} .

Es sei die Punktmenge \mathfrak{A} der Funktion $f(z)$ so angepaßt, daß ϱ der Konvergenzradius der Reihe (1) ist und $f(z)$ in der Komplementärmenge von \mathfrak{A} regulär und eindeutig bleibt.

Falls n so mit k verbunden ist, daß

$$(3) \quad \lim_{k \rightarrow \infty} \frac{k-1}{n+k-1} = \varkappa > 0$$

⁴⁾ a) G. Pólya, Über Potenzreihen mit ganzzahligen Koeffizienten, *Math. Annalen* 77 (1916), S. 497–513; b) F. Carlson, Über Potenzreihen mit ganzzahligen Koeffizienten, *Math. Zeitschr.* 9 (1921), S. 1–13; c) G. Pólya, Sur les séries entières à coefficients entiers, *Proceedings of the London Math. Society* (2) 21 (1922), p. 22–38. Vgl. ferner d) G. Pólya, Arithmetische Eigenschaften und analytischer Charakter, *Jahresbericht d. deutsch. Math. Ver.* 31 (1922), S. 107–115 und e) G. Szegő, Tschebyscheffsche Polynome und nicht-fortsetzbare Potenzreihen, *Math. Annalen* 87 (1922), S. 90–111.

⁵⁾ M. Fekete, Über die Verteilung der Wurzeln bei gewissen algebraischen Gleichungen mit ganzzahligen Koeffizienten, *Math. Zeitschr.* 17 (1923), S. 228–249. Ein wesentlicher Spezialfall schon bei G. Faber, Über Tschebyscheffsche Polynome, *Journal für die reine u. angew. Math.* 150 (1920), S. 70–106. Vgl. ferner G. Szegő, Bemerkungen zu einer Arbeit von Herrn M. Fekete, *Math. Zeitschr.* 21 (1924), S. 203–208.

⁶⁾ Der transfinite Durchmesser der leeren Menge sei, definitionsgemäß, = 0.

ausfällt, ist

$$(4) \quad \overline{\lim}_{k \rightarrow \infty} |A_n^{(k)}|^{\frac{1}{(n+k-1)k}} \leq \sigma^{1-\kappa} \tau^\kappa.$$

Zusatz I. Es ist

$$(5) \quad \varrho \geq \sigma \geq \tau.$$

Zusatz II. Es gilt (4) gleichmäßig in folgendem Sinne: Zu den gegebenen positiven Zahlen ε (klein) und ω (groß) gehört ein K , so beschaffen, daß, sobald

$$(6) \quad n < \omega k$$

und $k > K$ ist, notwendigerweise die Ungleichung

$$(7) \quad |A_n^{(k)}|^{\frac{1}{(n+k-1)k}} < (\sigma + \varepsilon)^{\frac{n}{n+k-1}} (\tau + \varepsilon)^{\frac{k-1}{n+k-1}}$$

stattfindet.

Ein Sonderfall von (4), nämlich die Ungleichung

$$(8) \quad \overline{\lim}_{k \rightarrow \infty} |A_0^{(k)}|^{\frac{1}{k(k-1)}} \leq \tau$$

verdient besonders hervorgehoben zu werden. Es stellt sich heraus, daß in vielen Beispielen die Punktmenge \mathfrak{A} der Funktion $f(z)$ so angepaßt werden kann, daß in (8) das Gleichheitszeichen gilt. Durch diese Beispiele wird die Frage nahegelegt, ob die im folgenden für die Fortsetzbarkeit als notwendig zu erkennenden Kriterien auch hinreichend sind.

Die Arbeit zerfällt in drei Teile. Der erste ist im wesentlichen dem Beweis eines etwas komplizierten, in Nr. 3 zu formulierenden Hilfssatzes gewidmet, der zweite dem Beweis des eben ausgesprochenen Hauptsatzes, und der dritte bringt Beispiele und Anwendungen, die insbesondere die zur Ungleichung (8) gemachte Bemerkung rechtfertigen werden.

I. Über Polynome, die in einer gegebenen Punktmenge wenig von Null abweichen.

1. Zusammenstellung bekannter Hilfssätze. Es sollen \mathfrak{A} , \mathfrak{A}^* beschränkte und abgeschlossene Punktfolgen in der z -Ebene bezeichnen, $T_n(z)$ bzw. $T_n^*(z)$ die dazugehörigen Tschebyscheffischen Polynome n -ten Grades, schließlich τ , τ^* die transfinite Durchmesser von \mathfrak{A} bzw. \mathfrak{A}^* .

I. Wenn \mathfrak{A} unendlich viele Punkte enthält, liegen die Nullstellen der zur Punktmenge \mathfrak{A} gehörigen Tschebyscheffischen Polynome $T_1(z)$, $T_2(z)$, ... in der konvexen Hülle von \mathfrak{A} .

II. Wenn \mathfrak{A}^* aus \mathfrak{A} mittels Verschiebung, Drehung und Dehnung im Linearverhältnis $1:c$ entsteht, ist $\tau^* = c\tau$.

III. Wenn \mathfrak{A} eine Teilmenge von \mathfrak{A}^* ist, ist $\tau \leq \tau^*$.

IV. Es sei gegeben die abgeschlossene Menge \mathfrak{A} und die Zahl δ , $\delta > 0$. Es gibt eine abgeschlossene Menge \mathfrak{A}^* mit folgenden Eigenschaften:

a) \mathfrak{A}^* besteht aus einer endlichen Anzahl von punktfremden, einfach zusammenhängenden Bereichen, deren jeder durch eine reguläre geschlossene Kurve begrenzt ist.

b) Jeder Punkt von \mathfrak{A} ist innerer Punkt von \mathfrak{A}^* , und jede Randkurve von \mathfrak{A}^* schließt Punkte von \mathfrak{A} ein.

c) Jeder Randpunkt von \mathfrak{A}^* hat von \mathfrak{A} eine kleinere Entfernung als δ .

d) Es ist $\tau^* < \tau + \delta$.

Unter einer „regulären geschlossenen Kurve“ soll in dieser Arbeit eine geschlossene Kurve ohne Doppelpunkte verstanden werden, die aus einer endlichen Anzahl von Kreisbögen besteht. (Man könnte sich an Stelle von „Kreisbögen“ auch auf „Geradenstücke“ oder auf „analytische Bögen“ festlegen.)

V. Der transfinite Durchmesser einer abgeschlossenen Menge \mathfrak{A} , deren (den Punkt $z = \infty$ enthaltende) Komplementärmenge \mathfrak{G} einfach zusammenhängend ist, ist der Abbildungsradius von \mathfrak{G} , d. h. der Radius desjenigen Kreises, auf dessen Äußere sich \mathfrak{G} bei Festhaltung des Punktes $z = \infty$ und des Linienelementes daselbst schlicht und konform abbilden läßt.

Zur Auffassung des Satzes V ist es vorteilhaft, \mathfrak{A} und \mathfrak{G} auf der Zahlenkugel sich vorzustellen.

Die Hilfssätze II und III folgen unmittelbar aus den Definitionen, IV ergibt sich auch leicht⁷⁾, I ist ein Satz von L. Fejér⁸⁾, V wurde zu-

⁷⁾ Wird das ganzzahlige m genügend groß und das positive ε genügend klein gewählt, so ist $\tau_m^{1/m} + \varepsilon < \tau + \delta$. Es sei $0 < \eta < \delta$ und η kleiner als die kleinste Entfernung zwischen \mathfrak{A} und dem Rand des Gebietes, in dem

$$|T_m(z)| < (\tau_m^{1/m} + \varepsilon)^m.$$

Man kann endlich viele in diesem Gebiet liegende Kreise vom Radius η finden, die \mathfrak{A} voll überdecken, und zwar so, daß der Mittelpunkt jedes solchen Kreises ein Punkt von \mathfrak{A} ist und keine zwei Kreise sich berühren; letzteres wird durch zweckmäßige Wahl von η erzielt. Die Vereinigungsmenge dieser Kreisflächen heiße \mathfrak{A}^{**} . Man betrachte im Komplement von \mathfrak{A}^{**} dasjenige zusammenhängende Teilgebiet, das $z = \infty$ enthält; das Komplement dieses Teilgebietes kann als \mathfrak{A}^* angenommen werden. Vgl. die Bemerkungen bei Fekete a. a. O.⁵⁾, S. 234, Fußnote 4). Vgl. auch Szegő, a. a. O.⁶⁾, S. 207.

⁸⁾ L. Fejér, Über die Lage der Nullstellen von Polynomen, die aus Minimumforderungen gewisser Art entspringen, Math. Annalen 85 (1922), S. 41–48.

erst durch G. Faber ⁵⁾ bewiesen, allerdings mit Einschränkungen betreffend den Rand von \mathcal{G} ; die Einschränkungen wurden nachher durch M. Fekete ⁵⁾ gehoben.

Es sollen zunächst nur I und IV herangezogen werden, und zwar zum Beweis eines neuen Hilfssatzes, dessen etwas schwerfällige Fassung sich in Nr. 3 befindet, und der von einer anderen Seite her durch die nächste Nummer vorbereitet wird.

2. Hilfssatz über die transfiniten Derivierten einer Punktmenge. *Es sei $\mathfrak{A}^{(\alpha)}$ die α -te derivierte Menge der beschränkten abgeschlossenen ebenen Punktmenge \mathfrak{A} ; hierbei bedeutet α eine finite oder transfiniten Ordinalzahl. Es sei \mathfrak{D} eine offene Punktmenge, die $\mathfrak{A}^{(\alpha)}$ enthält. Falls es überhaupt derivierte Mengen gibt, die nicht ganz in \mathfrak{D} enthalten sind, gibt es unter ihnen eine $\mathfrak{A}^{(\alpha_1)}$ von höchstem Index α_1 , und es liegen nur endlich viele Punkte von $\mathfrak{A}^{(\alpha_1)}$ außerhalb \mathfrak{D} . Es sei $\mathfrak{A}^{(\alpha_1)}$ in der offenen Punktmenge \mathfrak{D}_1 enthalten und die Derivierte von höchstem Index, die nicht ganz innerhalb \mathfrak{D}_1 liegt, sei (falls sie existiert) mit $\mathfrak{A}^{(\alpha_2)}$ bezeichnet usw. Es ist $\alpha > \alpha_1 > \alpha_2 > \alpha_3 > \dots$ und das Verfahren bricht nach endlich vielen Schritten ab; d. h. es existiert eine natürliche Zahl l , so beschaffen, daß die offene Punktmenge \mathfrak{D}_l , die $\mathfrak{A}^{(\alpha)}$ enthält, auch \mathfrak{A} enthält.*

Beweis. Da die Ordinalzahlen wohlgeordnet sind, gibt es unter den Derivierten von \mathfrak{A} eine von kleinstem Index β , die ganz in \mathfrak{D} enthalten ist, $\beta \leq \alpha$. Ich behaupte, daß β keine Limeszahl ist. Denn wäre β Grenzwert der unendlichen wachsenden Folge $\beta_1, \beta_2, \dots, \beta_n, \dots$, so hätten, wenn $\bar{\mathfrak{D}}$ die Komplementärmenge von \mathfrak{D} bezeichnet, die ineinandergeschachtelten abgeschlossenen Mengen $\mathfrak{A}^{(\beta_1)}\bar{\mathfrak{D}}, \mathfrak{A}^{(\beta_2)}\bar{\mathfrak{D}}, \dots, \mathfrak{A}^{(\beta_n)}\bar{\mathfrak{D}}, \dots$ sämtlich Punkte, und so wäre auch $\mathfrak{A}^{(\beta)}\bar{\mathfrak{D}}$ nicht leer, was der Voraussetzung widerspricht. — Somit kann man, falls überhaupt $\beta > 0$ ist, $\beta = \alpha_1 + 1$ setzen. Die Menge $\mathfrak{A}^{(\alpha_1)}\bar{\mathfrak{D}}$ ist abgeschlossen und nicht leer; hätte sie unendlich viele Punkte, so wäre auch $\mathfrak{A}^{(\alpha_1+1)}\bar{\mathfrak{D}} = \mathfrak{A}^{(\beta)}\bar{\mathfrak{D}}$ nicht leer, was der Bedeutung von β widersprechen würde; also hat $\mathfrak{A}^{(\alpha_1)}\bar{\mathfrak{D}}$ nur endlich viele Punkte. — Ist \mathfrak{A} in \mathfrak{D}_1 noch nicht enthalten, so bestimmt man ähnlich die Zahl α_2 usw. Die absteigende Folge von Ordinalzahlen $\alpha, \alpha_1, \alpha_2, \alpha_3, \dots$ muß wegen der Wohlordnung nach endlich vielen Gliedern abbrechen. (Das Abbrechen der Folge $\alpha, \alpha_1, \alpha_2, \dots$ ist übrigens auch aus dem Borelschen Überdeckungssatz klar.)

Es sei hervorgehoben, daß Satz und Beweis ihren Sinn behalten, wenn die Menge $\mathfrak{A}^{(\alpha)}$ leer ist, wie auch dann, wenn $\mathfrak{A}^{(\alpha)}$ und \mathfrak{D} zugleich leer sind.

3. Hilfssatz über Polynome, die wenig von Null abweichen. *Es sei \mathfrak{A} eine beschränkte, abgeschlossene, ebene Punktmenge, \mathfrak{P} ihr per-*

fekter Kern, d. h. die Gesamtheit der Verdichtungspunkte von \mathfrak{A} , und es sei τ jetzt als der transfinite Durchmesser von \mathfrak{B} (nicht als der von \mathfrak{A}) definiert. Es sei ferner gegeben die Zahl ε , $\varepsilon > 0$. Dann kann man endlich viele reguläre geschlossene Kurven $\mathfrak{L}_1, \mathfrak{L}_2, \dots, \mathfrak{L}_m$, die zu je zwei punktfremd sind, und eine unendliche Folge von Polynomen $Q_0(z), Q_1(z), Q_2(z), \dots, Q_n(z), \dots$ folgenden Bedingungen gemäß bestimmen:

a) Jeder Punkt von \mathfrak{A} liegt im Innern einer Kurve \mathfrak{L}_μ ($1 \leq \mu \leq m$).

b) Jeder Punkt von \mathfrak{L}_μ hat von \mathfrak{A} eine kleinere Entfernung als ε ($\mu = 1, 2, \dots, m$). Falls \mathfrak{L}_μ Punkte von \mathfrak{B} umschließt, so hat jeder Punkt von \mathfrak{L}_μ schon von \mathfrak{B} eine kleinere Entfernung als ε .

c) $Q_0(z) = 1$, $Q_n(z) = z^n + \dots$ ist vom n -ten Grad und hat 1 zum höchsten Koeffizienten ($n = 0, 1, 2, 3, \dots$).

d) Entlang solcher Kurven \mathfrak{L}_μ , die Punkte von \mathfrak{B} umschließen, bleibt die Folge

$Q_0(z), Q_1(z)(\tau + \varepsilon)^{-1}, Q_2(z)(\tau + \varepsilon)^{-2}, \dots, Q_n(z)(\tau + \varepsilon)^{-n}, \dots$ beschränkt, entlang solcher Kurven \mathfrak{L}_μ , die keinen Punkt von \mathfrak{B} umschließen, bleibt sogar die Folge

$$Q_0(z), Q_1(z)\varepsilon^{-1}, Q_2(z)\varepsilon^{-2}, \dots, Q_n(z)\varepsilon^{-n}, \dots$$

beschränkt.

Beweis. Die Menge \mathfrak{B} der Verdichtungspunkte von \mathfrak{A} kann bekanntlich als die α -te Derivierte von \mathfrak{A} aufgefaßt werden, wobei α eine finite oder transfinite Ordinalzahl bedeutet. Ich werde die unter Nr. 2 beschriebene Konstruktion auf $\mathfrak{A}^{(\alpha)} = \mathfrak{B}$ anwenden.

Es treten im folgenden gewisse positive Zahlen $\delta, \delta_1, \delta_2, \dots$ auf, die sämtlich kleiner sind als ε ; die weitere Fixierung ihrer Werte sei noch vorbehalten.

Man wende den Hilfssatz IV auf $\mathfrak{A}^{(\alpha)}$ (anstatt auf \mathfrak{A}) an⁹⁾. Es heiße auch jetzt \mathfrak{A}^* die durch reguläre geschlossene Kurven begrenzte Punktmenge vom transfiniten Durchmesser τ^* , die in bezug auf $\mathfrak{A}^{(\alpha)}$ den Bedingungen a), b), c), d) von Hilfssatz IV genügt. Die inneren Punkte von \mathfrak{A}^* bilden die offene Punktmenge \mathfrak{D} , die im Sinne der Konstruktion unter Nr. 2 zu verwenden ist. $\mathfrak{A}^{(\alpha)}$ sei, wie in Nr. 2, die Derivierte von höchstem Index die durch \mathfrak{D} nicht völlig überdeckt ist; $\mathfrak{A}^{(\alpha)}$ soll k_1 Punkte außerhalb \mathfrak{D} besitzen. Man konstruiere k_1 Kreise, deren Zentra diese k_1 Punkte sind, und zwar folgendermaßen:

⁹⁾ Im Text wird der Fall vor Augen behalten, in dem $\mathfrak{B} = \mathfrak{A}^{(\alpha)}$ nicht leer ist. Wenn \mathfrak{B} leer ist, kann die Konstruktion des Textes vereinfacht, oder aber in folgender Weise uminterpretiert werden: \mathfrak{A}^* ist ein willkürlich gewählter Punkt a von \mathfrak{A} , \mathfrak{D} ist leer, $\tau = 0$, $T_n^*(z) = (z - a)^n$.

1. Der Radius jeden Kreises ist $< \delta_1$.

2. Die offenen Flächen dieser Kreise bilden zusammen mit \mathfrak{D} ein offenes Gebiet \mathfrak{D}_1 , das das Innere von endlich vielen, zueinander punktfremden regulären geschlossenen Kurven ausmacht.

3. Die Punkte derjenigen der eben erwähnten Kurven, die Punkte von \mathfrak{B} umschließen, sind von \mathfrak{B} in kleinerer Entfernung als δ .

Dies alles läßt sich durch die Wahl genügend kleiner Kreise erreichen, wie man bei genauer Erwägung aller Möglichkeiten (z. B. daß einige der k_1 Punkte auf den Rand von \mathfrak{D} fallen können) findet.

Wie \mathfrak{D}_1 aus \mathfrak{D} , so sei \mathfrak{D}_2 aus \mathfrak{D}_1 konstruiert: Um die k_2 Punkte von $\mathfrak{A}^{(\alpha_2)}$, die außerhalb \mathfrak{D}_1 geblieben sind, werden k_2 Kreise vom Radius $< \delta_2$ geschlagen, deren inneren Punkte mit \mathfrak{D}_1 zusammen \mathfrak{D}_2 bilden, das die sinngemäß modifizierten Bedingungen 2, 3 (\mathfrak{D} ist jetzt in \mathfrak{D}_1 und \mathfrak{D}_1 in \mathfrak{D}_2 zu verwandeln) erfüllt. Ähnlich konstruiert man $\mathfrak{D}_3, \mathfrak{D}_4, \dots$ durch Hinzufügen von k_3, k_4, \dots Kreisen, deren Radien bzw. $< \delta_3, \delta_4, \dots$ sind, unter Beobachtung der sinngemäß modifizierten Bedingungen 2, 3. So gelangt man durch sukzessive Hinzufügung von offenen Kreisflächen schließlich zur Punktmenge \mathfrak{D}_l , die \mathfrak{A} als Teilmenge enthält, und das Innengebiet von endlich vielen, regulären, geschlossenen, zueinander punktfremden Kurven $\mathfrak{Q}_1, \mathfrak{Q}_2, \dots, \mathfrak{Q}_m$ ausmacht; diese erfüllen, gemäß Konstruktion, die geometrischen Bedingungen a), b) des zu beweisenden Satzes; die letztere ist wegen 3. und $\delta < \varepsilon, \delta_1 < \varepsilon, \delta_2 < \varepsilon, \dots$ erfüllt.

$T_0^*(z), T_1^*(z), \dots, T_n^*(z), \dots$ seien die zum Gebiet \mathfrak{A}^* gehörigen Tschebyscheffschen Polynome ($T_0^*(z) = 1, T_n^*(z) = z^n + \dots$). Am Rande von \mathfrak{D} (oder von \mathfrak{A}^* , was dasselbe ist) gilt, nach Konstruktion ($\tau^* < \tau + \delta$) und gemäß der Definition der Tschebyscheffschen Polynome und der Zahl τ^*

$$(9) \quad T_n^*(z) = O((\tau + \delta)^n) \quad \text{für } n \rightarrow \infty.$$

$P_l(z) = z^{k_l} + \dots$ sei dasjenige Polynom mit höchstem Koeffizienten 1, dessen Nullstellen die k_l Punkte von $\mathfrak{A}^{(\alpha_l)}$ sind, die außerhalb \mathfrak{D}_{l-1} liegen und in der Konstruktion als Mittelpunkte der Kreise dienen, durch deren Hinzufügen \mathfrak{D}_l aus \mathfrak{D}_{l-1} entstanden ist (\mathfrak{D}_0 bedeutet \mathfrak{D}). Daher ist in solchen Randpunkten von \mathfrak{D}_l , die nicht Randpunkte von \mathfrak{D}_{l-1} sind, einer der k_l Linearfaktoren von $P_l(z)$ dem Betrage nach $< \delta_l$.

Es sei $0 < \gamma < \frac{1}{2}$; auch der Wert von γ wird erst später genauer fixiert. Es sei

$$(10) \quad p = n - k_1 \left[\frac{\gamma n}{k_1} \right] - k_2 \left[\frac{\gamma^2 n}{k_2} \right] - \dots - k_l \left[\frac{\gamma^l n}{k_l} \right]$$

($[x]$ heißt ganzer Teil von x); es ist

$$p \geq n(1 - \gamma - \gamma^2 - \dots - \gamma^l) > 0.$$

Ferner sei

$$(11) \quad Q_n(z) = T_p^*(z) P_1(z)^{[\gamma n k_1^{-1}]} P_2(z)^{[\gamma^2 n k_2^{-1}]} \dots P_l(z)^{[\gamma^l n k_l^{-1}]}$$

gesetzt. $Q_n(z)$ erfüllt offenbar die Bedingung c). Alles kommt darauf hinaus, zu zeigen, daß durch passende Wahl von $\gamma, \delta, \delta_1, \delta_2, \dots$ die Bedingung d) erfüllbar ist.

4. Fortsetzung des Beweises. Es sei \mathfrak{R} der kleinste abgeschlossene konvexe Bereich, der alle Punkte umfaßt, die von \mathfrak{A} eine kleinere Entfernung als ε besitzen. Alle Kurven $\mathfrak{Q}_1, \mathfrak{Q}_2, \dots, \mathfrak{Q}_m$ und alle Nullstellen von $Q_n(z)$ liegen in \mathfrak{R} ; die Nullstellen von $P_\lambda(z)$ sind Punkte von \mathfrak{A} , für die von $T_p^*(z)$ vgl. I unter Nr. 1. Bedeutet z einen auf \mathfrak{Q}_μ liegenden Punkt, so stellt jeder Linearfaktor von $Q_n(z)$ den Verbindungsvektor zweier in \mathfrak{R} gelegener Punkte dar.

Es sei d der Maximalabstand zweier in \mathfrak{R} gelegener Punkte, der sog. Durchmesser von \mathfrak{R} ; es ist $d \geq 2\varepsilon$.

Die Zahl δ sei bloß der (schon ausgesprochenen) Bedingung

$$(12) \quad 0 < \delta < \varepsilon$$

unterworfen. Es sei die positive Zahl A so gewählt, daß

$$(13) \quad d < A,$$

$$(14) \quad \tau + \delta < A.$$

Es sei γ so gewählt, daß

$$(15) \quad (\tau + \delta) \left(\frac{A}{\tau + \delta} \right)^{\frac{\gamma}{1-\gamma}} < \tau + \varepsilon.$$

Schließlich sei δ_λ der Ungleichung

$$(16) \quad 0 < \delta_\lambda < A \left(\frac{\varepsilon}{A} \right)^{k_\lambda \gamma^{-\lambda}}$$

unterworfen. k_λ , also auch δ_λ ergibt sich erst nach der Wahl von $\delta_{\lambda-1}$ ($\delta_0 = \delta$) und l ergibt sich nach der Wahl von $\delta, A, \gamma, \delta_1, \delta_2, \delta_3, \dots$

Auf dem Rand von \mathfrak{D} (= Rand von \mathfrak{A}^*) ist, wie man durch sukzessive Heranziehung von (11), (13), (9), (10), (14) und (15) findet,

$$\begin{aligned} |Q_n(z)| &\leq |T_p^*(z)| A^{n-p} = O\left((\tau + \delta)^n \left(\frac{A}{\tau + \delta}\right)^{n-p}\right) \\ &= O\left((\tau + \delta)^n \left(\frac{A}{\tau + \delta}\right)^{\frac{\gamma n}{1-\gamma}}\right) = O((\tau + \varepsilon)^n). \end{aligned}$$

Auf demjenigen Stück des Randes von \mathfrak{D}_λ , das nicht Rand von $\mathfrak{D}_{\lambda-1}$ ist ($\lambda = 1, 2, \dots, l$), hat ein gewisser Linearfaktor von $P_\lambda(z)$ einen Betrag $< \delta_\lambda$, gemäß der Konstruktion von $P_\lambda(z)$, und daher ergibt es sich

da sogar, mit Beachtung von (11), (13) und (16),

$$|Q_n(z)| \leq \delta_\lambda^{[\gamma^2 n k_\lambda^{-1}]} A^{n - [\gamma^2 n k_\lambda^{-1}]} < \left(\frac{\delta_\lambda}{A}\right)^{\gamma^2 n k_\lambda^{-1} - 1} A^n = O(\varepsilon^n).$$

Nun ist irgendein Punkt von \mathcal{Q}_μ entweder Randpunkt von \mathfrak{D} oder von \mathfrak{D}_1 oder von $\mathfrak{D}_2 \dots$ oder schließlich von \mathfrak{D}_i . Beachtet man, daß solche Kurven \mathcal{Q}_μ , die keine Punkte von \mathfrak{B} einschließen, nur aus Randpunkten von $\mathfrak{D}_1, \mathfrak{D}_2, \dots, \mathfrak{D}_i$ bestehen, so findet man den am Anfang der Nr. 3 ausgesprochenen Hilfssatz voll bestätigt.

5. Korollar. Da $Q_n(z)$ auf den Kurven \mathcal{Q}_μ (und so auch im Innern der Kurven \mathcal{Q}_μ) also in allen Punkten von \mathfrak{A} gleichmäßig $O((\tau + \varepsilon)^n)$ ist, ist der transfinite Durchmesser von \mathfrak{A} (nach Definition, vgl. die Einleitung) nicht größer als $\tau + \varepsilon$, also $\leq \tau$, da ε beliebig ist. Er ist auch $\geq \tau$, da die abgeschlossene Menge \mathfrak{A} die Menge ihrer Verdichtungspunkte \mathfrak{B} , als deren transfiniter Durchmesser τ definiert wurde, enthält (vgl. III unter Nr. 1). Er ist somit $= \tau$. So haben wir folgenden Satz gefunden: *Eine beschränkte abgeschlossene Punktmenge und ihr perfekter Kern haben denselben transfiniten Durchmesser¹⁰⁾*. Dies Ergebnis kann man, auf Grund von III unter Nr. 1, auch so aussprechen: *Zwei beschränkte und abgeschlossene Punkt Mengen, die sich nur in abzählbar vielen Punkten unterscheiden, besitzen denselben transfiniten Durchmesser.*

Der perfekte Kern der in dem Hauptsatz erwähnten Punktmenge \mathfrak{A} hat, nach dem eben Gesagten, denselben transfiniten Durchmesser wie \mathfrak{A} , nämlich τ . Er ist in der Kreisfläche $|z| \leq \sigma$ enthalten, dessen transfiniter Durchmesser, nach Hilfssatz V, gleich ihrem Radius σ ist. Somit folgt aus Hilfsatz III, daß $\tau \leq \sigma$ ist, und hierin besteht der nichttriviale Teil des Zusatzes I zu dem Hauptsatz.

II. Beweis des Hauptsatzes.

6. Der Beweis steckt sich als Zielpunkt den verschärfenden Zusatz II und macht vollen Gebrauch von dem in Nr. 3 formulierten Hilfssatz.

a) Die Funktion $f(z)$ ist, nach Voraussetzung, in dem m -fach zusammenhängenden Außenraum der Kurven $\mathcal{Q}_1, \mathcal{Q}_2, \dots, \mathcal{Q}_m$ regulär und eindeutig. Daher ist der n -te Koeffizient der Reihe (1)

$$a_n = \frac{1}{2\pi i} \oint_{\mathfrak{L}} f(z) z^n dz,$$

wo der Integrationsweg \mathfrak{L} die Gesamtheit der Kurven $\mathcal{Q}_1, \mathcal{Q}_2, \dots, \mathcal{Q}_m$,

¹⁰⁾ Ein Spezialfall hiervon ist der bei Szegő a. a. O. ⁵⁾, S. 208, Fußnote ⁵⁾ angegebene Satz.

alle im positiven Sinne durchlaufen, bedeutet. Hieraus erhält man durch geeignete Kombination der Zeilen und der Kolonnen¹¹⁾ folgenden Ausdruck der Determinante (2):

$$(17) \quad A_n^{(k)} = \left| \frac{1}{2\pi i} \oint_{\mathfrak{Z}} f(z) z^n Q_\lambda(z) Q_\mu(z) \right|_{\lambda, \mu = 0, 1, \dots, k-1}.$$

$Q_n(z)$ bedeutet das aus dem Hilfssatz der Nr. 3 bekannte Polynom, $n = 0, 1, 2, \dots$, und die Umformung stützt sich wesentlich auf die dort unter c) genannte Eigenschaft.

b) Wenn z die m Kurven $\mathfrak{L}_1, \mathfrak{L}_2, \dots, \mathfrak{L}_m$ durchläuft, erreicht $|f(z)|$ ein bestimmtes Maximum M .

c) Diejenigen Kurven \mathfrak{L}_μ , die Punkte von \mathfrak{Z} umschließen, seien generell mit \mathfrak{L}_α , diejenigen, die Punkte von \mathfrak{Z} nicht umschließen, mit \mathfrak{L}_β bezeichnet. Die Kurven \mathfrak{L}_α liegen alle in der Kreisscheibe $|z| < \sigma + \varepsilon$ und die Kurven \mathfrak{L}_β in der Kreisscheibe $|z| < \varrho + \varepsilon$. Vgl. den Hilfssatz in Nr. 3, unter b).

d) Es gibt eine Konstante C , so beschaffen, daß

$$\begin{aligned} |Q_p(z)| &< C(\tau + \varepsilon)^p \quad \text{entlang } \mathfrak{L}_\alpha, \\ |Q_p(z)| &< C\varepsilon^p \quad \text{entlang } \mathfrak{L}_\beta \\ &(p = 0, 1, 2, \dots). \end{aligned}$$

Es sei A die Gesamtlänge der Kurven \mathfrak{L}_α und B die der Kurven \mathfrak{L}_β . Auf Grund von b), c), d) ergibt sich für das allgemeine Element der Determinante (17) die Abschätzung

$$(18) \quad \left| \frac{1}{2\pi i} \oint_{\mathfrak{Z}} f(z) z^n Q_\lambda(z) Q_\mu(z) dz \right| < \frac{A}{2\pi} M(\sigma + \varepsilon)^n C^2 (\tau + \varepsilon)^{\lambda + \mu} + \frac{B}{2\pi} M(\varrho + \varepsilon)^n C^2 \varepsilon^{\lambda + \mu} = A(\sigma + \varepsilon)^n (\tau + \varepsilon)^{\lambda + \mu} + B(\varrho + \varepsilon)^n \varepsilon^{\lambda + \mu},$$

wo A, B von n, λ, μ unabhängig sind.

Im folgenden Beweis soll der schwierigere Fall, in dem $\tau > 0$ ist, ins Auge gefaßt werden. Es sei gegeben eine (kleine) positive Zahl η . Wir denken uns ein positives ε von vornherein so klein gewählt, daß

$$(19) \quad \left(\frac{\varrho}{\sigma}\right)^\omega \left(\frac{\varepsilon}{\tau + \varepsilon}\right)^\eta < 1$$

(ω kommt unter (6) vor); dies ist möglich, falls $\tau > 0$ ist. Wir haben die Elemente in den ersten $[k\eta]$ Kolonnen der Determinante (17) anders abzuschätzen, als die in den letzten $k - [k\eta]$.

¹¹⁾ Vgl. a. a. O. ⁴⁰⁾ S. 25–26.

1. $\mu < [k\eta]$. Die letzte Zeile von (18) ist

$$< (A + B)(\varrho + \varepsilon)^n (\tau + \varepsilon)^{\lambda + \mu}.$$

2. $\mu \geq [k\eta]$. In diesem Falle ist die letzte Zeile von (18)

$$\begin{aligned} &= \left(A + B \left(\frac{\varrho + \varepsilon}{\sigma + \varepsilon} \right)^n \left(\frac{\varepsilon}{\tau + \varepsilon} \right)^\lambda \left(\frac{\varepsilon}{\tau + \varepsilon} \right)^\mu \right) (\sigma + \varepsilon)^n (\tau + \varepsilon)^{\lambda + \mu} \\ &\leq \left(A + B \left(\frac{\varrho + \varepsilon}{\sigma + \varepsilon} \right)^{k\omega} \left(\frac{\varepsilon}{\tau + \varepsilon} \right)^{[k\eta]} \right) (\sigma + \varepsilon)^n (\tau + \varepsilon)^{\lambda + \mu} \\ &< (A + B)(\sigma + \varepsilon)^n (\tau + \varepsilon)^{\lambda + \mu} \end{aligned}$$

für genügend großes k , mit Rücksicht auf (6), (5) und (19).

Wir heben aus der $(\lambda + 1)$ -ten Zeile der Determinante (17) die Potenz $(\tau + \varepsilon)^\lambda$ hervor und aus der $(\mu + 1)$ -ten Kolonne $(A + B)(\varrho + \varepsilon)^n (\tau + \varepsilon)^\mu$ bzw. $(A + B)(\sigma + \varepsilon)^n (\tau + \varepsilon)^\mu$, je nachdem $\mu < [k\eta]$ oder $\mu \geq [k\eta]$ ist. Wir erhalten so für genügend großes k , daß

$$\begin{aligned} |A_n^{(k)}| &< (\tau + \varepsilon)^{k(k-1)} (A + B)^k (\varrho + \varepsilon)^{n[k\eta]} (\sigma + \varepsilon)^{n(k-[k\eta])} k!, \\ |A_n^{(k)}|^{\frac{1}{(n+k-1)k}} &< ((A + B)k)^{\frac{1}{n+k-1}} \left(\frac{\varrho + \varepsilon}{\sigma + \varepsilon} \right)^\eta (\sigma + \varepsilon)^{\frac{n}{n+k-1}} (\tau + \varepsilon)^{\frac{k-1}{n+k-1}}. \end{aligned}$$

Mit Rücksicht darauf, daß η und ε beliebig klein gewählt werden können, ist die letzte Aussage mit (7) gleichbedeutend.

III. Beispiele und Anwendungen.

7. Beispiel. Wir betrachten die Funktion

$$(20) \quad \log \frac{z}{z-1} = \frac{1}{z} + \frac{1}{2z^2} + \frac{1}{3z^3} + \dots$$

Für die Funktion (20) ist

$$\alpha_n = \frac{1}{n+1},$$

$$(21) \quad A_n^{(k)} = \begin{vmatrix} \frac{1}{n+1} & \frac{1}{n+2} & \cdots & \frac{1}{n+k} \\ \frac{1}{n+2} & \frac{1}{n+3} & \cdots & \frac{1}{n+k+1} \\ \cdot & \cdot & \cdot & \cdot \\ \frac{1}{n+k} & \frac{1}{n+k+1} & \cdots & \frac{1}{n+2k-1} \end{vmatrix};$$

es soll bewiesen werden, daß

$$(22) \quad \lim_{k \rightarrow \infty} |A_0^{(k)}|^{\frac{1}{(k-1)k}} = \frac{1}{4},$$

ferner daß

$$(23) \quad \lim_{k \rightarrow \infty} |A_n^{(k)}|^{\frac{1}{(n+k-1)k}} = \left(\frac{x^{2x^2}}{(1-x)^{(1-x)^2} (1+x)^{(1+x)^2}} \right)^{\frac{1}{2x}} < \frac{1}{4^x},$$

falls n so mit k verbunden ist, daß

$$(24) \quad \lim_{k \rightarrow \infty} \frac{k-1}{n+k-1} = x, \quad (0 < x < 1).$$

Zuerst soll die Gleichung unter (23) bewiesen werden. Man findet auf Grund einer bekannten Determinantenformel¹²⁾, daß

$$\begin{aligned} A_n^{(k)} &= \frac{[1^{k-1} 2^{k-2} 3^{k-3} \dots (k-1)^1]^2}{(n+1)^1 (n+2)^2 \dots (n+k)^k \dots (n+2k-2)^2 (n+2k-1)^1} \\ &= \frac{1}{k^k} \left(\frac{k}{n+k} \right)^{k^2} \frac{\left[\left(\frac{1}{k} \right)^{k-1} \left(\frac{2}{k} \right)^{k-2} \dots \left(\frac{k-1}{k} \right)^1 \right]^2}{\left[1 - \left(\frac{1}{n+k} \right)^2 \right]^{k-1} \left[1 - \left(\frac{2}{n+k} \right)^2 \right]^{k-2} \dots \left[1 - \left(\frac{k-1}{n+k} \right)^2 \right]^1}, \\ & \frac{\log A_n^{(k)}}{(n+k-1)k} = \frac{1}{n+k-1} \log \frac{1}{k} + \frac{k}{n+k-1} \log \frac{k}{n+k} \\ & + \frac{2k}{n+k-1} \sum_{\nu=1}^{k-1} \left(1 - \frac{\nu}{k} \right) \log \frac{\nu}{k} \cdot \frac{1}{k} - \frac{(n+k)^2}{(n+k+1)k} \sum_{\nu=1}^{k-1} \frac{k-\nu}{n+k} \log \left[1 - \left(\frac{\nu}{n+k} \right)^2 \right] \cdot \frac{1}{n+k} \\ & \rightarrow 0 + x \log x + 2x \int_0^1 (1-x) \log x \cdot dx - \frac{1}{x} \int_0^x (x-x) \log(1-x^2) \cdot dx \end{aligned}$$

unter der Bedingung (24). Die letzterhaltene Beziehung ist mit der Gleichung unter (23) gleichbedeutend. Die Gleichung (nicht die Ungleichung!) unter (23) bleibt, wie die Ableitung zeigt, auch für $x=1$ gültig und ergibt in diesem Grenzfall (22).

Nun soll die Ungleichung unter (23) bewiesen werden. Sie ist, wenn

$$(25) \quad \varphi(x) = 2x^2 \log 4x - (1-x)^2 \log(1-x) - (1+x)^2 \log(1+x)$$

gesetzt wird, mit

$$(26) \quad \varphi(x) < 0 \quad \text{für } 0 < x < 1$$

gleichbedeutend. Man findet (alle drei nun folgenden Werte sind durch die Forderung der Stetigkeit definiert), daß

$$(27) \quad \varphi(0) = 0, \quad \varphi(1) = 0,$$

$$(28) \quad \varphi'(0) = 0.$$

¹²⁾ Vgl. z. B. a. a. O. ²⁾ Bd. II, Aufgaben 3, 4 auf S. 98 und 299–300.

Gäbe es einen Wert $x = \kappa$, $0 < \kappa < 1$, so beschaffen, daß $\varphi(\kappa) = 0$ ist, so würde $\varphi'(x)$, gemäß (27), in mindestens zwei inneren Punkten des Intervalles $0, 1$ verschwinden, und folglich würde, gemäß (28), auch $\varphi''(x)$ mindestens zwei verschiedene Nullstellen im Innern desselben Intervalles besitzen. Nun hat aber

$$\varphi''(x) = 2 \log \frac{16x^2}{1-x^2}$$

nur die einzige Nullstelle $x = \frac{1}{\sqrt{17}}$ zwischen 0 und 1 . Also verschwindet die Funktion $\varphi(x)$ in keinem innern Punkt des Intervalles $0, 1$. Sie ist daselbst von konstantem, und zwar von negativem Vorzeichen, wie die Berechnung eines Wertes, z. B. von $\varphi(\frac{1}{2})$ zeigt.

8. Bemerkungen zum vorangehenden Beispiel. Es sei nun unter \mathfrak{G} irgendein einfach zusammenhängendes Gebiet in der z -Ebene (auf der z -Kugel) verstanden, das den Punkt $z = \infty$ enthält, und die Punkte $z = 0$ und $z = 1$ nicht enthält. Die Funktion (20) ist in \mathfrak{G} regulär und eindeutig. Als Gebiet hat \mathfrak{G} nur innere Punkte. Die Komplementärmenge von \mathfrak{G} ist beschränkt und perfekt; sie sei \mathfrak{A} , ihr größter Abstand vom Nullpunkt σ und ihr transfiniter Durchmesser τ genannt. Es ist offenbar $\sigma \geq 1$ und τ ist, nach dem unter Nr. 1 zitierten Satz V, der Abbildungsradius von \mathfrak{G} .

Da für die jetzt betrachtete Funktion (20), wie eben berechnet, die Gleichung (22) gilt, ergibt der Fall (8) des Hauptsatzes, daß

$$\frac{1}{4} \leq \tau.$$

Wir haben so einen bekannten wichtigen Satz über konforme Abbildung (die zuerst von L. Bieberbach erbrachte Wertbestimmung der Koebschen Verzerrungskonstanten) wiedergefunden¹³⁾: *Der Abbildungsradius eines einfach zusammenhängenden Gebietes, das den Punkt ∞ enthält und die Punkte 0 und 1 nicht enthält, ist mindestens $\frac{1}{4}$.*

Wie bekannt, ist $\frac{1}{4}$ hier die beste Grenze; die Funktion

$$z = \frac{1}{2} + \frac{1}{4} \left(4w + \frac{1}{4w} \right)$$

bildet das Kreisäußere $|w| > \frac{1}{4}$ schlicht auf die von 0 bis 1 geradlinig

¹³⁾ Vgl. z. B. a. a. O. ²⁾, Bd. II, Aufgaben 141, 142 auf S. 25 und S. 199. Auf Grund der Resultate von ³⁰⁾ wurde ein äquivalenter Satz schon vorher durch G. Szegő auf ähnliche Art bewiesen. Vgl. a. a. O. ^{4a)}, S. 110, Fußnote ²⁾. Für einen auf die (auch hier benutzte) Größenordnung der Tschebyscheffschen Polynome in direkter Weise zurückgreifenden Beweis vgl. G. Faber, Über Potentialtheorie und konforme Abbildung, Sitzungsber. d. Bayer. Akad. (1920), S. 49–64, insbes. S. 56–58.

aufgeschlitzte z -Ebene ab. Wird also in dem Hauptsatz unter $f(z)$ die Funktion (20) und unter \mathfrak{A} die abgeschlossene geradlinige Verbindungsstrecke der Punkte 0 und 1 verstanden, so gilt in (8) das Gleichheitszeichen. Hingegen kann für die Funktion (20) *das Gleichheitszeichen in (4) bei keiner Wahl von \mathfrak{A} gelten, wenn $0 < \kappa < 1$ ist.* Denn die linke Seite von (4) ist im vorliegenden Falle $< 4^{-\kappa}$, gemäß (23). Nun muß \mathfrak{A} ein die Punkte $z=0$ und $z=1$ verbindendes Kontinuum enthalten, sonst ist (20) in der Komplementärmenge von \mathfrak{A} nicht eindeutig. Also ist $\sigma \geq 1$ und, nach dem eben Gesagten, $\tau \geq \frac{1}{4}$, somit die rechte Seite von (4)

$$\sigma^{1-\kappa} \tau^{\kappa} \geq 1^{1-\kappa} \left(\frac{1}{4}\right)^{\kappa} = 4^{-\kappa}.$$

Schließlich sei noch bemerkt, daß das Beispiel (20) in dem folgenden allgemeineren Beispiel enthalten ist:

$$(29) \quad f(z) = \int_0^1 \frac{\chi(u) du}{z-u} = \frac{a_0}{z} + \frac{a_1}{z^2} + \frac{a_2}{z^3} + \dots$$

$\chi(u)$ bedeutet hierbei eine wesentlich positive, stetige Funktion der reellen Variablen u , $0 \leq u \leq 1$. (Um (20) aus (29) zu erhalten, ist $\chi(u) = 1$ zu setzen.) Gilt für $\chi(u)$ die Ungleichung

$$0 < m \leq \chi(u) \leq M \quad \text{für} \quad 0 \leq u \leq 1,$$

so steht die aus den Koeffizienten

$$a_m = \int_0^1 \chi(u) u^m du$$

gebildete Determinante $A_n^{(k)}$ zu der Determinante (21) in einem leicht abzuschätzenden Verhältnis. Es ist nämlich, wie sich leicht aus einem bekannten Determinantensatz ergibt¹⁴⁾,

$$m^k \left| \frac{1}{n+1+\lambda+\mu} \right|_{\lambda, \mu=0}^{k-1} \leq A_n^{(k)} \leq M^k \left| \frac{1}{n+1+\lambda+\mu} \right|_{\lambda, \mu=0}^{k-1}.$$

Daher bleiben (22) und (23) auch für die allgemeine Funktion (29) gültig.

¹⁴⁾ Für den Determinantensatz vgl. z. B. a. a. O. ²⁾, Bd. I, Aufgabe 68, S. 48 und S. 208. Eine allgemeinere Ungleichung findet sich bei G. Szegő, A Hankel-féle formákról, Math. és természettudományi értesítő **36** (1918), S. 497–538, vgl. S. 514. In dieser Arbeit ist übrigens der asymptotische Wert $A_0^{(k)}$ für die Funktion (29) genauer bestimmt als hier durch (22) geschehen ist. Für die spezielle Funktion (20) findet sich übrigens eine (22) enthaltende genauere Abschätzung schon bei D. Hilbert, Ein Beitrag zur Theorie des Legendreschen Polynoms, Acta Math. **18** (1894), S. 155–159.

9. Fortsetzbare Potenzreihen. Wenn die Potenzreihe (1) über ihren Konvergenzkreis $|z| = \varrho$ hinaus fortsetzbar ist, kann die in dem Hauptsatz erwähnte Punktmenge \mathfrak{A} so gewählt werden, daß ihre Komplementärmenge *einfach zusammenhängend* ist und das Kreisäußere $|z| > \varrho$ als *echtes* Teilgebiet umfaßt. In diesem Falle ist der transfinite Durchmesser τ von \mathfrak{A} , gemäß V unter Nr. 1, der Abbildungsradius der Komplementärmenge von \mathfrak{A} und als solcher, nach einem bekannten fundamentalen Satz über die konforme Abbildung¹⁵⁾, kleiner als der Abbildungsradius des Teilgebietes $|z| > \varrho$, also $\tau < \varrho$. Zusatz II zu dem Hauptsatz ergibt also:

Notwendig für die Fortsetzbarkeit der Reihe (1) vom Konvergenzradius ϱ ist die Bedingung, daß

$$(30) \quad \overline{\lim}_{k \rightarrow \infty} |A_n^{(k)}|^{\frac{1}{(n+k-1)k}} < \varrho$$

ist, wenn n so mit k verbunden ist, daß das Verhältnis n/k beschränkt bleibt.

Mit \leq an Stelle $<$ wäre (30) nichtssagend; es gilt \leq auf alle Fälle, d. h. auch für nichtfortsetzbare Reihen, wie es leicht festzustellen. Ein Teil der notwendigen Bedingung (30) (nämlich $n = 0$) läßt sich so formulieren: Die Reihe

$$a_0 + a_1 z + a_2 z^2 + \dots + a_n z^n + \dots$$

ist nur dann fortsetzbar, wenn ihr Konvergenzbereich ganz im Innern des Konvergenzbereiches der Reihe

$$A_0^{(1)} z + A_0^{(2)} z^4 + \dots + A_0^{(k)} z^{k^2} + \dots$$

enthalten ist.

10. Potenzreihen mit Hadamardschen Lücken. Man sagt, daß die Reihe

$$(31) \quad \frac{a_0}{z} + \frac{a_1}{z^2} + \dots + \frac{a_p}{z^{p+1}} + \frac{a_q}{z^{q+1}} + \dots = f(z),$$

worin $a_{p+1} = a_{p+2} = \dots = a_{q-2} = a_{q-1} = 0$ ist, zwischen a_p und a_q eine Lücke von der Länge $q - p - 1$ aufweist. Gibt es eine unendliche Folge von Koeffizientenpaaren a_p, a_q mit so großen Lücken dazwischen, daß

$$(32) \quad \underline{\lim} \frac{q}{p} > 1$$

ist, wenn p und q die zusammengehörigen Indizespaare der Folge durchlaufen (p die Indizes der vorderen, q die der hinteren Glieder), so soll gesagt werden, daß die Reihe (31) „Hadamardsche Lücken“ aufweist.

¹⁵⁾ Vgl. z. B. a. a. O. ²⁾, Aufgabe 121, S. 21 und S. 195.

Die Existenz der Lücke zwischen a_p und a_q in der Reihe (31) hat zur Folge, daß gewisse Determinanten (2) verschwinden und andere sich auf das Produkt der Glieder in der Nebendiagonale reduzieren. Auf diese Art findet man insbesondere

$$(33) \quad A_{p+1}^{(q-p)} = \pm a_q^{q-p},$$

$$(34) \quad A_{2p-q+1}^{(q-p)} = \pm a_p^{q-p}, \quad \text{falls } q \leq 2p + 1,$$

$$(35) \quad A_0^{(p+1)} = \pm a_p^{p+1}, \quad \text{falls } q \geq 2p + 1.$$

Wenn p die Indizes der vorderen Glieder der Koeffizientenpaare a_p, a_q durchläuft, die durch (32) ausgezeichnet sind, wachsen p, q und $q - p$ gleichzeitig ins Unendliche, und die Quotienten

$$\frac{p}{q-p} = \frac{1}{\frac{q}{p}-1}, \quad \frac{2p-q}{q-p} = -1 + \frac{1}{\frac{q}{p}-1}, \quad \frac{0}{p}$$

bleiben beschränkt. Die zur Anwendung des Zusatzes II nötige Voraussetzung ist also erfüllt, und wir finden, daß bei gegebenem $\varepsilon, \varepsilon > 0$, und für genügend großes p die Ungleichungen

$$(36) \quad |A_{p+1}^{(q-p)}|^{\frac{1}{q(q-p)}} = |a_q|^{\frac{1}{p}} < (\sigma + \varepsilon)^{\frac{p}{q}} (\tau + \varepsilon)^{1-\frac{p}{q}},$$

$$(37) \quad A_{2p-q+1}^{(q-p)}|^{\frac{1}{p(q-p)}} = |a_p|^{\frac{1}{p}} < (\sigma + \varepsilon)^{2-\frac{q}{p}} (\tau + \varepsilon)^{\frac{q}{p}-1}, \quad \text{falls } q \leq 2p + 1,$$

$$(38) \quad |A_0^{(p+1)}|^{\frac{1}{p(p+1)}} = |a_p|^{\frac{1}{p}} < \tau + \varepsilon, \quad \text{falls } q \geq 2p + 1,$$

erfüllt sind.

Ist die Potenzreihe (31) über ihren Konvergenzkreis fortsetzbar, so ist $\tau < \varrho$ zu wählen, vgl. unter Nr. 9. Da auf alle Fälle $\sigma \leq \varrho$ ist, finden wir auf Grund von (36), (37), (38): *Wenn eine Reihe mit Hadamardschen Lücken fortsetzbar ist, sind die diese Lücken begrenzenden Koeffizienten a_p, a_q „klein“, nämlich es ist*

$$(39) \quad \overline{\lim} |a_p|^{\frac{1}{p}} < \overline{\lim}_{n \rightarrow \infty} |a_n|^{\frac{1}{n}},$$

$$(40) \quad \overline{\lim} |a_q|^{\frac{1}{q}} < \overline{\lim}_{n \rightarrow \infty} |a_n|^{\frac{1}{n}}$$

(n durchläuft die lückenlose Zahlenreihe $0, 1, 2, \dots$).

Das durch (39) ausgedrückte Resultat ist durch G. Szegö bekannt¹⁶⁾,

¹⁶⁾ A. a O. 4 e), Satz 7, S. 109.

und beide Ungleichungen (39), (40) sind in allgemeineren Resultaten von A. Ostrowski enthalten¹⁷⁾, jedoch scheinen die expliziten Abschätzungen (36), (37), (38) neu zu sein¹⁸⁾.

11. Potenzreihen mit ganzzahligen Koeffizienten. Wenn die Koeffizienten a_0, a_1, a_2, \dots der Reihe (1) ganze Zahlen sind (gewöhnliche ganze Zahlen $0, \pm 1, \pm 2, \dots$ oder auch ganze Zahlen eines bestimmten imaginär-quadratischen Zahlkörpers \mathfrak{f}) und die durch die Reihe definierte Funktion $f(z)$ regulär und eindeutig bleibt in der zusammenhängenden Komplementärmenge einer abgeschlossenen Punktmenge \mathfrak{A} , deren transfiniter Durchmesser kleiner als 1 ist, so ist $f(z)$ eine rationale Funktion.

Es folgt aus (8), da jetzt $\tau < 1$ ist, daß die Determinanten der Folge

$$(41) \quad A_0^{(1)}, A_0^{(2)}, A_0^{(3)}, \dots, A_0^{(k)}, \dots$$

von einer gewissen an dem Betrage nach kleiner als 1 sind. Aber die Determinanten (41) sind auch ganze Zahlen (gewöhnliche oder solche des Zahlkörpers \mathfrak{f}) und die einzige gewöhnliche ganze Zahl (die einzige ganze Zahl des imaginär-quadratischen Zahlkörpers \mathfrak{f}), deren Betrag < 1 ist, ist die Zahl 0. Die Determinanten (41) sind also von einer gewissen an sämtlich $= 0$. Hieraus folgt nach einem bekannten Kriterium¹⁹⁾, daß die Potenzreihe (1) eine rationale Funktion darstellt, w. z. b. w.

¹⁷⁾ A. Ostrowski, Über Potenzreihen, die überkonvergente Abschnittsfolgen besitzen, Sitzungsberichte d. preußischen Akademie (1923) S. 185—192, Satz II.

¹⁸⁾ Der Hadamardsche Lückensatz ist sowohl in (39) wie in (40) offenbar enthalten, aber man kann noch weiteres aus diesen Ergebnissen folgern. In den bekannten Untersuchungen von E. Fabry ist der folgende Satz enthalten: „Der Konvergenzradius der Reihe $\sum a_n z^n$ sei 1. Es sei δ gegeben, $0 < \delta < 1$, und n_1, n_2, \dots eine wachsende Folge von Indizes. Die Anzahl derjenigen nichtverschwindenden a_ν , für welche $n_k(1-\delta) < \nu < n_k$ ist, sei mit l_k , die Anzahl derjenigen, für welche $n_k < \nu < n_k(1+\delta)$ ist, mit r_k bezeichnet. Wenn

$$(a) \quad \lim_{k \rightarrow \infty} \frac{\log |a_{n_k}|}{n_k} = 0,$$

$$(b) \quad \lim_{k \rightarrow \infty} \frac{l_k}{n_k} = 0,$$

$$(c) \quad \lim_{k \rightarrow \infty} \frac{r_k}{n_k} = 0$$

ist, so ist der Kreis $|z| = 1$ natürliche Grenze für $\sum a_n z^n$.“ (Vgl. E. Fabry, Sur les séries de Taylor qui ont une infinité de points singuliers, Acta Math. 22 (1899) S. 65—87, ferner a. a. O. ³⁾ S. 47). Kombiniert man das durch (39) und (40) ausgedrückte Resultat des Textes mit einer zu ähnlichen Zwecken häufig brauchbaren Methode von G. Faber (vgl. Über die Nichtfortsetzbarkeit gewisser Potenzreihen, Sitzungsberichte der bayrischen Akademie 34 (1904) S. 63—74; vgl. ferner a. a. O. ³⁾ S. 70—73), so kann man den zitierten „zweiseitigen“ Satz von Fabry durch einen schärferen „einseitigen“ ersetzen: Die Behauptung des Fabry'schen Satzes bleibt bestehen, wenn man eine der beiden Voraussetzungen (b), (c) fortläßt.

¹⁹⁾ Vgl. a. a. O. ^{4c)} S. 23—24 oder a. a. O. ²⁾ Aufgabe 24, S. 103 und S. 305.

Der bewiesene Satz ist etwas allgemeiner, als die bisher bewiesenen ähnlichen Sätze²⁰⁾.

Der Spezialfall, in dem \mathfrak{A} abzählbar, also $\tau = 0$ ist, liefert folgende Aussage: *Wenn die Koeffizienten a_0, a_1, a_2, \dots der Potenzreihe*

$$a_0 + a_1 z + a_2 z^2 + \dots = F(z)$$

ganze Zahlen sind, so muß für die Funktion $F(z)$ einer der folgenden drei Fälle zutreffen: Entweder ist $F(z)$ eine rationale Funktion, oder ist $F(z)$ eine mehrdeutige Funktion, oder besitzt $F(z)$ un abzählbar viele singuläre Punkte.

Die Menge der singulären Punkte ist hier für eindeutige Funktionen als die Komplementärmenge der Menge der regulären Punkte definiert (was zulässig ist). Um dieselbe Tatsache noch etwas anders fassen zu können, bediene ich mich folgender Definition²¹⁾: Man sagt, daß die Potenzreihe

$$(42) \quad c_0 + c_1 z + c_2 z^2 + \dots + c_n z^n + \dots$$

mit rationalzahligen Koeffizienten c_0, c_1, c_2, \dots der *Eisensteinschen Bedingung* genügt, wenn eine ganze Zahl c existiert, so beschaffen, daß die Koeffizienten der Reihe

$$c_1 c z + c_2 c^2 z^2 + \dots + c_n c^n z^n + \dots$$

ganzzahlig sind. Mit Hilfe dieser Wendung läßt sich die Vereinigung des letzterwähnten Satzes mit einem trivialen Spezialfall des Eisensteinschen Satzes so ausdrücken: *Wenn die Koeffizienten c_0, c_1, c_2, \dots der Reihe (42) rationale Zahlen sind, und die Reihe (42) eine eindeutige Funktion mit höchstens abzählbar vielen singulären Punkten darstellt, so ist die dargestellte Funktion rational oder transzendent, je nachdem die Reihe (42) der Eisensteinschen Bedingung genügt oder nicht genügt.*

12. Schlussbemerkung. Die in den Nr. 7—11 besprochenen Beispiele lassen mannigfaltig verschiedene Fälle erkennen, in denen die Punktmenge \mathfrak{A} so auf die Funktion $f(z)$ abgestimmt ist, daß in (8) das Gleichheitszeichen am Platze ist.

²⁰⁾ In 4a), 4b) und 4e) wurden Kriterien benutzt, die in (30) als Spezialfälle enthalten sind, nämlich für $k = n$ und $k = n + 1$; dies entspricht dem Sonderfall $\kappa = \frac{1}{2}$ von (4). Das in 4c) benutzte Kriterium entspricht dem Sonderfall $\kappa = 1$ von (4) und unterscheidet sich von (8) in derselben Spezialisierung, wie das in 4c) erreichte Resultat von dem Ergebnis des Textes: Es wurde derzeit \mathfrak{A} der (mit dem Satz V der Nr. 1 zusammenhängenden) Restriktion unterworfen, daß entweder die Komplementärmenge von \mathfrak{A} oder die von \mathfrak{A}' einfach zusammenhängend ist.

²¹⁾ Vgl. a. a. O. 4a) S. 508. Rationalzahlig heißt: entweder sind alle c_0, c_1, c_2, \dots rational im gewöhnlichen Sinne, oder sie sind alle einem quadratisch-imaginären Körper \mathfrak{k} entnommen, zu dem dann auch die später zu erwähnende ganze Zahl c gehören muß.

a) Es ist $f(z)$ durch (20) oder (29) gegeben, und \mathfrak{A} ist die abgeschlossene geradlinige Verbindungsstrecke der Punkte 0 und 1; beide Seiten von (8) sind dann gleich $\frac{1}{4}$, wie in Nr. 7—8 besprochen wurde. ($f(z)$ kann über \mathfrak{A} hinaus fortsetzbar sein, wird aber dann notwendigerweise mehrdeutig; dies ist wohlbekannt, kann aber auch aus dem Hauptsatz dieser Arbeit geschlossen werden.)

b) Es seien $\lambda_1, \lambda_2, \lambda_3, \dots$ ganz, $\lambda_1 \geq 1, \lambda_{n+1} > 2\lambda_n$, das Konvergenzgebiet der Potenzreihe

$$a_{\lambda_1} z^{-\lambda_1} + a_{\lambda_2} z^{-\lambda_2} + \dots + a_{\lambda_n} z^{-\lambda_n} + \dots = f(z)$$

sei das Kreisäußere $|z| > \rho$ und \mathfrak{A} sei die abgeschlossene Kreisscheibe $|z| \leq \rho$. In diesem Falle ist

$$A_0^{(\lambda_{n+1})} = \pm a_{\lambda_n}^{\lambda_{n+1}}$$

und beide Seiten von (8) sind $= \rho$. (Hieraus folgt nach dem Kriterium der Nr. 9, daß der Rand von \mathfrak{A} natürliche Grenze für $f(z)$ ist, wie auch sonst bekannt. Die Überlegungen unter Nr. 10 können noch weitere, etwas anders geartete Beispiele liefern.)

c) Es sei die Komplementärmenge von \mathfrak{A} einfach zusammenhängend und der transfinite Durchmesser von \mathfrak{A} (der Abbildungsradius der Komplementärmenge von \mathfrak{A}) sei $= 1$. Es existiert²²⁾ eine Potenzreihe (1) mit ganzzahligen Koeffizienten a_0, a_1, a_2, \dots , deren Fortsetzung außerhalb \mathfrak{A} regulär ist, die aber keine rationale Funktion von z darstellt. Die rechte Seite von (8) ist dann $= 1$, und die linke Seite muß ≥ 1 (also $= 1$) sein, da doch alle $A_0^{(k)}$ ganzzahlig und unendlich viele unter ihnen von 0 verschieden sind. (Ist der Rand von \mathfrak{A} eine einfache geschlossene Kurve, so hat $f(z)$ diesen Rand zur natürlichen Grenze, wie der Hauptsatz und die Theorie der konformen Abbildung²³⁾ ergeben.)

d) Wie es der Reihe (1) unmittelbar anzusehen ist, gelangt man im Übergang

$$\text{von } f(z) \text{ zu } \frac{1}{c} f\left(\frac{z}{c}\right),$$

$$\text{von } a_m \text{ zu } a_m c^m,$$

$$\text{also von } A_n^{(k)} \text{ zu } A_n^{(k)} c^{(n+k-1)k};$$

c ist als nicht verschwindende Konstante angenommen. Es ist ferner leicht nachzurechnen, daß im Übergang

$$\text{von } f(z) \text{ zu } f(z - c),$$

$$\text{also von } a_m \text{ zu } a_m + \binom{m}{1} a_{m-1} c + \binom{m}{2} a_{m-2} c^2 + \dots + a_0 c^m$$

²²⁾ A. a. O. 4c), S. 32—34.

²³⁾ Aus dem schon einmal zitierten Satz a. a. O. ¹⁵⁾; der Schluß ist ausgeführt in 4c) S. 34.

$A_0^{(k)}$ unverändert bleibt (nicht $A_n^{(k)}$ für $n \geq 1$!); man kann zum Nachweis in Formel (17) $n = 0$ und $Q_\lambda(z) = (z + c)^\lambda = z^\lambda + \dots$ setzen, $\lambda = 0, 1, 2, \dots$

Man kann diese Tatsachen mit Satz II unter Nr. 1 so zusammenfassen: Bei linearen Transformationen der z -Ebene, die die beiden Punkte $z = 0$ und $z = \infty$ unverändert lassen, ändern sich die beiden Seiten von (4) auf die gleiche Weise, nämlich wie die Distanzen der singulären Punkte von $f(z)$ vom Nullpunkt. Bei linearen Transformationen der z -Ebene, die bloß den Punkt $z = \infty$ unverändert lassen, ändern sich die beiden Seiten von (8) auf die gleiche Weise, nämlich wie die Distanzen der singulären Punkte von $f(z)$ voneinander.

Diese Bemerkung gestattet die Beispiele unter a), b), c) durch Bewegungen und Dehnungen der Ebene zu vermehren.

(Eingegangen am 24. 10. 1927.)

Über quasi-normale Funktionsscharen und eine Verschärfung des Picardschen Satzes.

Von

Walter Saxer in Zürich.

Inhaltsübersicht.

	Seite.
Einleitung	708
I. Kapitel. Quasi-normale Funktionsscharen endlicher Ordnung	712
§ 1. Irreguläre Punkte endlicher Ordnung von konvergenten Folgen	712
§ 2. Definition und Eigenschaften quasi-normaler Funktionsscharen endlicher Ordnung	716
II. Kapitel. Juliasche Folgen und Quasi-Ausnahmefunktionen	717
§ 1. Problemstellung und allgemeine Sätze	717
§ 2. Verallgemeinerung eines Hilfssatzes von Ostrowski	718
§ 3. Notwendige Bedingungen für die Quasi-Ausnahmefunktionen	722
§ 4. Notwendige Bedingungen für die J -Ausnahmefunktionen	723
§ 5. Verteilung der Nullstellen und Pole der Quasi-Ausnahmefunktionen	725
§ 6. Die Darstellung der Quasi-Ausnahmefunktionen als kanonische Produkte	729
§ 7. Nachweis der Quasi-Ausnahmeeigenschaft als Quotienten kanonischer Produkte als Folge der Bedingungen (1) und (2)	730
§ 8. Die beiden Hauptsätze	735

Einleitung.

Der Picardsche Satz stellt eines der klassischen Resultate in der Theorie der ganzen Funktionen dar und hat zu einer Fülle von schönen und tiefen Untersuchungen Anlaß gegeben. Diese Arbeit knüpft an eine bestimmte Gruppe derselben an, welche durch die drei Namen Montel, Julia und Ostrowski charakterisiert wird. Julia¹⁾ ist es unter Anwendung und Verfeinerung einer Montelschen Begriffsbildung gelungen, nicht nur Aussagen über die Anzahl der a -Stellen einer ganzen Funktion, sondern auch über deren Argumente zu machen. Sein in dieser Richtung gewonnener Hauptsatz lautet:

Es sei $z = \sigma(t)$ ($1 \leq t < \infty$) die Parameterdarstellung einer stetigen vom Punkte $z = 1$ sich ins ∞ erstreckenden Kurve. Dann gibt es stets Kreise mit beliebig kleinem Radius und mit dem Zentrum z_0 mit folgender Eigenschaft: Wenn man diese Kurve $z = \sigma(t)$ um den Nullpunkt so weit rotiert, bis sie durch den Punkt z_0 geht und den gegebenen Kreis längs der rotierten Kurve ins ∞ verschiebt unter gleichzeitiger Vergrößerung seines Radius proportional dem Abstand vom Nullpunkt, dann nimmt eine ganze transzendente Funktion in diesem Gebiete sämtliche Funktionswerte bis auf höchstens einen einzigen ∞ oft an. Man kann als Kurve z. B. einen vom Nullpunkt ausgehenden Halbstrahl annehmen und erhält für dieses Gebiet einen Winkelraum mit beliebig kleinem Öffnungswinkel. Eine solche σ -Kurve bezeichnen wir auf Grund einer von Ostrowski eingeführten Terminologie als eine J -Kurve. Julia hat auch gezeigt, daß dieser Satz für weite Klassen meromorpher Funktionen richtig bleibt, wenn man statt einem zwei Ausnahmewerte zuläßt. Umgekehrt hat er an Beispielen festgestellt, daß der Satz nicht für alle meromorphen Funktionen richtig ist. Der obige Hauptsatz ist eine leichte Folgerung des wichtigsten Satzes in der Theorie der von Montel eingeführten normalen Funktionsscharen.

Julia hat weiter die folgende Frage untersucht: Gegeben eine beliebige Folge von ins ∞ strebenden Zahlen $\sigma_1, \sigma_2, \dots$. Wann ist die Funktionsschar $f(\sigma_n, z)$ in der Umgebung eines beliebigen Punktes $z \neq 0, \infty$ nicht normal, wobei $f(z)$ eine meromorphe Funktion bedeutet? Gehört nämlich zu einer Folge σ_n ein solcher Punkt z_0 , dessen Umgebung wir als einen Kreis mit dem festen, aber beliebig kleinen Radius δ annehmen, dann gibt

¹⁾ Julia hat seine Ergebnisse in zahlreichen Noten der C. R. 168 (1919) und 170 (1920) publiziert, ebenso in drei zusammenfassenden Arbeiten in den Annales de l'École Normale Supérieure (3) 36 (1919), (3) 37 (1920), (3) 38 (1921). Ferner findet sich eine zusammenfassende Darstellung in seinem Buche: G. Julia, Leçons sur les fonctions uniformes à point singulier essentiel isolé, Paris (1924) (Collection Borel).

es eine Folge von ins ∞ strebenden Kreisen mit den Zentren $z_0 \sigma_{n_k}$ und den Radien $\delta |\sigma_{n_k}|$, in denen $f(z)$ sämtliche Funktionswerte bis auf höchstens zwei Ausnahmewerte annimmt. Eine solche Folge werden wir als J -Folge und den dazu gehörigen Punkt z_0 als J -Punkt bezeichnen.

An dieser Stelle hat Ostrowski²⁾ mit einer wichtigen Untersuchung eingesetzt, die im wesentlichen zwei Ziele verfolgt. Der erste Teil enthält eine bedeutende Vertiefung der bisherigen Theorie der normalen Funktionsscharen. Sie gibt besonders eine genaue Analyse der Struktur einer nicht normalen Funktionsschar und stellt einen umkehrbaren Zusammenhang zwischen dem Charakter einer solchen und dem Wertvorrat der in dieser Schar vorkommenden Funktionen her.

Im 2. Teil werden sämtliche meromorphe Ausnahmefunktionen bestimmt, d. h. sämtliche für $|z| \geq R$ eindeutige und meromorphe Funktionen, für welche der zitierte Juliasche Satz nicht mehr gültig ist. Wir werden diese Funktionen im folgenden als J -Ausnahmefunktionen bezeichnen. Eine solche Ausnahmefunktion besitzt dann weder eine J -Kurve im vorhergehenden Sinn noch eine J -Folge, indem Ostrowski bewiesen hat, daß beide Erscheinungen nur simultan auftreten können. Ostrowski zeigt, daß diese Ausnahmefunktionen sich leicht aus denjenigen Ausnahmefunktionen bestimmen lassen, welche sich in der ganzen z -Ebene meromorph verhalten und daß diese letztere eine wohlbestimmte Klasse von der Ordnung Null sind.

Der Picardsche Satz sagt also aus, daß eine ganze Funktion in der z -Ebene sämtliche Werte mit Ausnahme von höchstens einem einzigen ∞ oft annimmt. Der Satz von Julia behauptet, daß jede ganze Funktion $g(z)$ mindestens eine ins ∞ strebende Folge von Kreisen besitzt, die man vom Nullpunkt aus unter einem beliebig kleinen Winkel sieht, mit der Eigenschaft, daß $g(z)$ jeden Funktionswert mit Ausnahme von höchstens einem einzigen in jenen Kreisen annimmt. Ungenauer kann man diesen Satz so formulieren, daß jede ganze Funktion mindestens einen beliebig schmalen Winkelraum mit dem Nullpunkt als Scheitel besitzt, indem sie sämtliche

²⁾ A. Ostrowski, Über Folgen analytischer Funktionen und einige Verschärfungen des Picardschen Satzes, *Math. Zeitschr.* 24 (1925), S. 215—258. Diese Arbeit wird im folgenden mit O. zitiert. Man vgl. auch: G. Valiron, Remarque sur un théorème de M. Julia, *Bulletin des Sciences Mathématiques* 49 (1925), p. 168—174.

Mit denselben Fragestellungen, aber mit anderen Methoden, befassen sich noch folgende Arbeiten: Milloux, *Journal de Mathématiques p. e. a.* (9) 3 (1924), p. 345—401. Valiron, Sur la distribution des valeurs des fonctions méromorphes, *Acta math.* 47 (1925), p. 117—141. Valiron, Sur une propriété des fonctions méromorphes d'ordre positif, *Bull. Soc. math.* 50 (1926), p. 168—174. Valiron, Compléments au théorème de Picard-Julia, *Bulletin des Sciences Mathématiques* 51 (1927), p. 167—183.

Funktionswerte mit Ausnahme von höchstens einem einzigen ∞ oft annimmt. Der Satz von Ostrowski behauptet, daß dies auch für eine meromorphe Funktion richtig bleibt, wenn man an Stelle von einem zwei Ausnahmewerte zuläßt und die sehr spezielle Klasse der J -Ausnahmefunktionen ausschließt. Diese Entwicklung legt die Frage nahe:

Welche meromorphen Funktionen $f(z)$ besitzen für jeden beliebig kleinen Winkel φ mindestens eine ins ∞ strebende Folge von Kreisen, die man vom Nullpunkt aus unter diesem Winkel sieht, mit der Eigenschaft, daß $f(z)$ in jenen Kreisen sämtliche Funktionswerte bis auf höchstens zwei feste Ausnahmewerte in unbeschränkter Anzahl annimmt?

Die Verschärfung gegenüber den bisherigen Fragestellungen beruht im Worte „unbeschränkt“. Die Beantwortung dieser Frage bedeutet einen gewissen Abschluß dieser Fragestellungen, die von der Ebene über den beliebig schmalen Winkelraum zu einem Kreise führte, der vom Nullpunkt aus unter beliebig kleinem Winkel gesehen wird.

Die Untersuchung dieser Frage kann in analoger Weise wie diejenige von Ostrowski durchgeführt werden. Man hat aber an Stelle der normalen Funktionsscharen *quasi-normale Funktionsscharen endlicher Ordnung* zu verwenden, ein Begriff, der ebenfalls von Montel³⁾ eingeführt wurde und der sich als ebenso tragfähig wie derjenige der normalen Funktionsscharen erweist. Eine beliebige Schar von in einem Gebiete G meromorphen Funktionen bildet eine quasi-normale Schar endlicher Ordnung, wenn man aus jeder Folge eine in einem zu G gehörigen beliebigen Bereiche B , abgesehen von endlich vielen irregulären Punkten endlicher Ordnung, gleichmäßig konvergente Teilfolge auswählen kann. Die Einführung der Ordnung eines irregulären Punktes einer gleichmäßig konvergenten Folge von meromorphen Funktionen wird nahegelegt durch den Satz, daß die Gesamtheit dieser Funktionen in einer beliebig kleinen Umgebung eines solchen Punktes sämtliche Funktionswerte bis auf höchstens einen annimmt. Einen irregulären Punkt einer in einem Bereiche B , abgesehen von diesem Punkte gleichmäßig konvergenten Reihe, bezeichnen wir nun als von endlicher Ordnung, wenn sämtliche Funktionen dieser Folge jeden Funktionswert mit Ausschluß von höchstens einem einzigen festen Ausnahmewert in der Umgebung dieses Punktes in beschränkter Anzahl annehmen. Montel zeigte, daß diejenigen meromorphen Funktionen, die in einem bestimmten Bereiche B drei Funktionswerte abc in beschränkter

³⁾ a) P. Montel, Sur les familles quasi-normales de fonctions holomorphes, Mémoires de l'Académie royale de Belgique, classe des Sciences, (2) 6 (1922), p. 1–41.

b) P. Montel, Sur les familles quasi-normales de fonctions analytiques, Bulletin de la Société Mathématique de France 52 (1924), p. 85–114.

Anzahl annehmen, eine solche quasi-normale Schar endlicher Ordnung in diesem Bereiche bilden.

Die Beantwortung unseres Problems ist mit der Lösung der folgenden Frage gleichbedeutend: Eine meromorphe Funktion $f(z)$ und eine beliebige Folge von komplexen Zahlen $\sigma_1, \sigma_2, \dots$ mit $\lim_{n \rightarrow \infty} \sigma_n = \infty$ sei gegeben. Wie muß die Funktion beschaffen sein, damit die Funktionsschar $f(\sigma_n z)$ für sämtliche Werte von $z \neq 0, \infty$ quasi-normal von endlicher Ordnung ist? Eine Funktion, welche diese Eigenschaft besitzt, bezeichnen wir als *Quasi-Ausnahmefunktion*. Offenbar genügen dann sämtliche meromorphe Funktionen mit Ausschluß dieser Quasi-Ausnahmefunktion infolge des obigen Montelschen Satzes unserer Verallgemeinerung des Picard-Julia-Ostrowskischen Satzes.

Für die Bestimmung dieser Quasi-Ausnahmefunktionen ist eine Eigenschaft der irregulären Punkte endlicher Ordnung sehr wesentlich, die aus der Montelschen Darstellung nicht direkt ersichtlich ist und von ihm auch nicht erwähnt wurde. Es zeigt sich nämlich, daß ein bekannter Hurwitzscher Satz für eine gleichmäßig konvergente Folge meromorpher Funktionen mit endlich vielen irregulären Punkten endlicher Ordnung noch gültig bleibt. Wir behaupten nämlich: *Wenn eine Folge von meromorphen Funktionen in einem Bereiche B , abgesehen von endlich vielen irregulären Punkten endlicher Ordnung, gleichmäßig konvergiert und die Funktionen in diesem Bereiche eine unbeschränkte Anzahl von Nullstellen besitzen, dann ist die Grenzfunktion identisch Null.* Das I. Kapitel enthält im wesentlichen den Nachweis dieses Satzes und seiner Umkehrung.

Im II. Kapitel bestimme ich sämtliche Quasi-Ausnahmefunktionen. Es sind ebenfalls Funktionen von der *Ordnung Null* mit wohlbestimmten Eigenschaften, die in gewissem Sinne sogar einfacher sind als diejenigen der *J*-Ausnahmefunktionen. Die letzteren lassen sich aus den ersteren sehr leicht herleiten. Zum Unterschied gegenüber den *J*-Ausnahmefunktionen gibt es *ganze Quasi-Ausnahmefunktionen*. Die Bestimmung dieser Quasi-Ausnahmefunktionen geschieht im wesentlichen mit den gleichen Hilfsmitteln wie bei Ostrowski. Es zeigt sich aber, daß sie voll ausgenutzt werden, ein Zeichen mehr, daß mit dieser Untersuchung die ganze Fragestellung einen natürlichen Abschluß fand.

Endlich möchte ich beifügen, daß ich einen Teil der Resultate schon in einer Note der C. R. der Akademie von Paris publizierte⁴⁾.

⁴⁾ W. Saxon, Sur les fonctions méromorphes quasi-exceptionnelles, C. R. de l'Académie des Sciences 184 (1927), p. 264—267.

I. Kapitel.

Quasi-normale Funktionsscharen endlicher Ordnung.

§ 1.

Irreguläre Punkte endlicher Ordnung von konvergenten Folgen.

1. Eine Folge von in einem beliebigen zusammenhängenden Gebiete G meromorphen Funktionen $f_1(z), f_2(z), \dots$ sei gegeben. Wir sagen in bekannter Weise, diese Folge konvergiere gleichmäßig in G , wenn für jeden innern Punkt P von G entweder die Folge $f_n(z)$ oder die Folge $\frac{1}{f_n(z)}$ in einer hinreichend kleinen Umgebung von P in gewöhnlichem Sinne gleichmäßig konvergiert. Bekanntlich ist dann die Grenzfunktion in G ebenfalls meromorph und kann speziell identisch ∞ werden. Diese Definition ist invariant gegenüber einer beliebigen, aber festen linearen Transformation der Funktionen $f_n(z)$. Zudem gilt der folgende, leicht beweisbare

Hilfssatz 1 von Ostrowski⁵⁾: *Es sei $f_n(z)$ eine Folge von in einem den unendlich fernen Punkt nicht im Innern enthaltenden Gebiet G meromorphen Funktionen, die in G gleichmäßig gegen $f(z)$ konvergiert, ϑ_n eine Folge gegen 1 konvergierender Konstanten. Dann konvergiert auch die Funktionsfolge $f_n(\vartheta_n z)$ gleichmäßig in G gegen $f(z)$.*

2. Es sei nun die Folge $f_n(z)$ in G im obigen Sinne in allen innern Punkten mit Ausnahme eines einzigen, den wir als den Nullpunkt wählen dürfen, gleichmäßig konvergent. Darunter verstehen wir folgendes: Wenn man aus diesem Gebiet G eine beliebig kleine Umgebung des Nullpunktes, also speziell einen beliebig kleinen Kreis mit dem Nullpunkt als Zentrum ausschneidet und in dieser Weise das Gebiet \bar{G} erhält, so ist die Folge $f_n(z)$ in \bar{G} gleichmäßig konvergent. Wir bezeichnen den Nullpunkt als einen zur Folge $f_n(z)$ gehörigen *irregulären Punkt*. Die Grenzfunktion $f(z)$ der Folge $f_n(z)$ wird dann im allgemeinen in einem irregulären Punkt einen wesentlich singulären Punkt besitzen. Wir wollen zeigen, daß sie sich in gewissen Fällen auch noch meromorph in einem irregulären Punkte verhält.

3. Es existiere ein solcher Wert a , so daß die Funktionen $f_n(z) - a$ in der Umgebung des irregulären Punktes, abgesehen von endlich vielen, keine Nullstellen besitzen, wobei zudem noch $f(z) \equiv a$. Wir behaupten, daß diese Voraussetzungen sich gegenseitig nicht vertragen, daß also der Nullpunkt in diesem Falle kein irregulärer Punkt sein kann. Das heißt: *Die meromorphen Funktionen einer in einem beliebigen Gebiete, abgesehen*

⁵⁾ Vgl. O., S. 226—227.

von einem innern irregulären Punkte gleichmäßig konvergenten Folge, nehmen in der Umgebung dieses irregulären Punktes sämtliche Funktionswerte bis auf höchstens einen einzigen an. Dieser Ausnahmewert müßte dann gleich der Grenzfunktion sein⁶⁾.

Beweis. Es sei $a \neq \infty$. Wir bilden die Funktionen $\frac{1}{f_n(z) - a}$. Diese Funktionen sind, abgesehen von endlich vielen, in einem genügend kleinen Kreise um den Nullpunkt regulär. Auf der Peripherie dieses Kreises konvergieren sie gleichmäßig gegen die meromorphe Funktion $\frac{1}{f(z) - a}$, die nicht identisch ∞ ist. Daraus folgt nach einem bekannten Satze von Weierstraß, daß die Folge $\frac{1}{f_n(z) - a}$ und also auch $f_n(z)$ im Innern dieses Kreises gleichmäßig konvergiert. Ist $a = \infty$, so betrachtet man die Folge $f_n(z)$.

4. Wir wollen die vorhergehende Betrachtung verallgemeinern. Es existiere ein solcher Wert a , wobei $f(z) \not\equiv a$, so daß die Funktionen $f_n(z) - a$, abgesehen von endlich vielen, in der Umgebung des Nullpunktes genau μ Nullstellen besitzen. Wir behaupten, daß die Grenzfunktion $f(z)$ sich unter dieser Voraussetzung auch im Nullpunkte meromorph verhält.

Beweis. Wir dürfen diesen Wert $a = \infty$ annehmen, sonst betrachtet man die Folge $\frac{1}{f_n(z) - a}$. Dann ist also die Grenzfunktion $f(z) \not\equiv \infty$. Wir bezeichnen die μ Pole der Funktion $f_n(z)$ in der Umgebung des Nullpunktes mit $z_1^{(n)} \dots z_\mu^{(n)}$, wobei $\lim_{n \rightarrow \infty} z_\nu^{(n)} = 0$, $1 \leq \nu \leq \mu$. Nun bilden wir die Funktionen

$$\varphi_n(z) = f_n(z)(z - z_1^{(n)}) \dots (z - z_\mu^{(n)}).$$

Diese Funktionen verhalten sich, abgesehen von endlich vielen, auf einem genügend kleinen Kreis um den Nullpunkt regulär. Zudem konvergieren sie auf der Peripherie dieses Kreises gleichmäßig gegen die Funktion $\varphi(z) = f(z)z^\mu \not\equiv \infty$. Daraus folgt nach Weierstraß, daß die Folge $\varphi_n(z)$ auch im Innern dieses Kreises gleichmäßig gegen die im Nullpunkt holomorphe Funktion $\varphi(z) = f(z)z^\mu$ konvergiert. Man erhält

$$f(z) = \frac{\varphi(z)}{z^\mu}.$$

$f(z)$ verhält sich also im Nullpunkt wirklich meromorph und besitzt dort einen Pol von höchstens μ -ter Ordnung.

⁶⁾ Ähnliche, viel allgemeinere Sätze finden sich in einer andern Untersuchung von Ostrowski. Man vergleiche: A. Ostrowski, Über vollständige Gebiete gleichmäßiger Konvergenz von Folgen analytischer Funktionen, Abhandl. aus dem Math. Seminar der Hamburger Universität 1 (1922), S. 327–350, insbesondere S. 347–350.

5. Um diese Schlußweise anwenden zu können, war die Voraussetzung wesentlich, daß die Funktionen, abgesehen von endlich vielen in der Umgebung des irregulären Punktes, dieselbe Anzahl, und zwar endlich viele Pole besitzen, wobei $f(z) \not\equiv \infty$ angenommen wurde. Aus dieser Voraussetzung werden wir weiter schließen, daß die Funktionen $f_n(z) - a$, wobei a einen beliebigen Wert $\not\equiv f(z)$ bedeutet, abgesehen von endlich vielen, in der Umgebung des Nullpunktes die gleiche, und zwar endliche Anzahl von Nullstellen besitzen. Das heißt: Besitzt ein Funktionswert $a \not\equiv f(z)$ diese Eigenschaft, so gilt dies von allen andern. Der Beweis kann leicht mit Hilfe der obigen Darstellung für $f(z)$ geführt werden. Wir haben in der Tat die Nullstellen der Funktionen

$$f_n(z) - a = \frac{\varphi_n(z)}{(z - z_1^{(n)}) \dots (z - z_\mu^{(n)})} - a$$

in der Umgebung des Nullpunktes zu bestimmen, wobei $a \not\equiv \frac{\varphi(z)}{z^\mu}$ und im übrigen beliebig. Diese Nullstellen sind mit den Nullstellen der Funktionen $\psi_n(z)$ identisch, wobei

$$\psi_n(z) = \varphi_n(z) - a(z - z_1^{(n)}) \dots (z - z_\mu^{(n)}),$$

da $\varphi_n(z_\nu^{(n)}) \neq 0$, wenn $1 \leq \nu \leq \mu$. Die Funktionen $\psi_n(z)$ sind in einem genügend kleinen Kreis mit dem Nullpunkt als Zentrum holomorph und konvergieren in diesem Bereich gleichmäßig gegen die Funktion $\varphi(z) - az^\mu$. Nach einem bekannten Satz von Hurwitz besitzen deshalb die Funktionen $\psi_n(z)$ und damit auch $f_n(z) - a$ von einem genügend großen Index $n > N(a)$ an in diesem Kreis die gleiche Anzahl von Nullstellen wie die Grenzfunktion $\varphi(z) - az^\mu$. Infolge unserer Voraussetzung über $a \not\equiv f(z) = \frac{\varphi(z)}{z^\mu}$ kann diese nicht identisch verschwinden. $\varphi(z)$ besitze in der Umgebung des Nullpunktes die Taylor-Entwicklung

$$\varphi(z) = c_\nu z^\nu + \dots,$$

wobei ν eine positive ganze Zahl bedeutet, die auch Null sein kann. In bekannter Weise zeigt man, daß die Anzahl der Nullstellen dieser Funktion $\varphi(z) - az^\mu$ in der Umgebung des Nullpunktes folgende Werte besitzt:

$\nu < \mu$:	Anzahl der Nullstellen = ν ,
$\nu = \mu, a \neq c_\nu$:	" "	" " = $\mu = \nu$,
$\nu = \mu, a = c_\nu$:	" "	" " = $\nu + k$, wobei k die kleinste positive ganze Zahl bedeutet, so daß $c_{\nu+k} \neq 0$. Diese Zahl ist sicher endlich.
$\nu > \mu, a \neq 0$:	" "	" " = μ ,
$\nu > \mu, a = 0$:	" "	" " = ν .

Diese Tabelle, kombiniert mit den obigen Bemerkungen, ergibt ohne weiteres den Beweis für unsere Behauptung. Die Tabelle zeigt aber noch mehr. Wir wollen das gesamte Ergebnis in der folgenden Zusammenfassung aussprechen:

Eine Folge von meromorphen Funktionen $f_1(z), f_2(z), \dots$ sei in einem zusammenhängenden Gebiete G , abgesehen von einem innern irregulären Punkte P , gleichmäßig konvergent. Die Grenzfunktion werde mit $f(z)$ bezeichnet. Es existiere ein solcher Wert $a_0 \equiv f(z)$, so daß die Funktionen $f_n(z) - a_0$, abgesehen von endlich vielen, in der Umgebung des irregulären Punktes die gleiche, und zwar endliche Anzahl von Nullstellen besitzen. Dann gelten folgende Eigenschaften:

a) *Die Grenzfunktion $f(z)$ verhält sich auch im irregulären Punkte meromorph.*

b) *Für jedes a mit Ausnahme eines bestimmten Wertes \bar{a} existiert eine solche positive ganze Zahl $N(a)$, so daß für $n > N(a)$ die Funktionen $f_n(z) - a$ dieselbe Anzahl Nullstellen in der Umgebung des irregulären Punktes besitzen. Dabei ist diese Anzahl k endlich und von a unabhängig. Ist die Grenzfunktion keine Konstante, so haben auch die Funktionen $f_n(z) - \bar{a}$, abgesehen von endlich vielen in der Umgebung des irregulären Punktes, die gleiche Anzahl Nullstellen. Diese Anzahl ist ebenfalls endlich und $> k$. Ist die Grenzfunktion eine Konstante, so ist dieser Ausnahmewert \bar{a} gleich der Grenzfunktion. Die Anzahl der Nullstellen der Funktionen $f_n(z) - \bar{a}$ in der Umgebung des irregulären Punktes ist in diesem Falle beliebig.*

Wir bezeichnen mit Montel einen solchen irregulären Punkt von endlicher Ordnung k .

6. Aus der Eigenschaft b) ziehen wir sofort den folgenden Schluß:

c) *Eine in einem Bereiche B , abgesehen von einem irregulären Punkt endlicher Ordnung, gleichmäßig konvergente Folge von meromorphen Funktionen $f_1(z), z_2(z), \dots$ sei gegeben. Gibt es einen solchen Funktionswert a , so daß die Anzahl der Nullstellen der Funktionen $f_n(z) - a$ in B unbeschränkt ist, so muß die Grenzfunktion $f(z)$ notwendigerweise diese Konstante a sein. Der schon angewendete Satz von Hurwitz gilt also auch noch für den Fall, daß sich im Konvergenzbereich ein irregulärer Punkt endlicher Ordnung befindet.*

Wie man sich leicht überzeugt, gilt auch die *Umkehrung des Satzes*. Wir behaupten also folgendes: Eine Folge von meromorphen Funktionen sei mit Ausnahme eines innern irregulären Punktes in einem zusammenhängenden Bereiche gleichmäßig konvergent. Ist die Nullstellen-Anzahl der Funktionen $f_n(z) - a$ für einen beliebigen Wert a unbeschränkt, dann

soll die Grenzfunktion gleich a sein. Unter diesen Voraussetzungen muß der irreguläre Punkt notwendigerweise von endlicher Ordnung sein. Der Beweis ergibt sich sofort aus den Voraussetzungen und unsern vorangegangenen Ausführungen.

Selbstverständlich kann man aus der Tatsache, daß die Grenzfunktion sich im irregulären Punkt meromorph verhält, noch nicht schließen, daß er von endlicher Ordnung sei. Beispiel: Die Grenzfunktion sei eine Konstante. Die Eigenschaft c) ist also charakteristisch für einen irregulären Punkt endlicher Ordnung und kann zu dessen Definition dienen.

§ 2.

Definition und Eigenschaften quasi-normaler Funktionsscharen endlicher Ordnung.

Eine Menge von in einem Gebiete G meromorphen Funktionen heißt nach Montel eine in G *normale Funktionsschar*, wenn man aus jeder Folge eine in G gleichmäßig konvergente Teilfolge auswählen kann. Sie heißt eine *quasi-normale Funktionsschar*, wenn man aus jeder Folge eine Teilfolge so auswählen kann, daß sie, abgesehen von endlich vielen irregulären Punkten, gleichmäßig in jedem ganz im Innern von G gelegenen Bereich konvergiert. Kann man speziell die Teilfolge immer so auswählen, daß die zu ihr gehörigen *irregulären Punkte von endlicher Ordnung* sind, so nennen wir die *quasi-normale Funktionsschar von endlicher Ordnung*. Diese quasi-normalen Funktionsscharen endlicher Ordnung sind durch zwei Sätze gegenüber einer allgemeinen quasi-normalen Funktionsschar ausgezeichnet.

Satz 1. *Die Grenzfunktion $f(z)$ irgendeiner in einem innern Teilbereich B von G gleichmäßig konvergenten Folge einer quasi-normalen Funktionsschar endlicher Ordnung verhält sich meromorph in G .*

Beweis. Auf Grund der obigen Definition kann man eine Teilfolge aus dieser in B gleichmäßig konvergenten Teilfolge so auswählen, daß sie in G , abgesehen von endlich vielen irregulären Punkten endlicher Ordnung, gleichmäßig konvergiert. Die Grenzfunktion dieser Teilfolge muß mit der Grenzfunktion der gegebenen Folge auf Grund eines klassischen Satzes der Funktionentheorie übereinstimmen. Nach der Eigenschaft a) des §1 verhält sich diese Grenzfunktion auch in den irregulären Punkten meromorph, womit der Beweis geleistet ist.

Satz 2. *Eine beliebige in einem beliebigen innern Teilbereich B_1 von G gleichmäßig konvergente Folge $f_1(z), f_2(z), \dots$ einer quasi-normalen Schar endlicher Ordnung sei gegeben. Es existiere ein solcher Wert a , daß die Anzahl der Nullstellen der Funktionen $f_1(z) - a, f_2(z) - a, \dots$*

in einem beliebigen innern Teilbereich B_2 von G unbeschränkt ist. Dann ist die Grenzfunktion notwendigerweise a .

Beweis. Wir wählen eine solche Teilfolge, daß die Anzahl der Nullstellen in B_2 der betreffenden Funktionen $f_{n_1}(z) - a$, $f_{n_2}(z) - a$, ... gegen ∞ strebt. Aus dieser wählen wir eine in G , abgesehen von endlich vielen irregulären Punkten endlicher Ordnung, gleichmäßig konvergente Teilfolge. Auf Grund der Eigenschaft c) des § 1 muß die Grenzfunktion dieser Teilfolge und also auch der gegebenen Folge gleich a sein.

II. Kapitel.

Juliasche Folgen und Quasi-Ausnahmefunktionen.

§ 1.

Problemstellung und allgemeine Sätze.

Es sei $f(z)$ eine in der ganzen z -Ebene meromorphe Funktion. Es sei $z = \sigma(t)$ ($1 \leq t < \infty$) die Parameterdarstellung einer stetigen Kurve, die außerhalb des Kreises $|z| = 1$ vom Punkte $z = 1$ aus ins ∞ verläuft. Wir bilden die Funktionsschar $f(\sigma(t)z)$, wobei z irgendeinen Wert $\neq 0, \infty$ annehmen darf. Es sei ferner σ_n eine beliebige Folge komplexer Zahlen mit $\lim_{n \rightarrow \infty} \sigma_n = \infty$ eine sogenannte σ -Folge. Wir bilden die Funktionsschar $f(\sigma_n z)$ mit einem beliebigen $z \neq 0, \infty$.

Wir werden im folgenden alle diejenigen meromorphen Funktionen bestimmen, bei denen die soeben gebildeten Funktionsscharen $f(\sigma(t)z)$ resp. $f(\sigma_n z)$ quasi-normal von endlicher Ordnung sind, und bezeichnen meromorphe Funktionen mit dieser Eigenschaft als Quasi-Ausnahmefunktionen. Wir zeigen zunächst, daß es genügt, eine dieser beiden Scharen zu betrachten, d. h. wir behaupten den

Satz 1. *Es sei $f(z)$ eine in der ganzen z -Ebene meromorphe Funktion, $z = \sigma(t)$ die Parameterdarstellung einer vom Punkte $z = 1$ aus außerhalb des Kreises $|z| = 1$ ins ∞ strebenden stetigen Kurve und $\sigma_1, \sigma_2, \dots$ eine beliebige Folge von komplexen Zahlen mit $\lim_{n \rightarrow \infty} \sigma_n = \infty$. Wenn die Funktionsschar $f(\sigma(t)z)$ quasi-normal von endlicher Ordnung ist für jedes $z \neq 0, \infty$, so gilt dasselbe von der Funktionsschar $f(\sigma_n z)$ und umgekehrt.*

Beweis. Es gebe eine Folge σ_n mit einem zugehörigen Punkte z_0 , für welche die Schar $f(\sigma_n z_0)$ nicht quasi-normal endlicher Ordnung ist. Man bilde eine stetige Kurve, für welche eine bestimmte Folge von ins ∞ strebenden t -Werten t_1, t_2, \dots die Beziehung

$$\sigma(t_n) = \sigma_n$$

gilt. Die Schar $f(\sigma(t)z_0)$ wird ebenfalls nicht quasi-normal endlicher Ordnung sein. Die Umkehrung läßt sich ganz analog beweisen.

Gemäß diesem Satz 1 genügt es also, für die Bestimmung der Quasi-Ausnahmefunktionen nur σ_n -Folgen zu betrachten. Im nächsten Satz werden wir zeigen, daß man die Auswahl in den σ_n -Folgen bedeutend einschränken kann, ohne die Allgemeinheit zu beeinträchtigen. Dabei legen wir keinen Wert darauf, diese Einschränkung so weit als möglich zu treiben. Wir treffen lediglich eine praktische Auswahl in Hinsicht auf unser Problem.

Satz 2. *Es sei $f(z)$ eine in der ganzen z -Ebene meromorphe Funktion und $\sigma_1, \sigma_2, \dots$ eine beliebige Folge reeller Zahlen mit $\lim_{n \rightarrow \infty} \sigma_n = \infty$.*

Die Funktionsschar $f(\sigma_n z)$ sei für $z \neq 0, \infty$ quasi-normal endlicher Ordnung. Dann gilt dasselbe, wenn $\sigma_1, \sigma_2, \dots$ eine beliebige Folge komplexer Zahlen mit $\lim_{n \rightarrow \infty} \sigma_n = \infty$ bedeutet.

Kurz ausgedrückt: Man kann sich bei der Bestimmung der Quasi-Ausnahmefunktionen auf reelle σ_n -Folgen beschränken.

Beweis. Es sei $\sigma_1, \sigma_2, \dots$ irgendeine komplexe Folge mit $\lim_{n \rightarrow \infty} \sigma_n = \infty$.

Nach Voraussetzung kann man eine Teilfolge so auswählen, daß die zugehörige Folge $f(|\sigma_n|z)$ in der Umgebung eines beliebigen, aber festen Punktes $\neq 0, \infty$, abgesehen von endlich vielen irregulären Punkten endlicher Ordnung, gleichmäßig konvergiert. Es sei $\sigma_1, \sigma_2, \dots$ bereits diese Teilfolge, wobei $\sigma_n = |\sigma_n| e^{i\vartheta_n}$. Wir können diese Folge zudem so bestimmen, daß $\lim_{n \rightarrow \infty} \vartheta_n = \vartheta$. Die Folge $f(z e^{-i\vartheta} \sigma_n)$ konvergiert, abgesehen

von den gleichen irregulären Punkten, nach dem 1. Hilfssatz von Ostrowski gleichmäßig in der Umgebung des Punktes z , da $\lim_{n \rightarrow \infty} \frac{e^{-i\vartheta} \sigma_n}{|\sigma_n|} = 1$. Zudem

sind diese irregulären Punkte auch für diese neue Folge von endlicher Ordnung, wie aus der Definition der Ordnung hervorgeht. Damit ist gezeigt, daß auch die Funktionsschar $f(\sigma_n z)$ mit komplexen σ_n -Folgen quasi-normal von endlicher Ordnung ist.

Da unsere Schar $f(\sigma_n z)$ für irgendeine σ_n -Folge und für einen beliebigen Punkt $\neq 0, \infty$ quasi-normal endlicher Ordnung ist, so gilt dasselbe auch für die Schar $f(\sigma_n z)$, wobei $r_1 < |z| < r_2$ und $r_1, r_2 \neq 0, \infty$. Der Beweis folgt in bekannter Weise aus dem Borelschen Überdeckungssatz.

§ 2.

Verallgemeinerung eines Hilfssatzes von Ostrowski.

Ist $f(z) \neq 0, \infty$ eine für $|z| \leq r$ meromorphe Funktion, so setzen wir

$$(1) \quad M(r, f) = \frac{1}{2\pi} \int_0^{2\pi} \log |f(re^{i\varphi})| d\varphi.$$

Bezeichnet man mit α_ν resp. β_ν die Nullstellen resp. die Pole der Funktion $f(z)$ im Kreise $|z| \leq r$ und besitzt diese in der Umgebung des Nullpunktes die Taylor-Entwicklung

$$f(z) = c_p z^p + \dots,$$

wobei $c_p \neq 0$ und p eine ganze Zahl bedeutet, die auch Null sein kann, so gilt bekanntlich die Jensensche Formel

$$(2) \quad e^{M(r, f)} = |c_p| r^p \frac{\prod_{\mu=1}^r \frac{r}{|\alpha_\mu|}}{\prod_{\nu=1}^r \frac{r}{|\beta_\nu|}}.$$

Es gilt der folgende

Hilfssatz 2. *Es sei $f_n(z)$ eine Folge von in einem Kreisring $r_1 < |z| < r_2$ meromorphen Funktionen, die in jedem innern Bereich dieses Kreisringes, abgesehen von endlich vielen irregulären Punkten endlicher Ordnung, gleichmäßig gegen $f(z)$ konvergieren. Dann konvergiert auch die Folge $M(\varrho, f_n)$ gegen $M(\varrho, f)$ gleichmäßig für feste r'_1, r'_2 und alle ϱ mit $r_1 < r'_1 \leq \varrho \leq r'_2 < r_2$, wenn $M(\varrho, 0) = -\infty$, $M(\varrho, \infty) = +\infty$ gesetzt wird.*

Ostrowski hat diesen Hilfssatz ohne Voraussetzung von irregulären Punkten formuliert und bewiesen⁷⁾.

Beweis. Die Gültigkeit dieses Hilfssatzes ergibt sich leicht aus derjenigen des Hilfssatzes von Ostrowski. Da der Beweis für diesen sich ganz knapp darstellen läßt, stellen wir ihn der Übersichtlichkeit halber dar. Wenn $f(z) \equiv 0$ oder $\equiv \infty$, so ergibt sich der Beweis aus der Darstellung (1) für $M(r, f)$. Es sei also $f(z) \not\equiv 0, \infty$. Dann besitzt die Funktion $f(z)$ in einem Kreisring $r_1 < \frac{r'_1 + r_1}{2} \leq |z| \leq \frac{r'_2 + r_2}{2} < r_2$ endlich viele Nullstellen $\alpha_1 \dots \alpha_\mu$ und endlich viele Pole $\beta_1 \dots \beta_\nu$. Von einem bestimmten Index an besitzen die Funktionen $f_n(z)$ in diesem Kreisring ebenfalls μ Nullstellen $\alpha_1^{(n)} \dots \alpha_\mu^{(n)}$ und ν Pole $\beta_1^{(n)} \dots \beta_\nu^{(n)}$. Zudem gilt

$$\lim_{n \rightarrow \infty} \alpha_k^{(n)} = \alpha_k \quad (1 \leq k \leq \mu),$$

$$\lim_{n \rightarrow \infty} \beta_k^{(n)} = \beta_k \quad (1 \leq k \leq \nu).$$

Wir bilden nun

$$P_n(z) = z^p \prod_{k=1}^{\mu} (z - \alpha_k^{(n)}) \quad \text{resp.} \quad \prod_{k=1}^{\mu} (z - \alpha_k^{(n)}), \quad \text{wenn } p \geq 0 \text{ resp. } < 0,$$

$$Q_n(z) = \prod_{k=1}^{\nu} (z - \beta_k^{(n)}) \quad \text{resp.} \quad z^{-p} \prod_{k=1}^{\nu} (z - \beta_k^{(n)}), \quad \text{wenn } p \geq 0 \text{ resp. } < 0,$$

⁷⁾ Vgl. O., S. 243–245.

und setzen

$$f_n(z) = \frac{P_n(z)}{Q_n(z)} \varphi_n(z) \quad \text{und ebenso} \quad f(z) = \frac{P(z)}{Q(z)} \varphi(z),$$

wobei die Funktionen $\varphi_n(z)$ und $\varphi(z)$ in unserm Kreisring regulär und von Null verschieden sind. Offenbar gilt:

$$M(r, f_n) = M(r, \varphi_n) + M(r, P_n) - M(r, Q_n)$$

und

$$M(r, f) = M(r, \varphi) + M(r, P) - M(r, Q).$$

Der Hilfssatz von Ostrowski ist bewiesen, wenn wir zeigen können, daß

$$(3a) \quad \lim_{n \rightarrow \infty} M(r, \varphi_n) = M(r, \varphi),$$

$$(3b) \quad \lim_{n \rightarrow \infty} M(r, P_n) = M(r, P),$$

$$(3c) \quad \lim_{n \rightarrow \infty} M(r, Q_n) = M(r, Q).$$

Nun gilt in unserm Kreisring gleichmäßig

$$\lim_{n \rightarrow \infty} P_n(z) = P(z) \quad \text{und} \quad \lim_{n \rightarrow \infty} Q_n(z) = Q(z).$$

Daraus folgt für den Kreisring $r'_1 \leq |z| \leq r'_2$ gleichmäßig

$$\lim_{n \rightarrow \infty} \varphi_n(z) = \varphi(z).$$

Denn für einen Punkt $z \neq \alpha_k, \beta_k$ ist dies klar, und für einen solchen Punkt folgt diese Beziehung aus dem Theorem von Weierstraß. Für die Funktionen $\varphi_n(z)$ ist nun die Gleichung (3a) evident, da in unserm Kreisring ebenfalls gleichmäßig gilt:

$$\lim_{n \rightarrow \infty} \log |\varphi_n(z)| = \log |\varphi(z)|.$$

Die Beziehungen (3b) und (3c) ergeben sich direkt aus der Jensenschen Darstellung der Funktionen $M(r, P_n)$ und $M(r, Q_n)$.

Für den Beweis des allgemeinen Hilfssatzes können wir r'_1, r'_2 so annehmen, daß auf der Peripherie der Kreise $|z| = r'_1, r'_2$ keine irregulären Punkte unserer Folge $f_n(z)$ liegen. Im Kreisring r'_1, r'_2 liegen endlich viele irreguläre Punkte endlicher Ordnung. Ihre absoluten Beträge d_k seien so bezeichnet, daß

$$r'_1 < d_1 < d_2 < \dots < d_m < r'_2.$$

Bedeutet ε_1 eine beliebig kleine, feste, positive Zahl, so gilt der Hilfssatz (2) nach Ostrowski zunächst in den Kreisringen $(r'_1, d_1 - \varepsilon_1)$, $(d_1 + \varepsilon_1, d_2 - \varepsilon_1)$, \dots , $(d_m + \varepsilon_1, r'_2)$. Nun unterscheiden wir zwei Fälle:

a) $f(z)$ sei $\neq 0, \infty$.

Wir umgeben die Nullstellen und Pole von $f(z)$ sowie die irregulären Punkte in einem solchen kritischen Kreisring $(d_k - \varepsilon_1, d_k + \varepsilon_1)$ mit Kreisen, deren Radien ε_2 und $\varepsilon_2 < \varepsilon_1$ beliebig klein, aber fest gewählt werden. Außerhalb dieser Kreise $K(\varepsilon_2)$ gibt es dann zu jeder Zahl ε_3 eine solche ganze, positive Zahl $N(\varepsilon_2, \varepsilon_3)$, so daß im Kreisring $(d_k - \varepsilon_1, d_k + \varepsilon_1)$ gilt:

$$(4) \quad \log \left| \frac{f_n(z)}{f(z)} \right| < \varepsilon_3 \quad \text{für } n > N(\varepsilon_2, \varepsilon_3).$$

Die Funktionen $f_n(z)$ und $f(z)$ besitzen nach unseren früheren Ausführungen in der Umgebung eines irregulären Punktes endlicher Ordnung die Darstellung

$$(5) \quad f_n(z) = \frac{\psi_n(z)}{(z - \beta_{1,j}^{(n)}) \dots (z - \beta_{k,j}^{(n)})}, \quad f(z) = \frac{\psi(z)}{(z - \beta_j)^k}.$$

Der Index j charakterisiert den irregulären Punkt, k bedeutet dessen Ordnung. $\psi_n(z)$ und $\psi(z)$ verhalten sich in der Umgebung dieses irregulären Punktes holomorph. Zudem gilt im Kreise $K(\beta_j, \varepsilon_2)$ gleichmäßig

$$(6) \quad \lim_{n \rightarrow \infty} \psi_n(z) = \psi(z) \quad \text{und} \quad \lim_{n \rightarrow \infty} \beta_{\kappa j}^{(n)} = \beta_j \quad \text{für } 1 \leq \kappa \leq k.$$

Man erhält

$$(7) \quad M(r, f) = \frac{1}{2\pi} \int_{C_1} \log |f(re^{i\varphi})| d\varphi + \sum_{C_2} \int \log |f(re^{i\varphi})| d\varphi,$$

wobei $d_k - \varepsilon_1 < r < d_k + \varepsilon_1$, ganz entsprechend stellt man $M(r, f_n)$ dar. Das erste Integral erstreckt sich nur über diejenigen Punkte der Kreisperipherie $|z| = r$, die sich außerhalb der Kreise $K(\varepsilon_2)$ befinden. Die Summe der Integrale bezieht sich auf die Bögen des Kreises $|z| = r$, die innerhalb dieser Kreise liegen. Da der Radius eines solchen Kreises beliebig klein gewählt werden darf, die Anzahl der Summanden endlich ist, so wird diese Summe dank Gleichung (5) beliebig klein ausfallen. Kombiniert man diese Bemerkung mit Ungleichung (4), so ergibt sich

$$|M(r, f) - M(r, f_n)| < \varepsilon_3 + c\varepsilon_4 \quad \text{für } n > N(\varepsilon_2, \varepsilon_3).$$

c bedeutet eine endliche Zahl, ε_4 eine beliebig kleine, positive Zahl. Aus dieser Ungleichung folgt der Beweis für unsern Hilfssatz (2).

b) $f(z)$ sei z. B. $\equiv 0$. Man kann den Beweis genau gleich führen. Es gilt auch in diesem Fall in der Umgebung eines irregulären Punktes die Darstellung (5). Zudem gilt im Kreise $K(\beta_j, \varepsilon_2)$ gleichmäßig

$$\lim_{n \rightarrow \infty} \psi_n(z) = 0 \quad \text{und} \quad \lim_{n \rightarrow \infty} \beta_{\kappa j}^{(n)} = \beta_j \quad (1 \leq \kappa \leq k).$$

Zerlegt man $M(r, f_n)$ und $M(r, f)$ wieder nach der gleichen Weise wie in Gleichung (7), so folgt aus den obigen Bemerkungen der Hilfssatz (2).

§ 3.

Notwendige Bedingungen für die Quasi-Ausnahmefunktionen.

Auf Grund der allgemeinen Sätze über quasi-normale Funktionsscharen endlicher Ordnung und des Hilfssatzes (2) ist es leicht, zwei notwendige Bedingungen für die Quasi-Ausnahmefunktionen aufzustellen. Diese beiden Bedingungen erweisen sich nachher als hinreichend zur Bestimmung sämtlicher Quasi-Ausnahmefunktionen.

1. Bedingung. *Es existieren zwei nur von $f(z)$ abhängige positive Konstanten k_1 und k_2 mit folgender Eigenschaft: Wenn für einen Wert r eine der folgenden Ungleichungen erfüllt ist*

$$M(r, f) \geq -k_1 \quad \text{oder} \quad M(r, f) \leq +k_1,$$

dann besitzt die Quasi-Ausnahmefunktion $f(z)$ im Kreisring $r \leq |z| \leq 2r$ im ersten Fall höchstens k_2 Nullstellen und im zweiten Fall höchstens k_2 Pole.

Geometrisch gesprochen: Wenn die stückweise lineare Funktion $M(r, f)$ von $\log r$ in einem Punkte über die Gerade $y = -k_1$ resp. unter die Gerade $y = +k_1$ ragt, so kann der Richtungskoeffizient in dem sich diesem Punkte nach rechts anschließenden Intervall von der Breite $\log 2$ höchstens um k_2 zu- resp. abnehmen.

Wir beweisen z. B. den ersten Teil der Aussage indirekt. Zu einem beliebigen k_1 gäbe es eine Folge r_1, r_2, \dots mit $\lim_{n \rightarrow \infty} r_n = \infty$, so daß $M(r_k, f) \geq -k_1$ und die Anzahl der Nullstellen der Funktion $f(z)$ in den Kreisringen $r_k \leq |z| \leq 2r_k$ mit wachsendem k beliebig groß würde. Wir setzen

$$\frac{r_2}{r_1} = \sigma_1, \quad \frac{r_3}{r_1} = \sigma_2, \quad \dots, \quad \frac{r_{n+1}}{r_1} = \sigma_n, \quad \dots$$

und bilden die Funktionsschar $f(\sigma_n z) = f_n(z)$. Gemäß unserer Annahme muß diese Schar in jedem Punkt $z \neq 0, \infty$ und also auch im Kreisring $\frac{r_1}{2} < |z| < 4r_1$ quasi-normal endlicher Ordnung sein. Man kann deshalb aus unserer Folge eine Teilfolge so auswählen, daß sie, abgesehen von endlich vielen irregulären Punkten endlicher Ordnung im Kreisring $r_1 \leq |z| \leq 2r_1$ gleichmäßig konvergiert. Unsere Funktionen sind zudem meromorph in diesem Bereich. Man kann also den Satz (2) des § 2, Kap. I und den Hilfssatz (2) anwenden. Nach diesen Sätzen müßte die Grenzfunktion $f(z)$ Null sein und deshalb $M(r, f_n)$ im Kreisring $r_1 \leq |z| \leq 2r_1$ also speziell für $r = r_1$ gegen $-\infty$ streben, was unserer Voraussetzung über $M(r_k, f)$ widerspricht. Ganz analog läßt sich der zweite Teil unserer Behauptung beweisen.

2. Bedingung. Wir bezeichnen mit $N(r)$ die Differenz zwischen der Anzahl der Nullstellen von $f(z)$ und der Anzahl der Pole im Kreis $|z| = r$. Es existiert eine nur von $f(z)$ abhängige positive Konstante k_3 mit folgender Eigenschaft: Wenn für einen Wert r die Ungleichung

$$|M(r, f)| \leq k_1$$

erfüllt ist, dann gilt gleichzeitig

$$|N(r, f)| \leq k_3.$$

Geometrisch gesprochen: der Richtungskoeffizient der stückweise linearen Funktion $M(r, f)$ beträgt im Parallelstreifen $-k_1 \leq y \leq k_1$ höchstens k_3 .

Beweis (indirekt). Es gebe eine ins ∞ strebende Folge r_1, r_2, \dots mit folgenden Eigenschaften:

$$(8) \quad |M(r_k, f)| \leq k_1$$

$$(9) \quad \lim_{n \rightarrow \infty} N(r_n, f) = +\infty \quad \text{oder} \quad -\infty.$$

Wir bilden wie im vorhergehenden Fall die Funktionsschar $f(\sigma_n z) = f_n(z)$. Laut unserer Voraussetzung gilt

$$(10) \quad |M(r_1, f_n)| \leq k_1$$

$$(11) \quad \lim_{n \rightarrow \infty} N(r_1, f_n) = +\infty \quad \text{oder} \quad -\infty.$$

Man kann eine im Kreisring $r_1 \leq |z| \leq 2r_1$, abgesehen von endlich vielen irregulären Punkten endlicher Ordnung, gleichmäßig konvergente Teilfolge herausgreifen. Nach dem 2. Hilfssatz könnte die Grenzfunktion infolge Ungleichung (10) nicht identisch 0 oder ∞ sein. Umgekehrt müßte sie infolge Gleichung (11) identisch einen dieser beiden Werte annehmen, womit der Widerspruch hergestellt ist.

Diese Bedingungen werden für ganze Quasi-Ausnahmefunktionen sehr einfach. Denn für eine solche Funktion ist $M(r, f)$ entweder eine Konstante oder eine wachsende, gegen ∞ strebende Funktion. Man findet deshalb als notwendige Bedingung auf Grund unserer Bedingung (1): Die Anzahl der Nullstellen einer ganzen Quasi-Ausnahmefunktion in einem Kreisring mit beschränkter logarithmischer Breite ist beschränkt.

§ 4.

Notwendige Bedingungen für die J -Ausnahmefunktionen.

Ostrowski hat für die absoluten Beträge der Nullstellen und Pole einer Ausnahmefunktion die folgenden drei nicht vollständig unabhängigen, notwendigen Bedingungen aufgestellt:

a) Es gibt eine nur von $f(z)$ abhängige Konstante c_1 , $c_1 > 0$, so daß für alle $r > 0$

$$-c_1 \leq N(r) \leq c_1$$

gilt.

b) Für jedes $r > 0$ liegen im Kreisring $r \leq |z| \leq 2r$ weniger als c_2 Nullstellen bzw. Pole, wo die ganze Zahl c_2 nur von $f(z)$ abhängt.

c) Es gibt eine nur von $f(z)$ abhängige Konstante $c_3 > 0$ derart, daß für alle Nullstellen α_r bzw. Pole β_r

$$M(|\alpha_r|) \leq c_3, \quad M(|\beta_r|) \geq -c_3.$$

Die Bedingung (c) ist eine leichte Folge der Definition einer J -Ausnahmefunktion und des Hilfssatzes (2). Wäre sie nämlich nicht erfüllt, so könnte man eine solche Folge von Funktionen $f_n(z) = f(\sigma_n z)$ konstruieren, die in einem bestimmten Kreisring gleichmäßig gegen ∞ streben müßte. Gleichzeitig wären sämtliche Funktionen in einem innern Punkt dieses Kreisringes gleich Null, was mit der gleichmäßigen Konvergenz gegen ∞ nicht vereinbart werden kann. Offenbar können wir diese Konstante c_3 gleich unserer Konstanten k_1 setzen.

Die Bedingungen (a) und (b) sind nun eine leichte Folge unserer Bedingungen (1) und (2). Gemäß der Bedingung (c) kann nämlich $f(z)$ so lange keine Nullstellen bzw. Pole besitzen, als $M(r, f) > k_1$ bzw. $< -k_1$. Außerhalb unseres Parallelstreifens kann also $N(r, f)$ oberhalb nur abnehmen und unterhalb nur zunehmen. Da $|N(r, f)|$ nach unserer Bedingung (2) innerhalb desselben und in den Schnittpunkten mit den Grenzgeraden dieses Streifens beschränkt ist, so folgt daraus, daß $|N(r, f)|$ überhaupt beschränkt ist, d. h. die Bedingung (a). Solange das Bild von $M(r, f)$ unterhalb der oberen Grenzgeraden unseres Streifens verläuft, muß die Anzahl der Pole in einem Kreisring mit endlicher logarithmischer Breite nach unserer ersten Bedingung beschränkt sein. Da nun der beschränkte Richtungskoeffizient oberhalb nur abnehmen kann, so folgt daraus, daß die Anzahl der Pole in einem beliebigen Kreisring mit beschränkter logarithmischer Breite beschränkt sein muß, d. h. die Bedingung (b). Offenbar kann man in ganz analoger Weise diese Bedingung auch für die Nullstellen beweisen.

Zusammenfassend läßt sich sagen, daß die absoluten Beträge der Nullstellen und Pole einer Quasi-Ausnahmefunktion bzw. einer Ausnahmefunktion unsern Bedingungen (1) und (2) bzw. diesen Bedingungen plus Bedingung (c) genügen müssen.

§ 5.

Verteilung der Nullstellen und Pole der Quasi-Ausnahmefunktionen.

Wir haben gesehen, daß die Verteilung der Nullstellen und Pole einer J -Ausnahmefunktion sehr regelmäßig ist. In jedem Kreisring von beschränkter logarithmischer Breite befindet sich eine beschränkte Anzahl von Nullstellen und Polen. Zudem dürfen — wenn der Einfluß der Nullstellen so groß geworden ist, daß $M(r, f)$ eine gewisse Schranke überschreitet — mit wachsendem r nur noch Pole folgen, bis dieser Einfluß kompensiert ist.

Wir werden sehen, daß die Verteilung der Nullstellen und Pole bei den Quasi-Ausnahmefunktionen nicht mehr so regelmäßig ist. Allerdings ist die Anzahl der Nullstellen und Pole einer Quasi-Ausnahmefunktion in einem Kreis $|z| = r$ höchstens gleich der Anzahl der Nullstellen und Pole einer Ausnahmefunktion, aber ihre Verteilung ist bedeutend willkürlicher. Wir werden im folgenden eine Anzahl von Hilfssätzen formulieren, welche diese Verteilung beschreiben und unsere Behauptungen beweisen. Zunächst führen wir eine neue Hilfsgröße, die zu einem bestimmten r -Werte gehörige „Grenzzahl“ ein. Der Grund ihrer Einführung und der Sinn dieser Grenzzahl wird aus den folgenden Ausführungen klar hervorgehen.

Definition. Zu jedem Wert r einer Quasi-Ausnahmefunktion gehört eine bestimmte Grenzzahl $g(r)$, die durch folgende Gleichung definiert wird:

$$g(r) = \sqrt{\frac{2|M(r, f)|}{k_2 \log 2}}, \quad \text{wenn } |M(r, f)| > k_1,$$

$g(r) = g$, wobei g eine geeignete, positive Konstante bedeutet, wenn $|M(r, f)| \leq k_1$. Dieses „geeignet“ wird nachher präzisiert werden.

Auf Grund dieses Begriffes werden wir nun eine Reihe von Hilfssätzen beweisen.

Hilfssatz a. *Es sei ε_1 eine beliebig kleine, feste, positive Zahl. Dann gibt es zu jeder Zahl ε_1 eine Zahl $\bar{k}(\varepsilon_1)$ mit folgender Eigenschaft: Für alle r -Werte, welche die Ungleichung erfüllen:*

$$|M(r, f)| \geq \bar{k}(\varepsilon_1),$$

gilt gleichzeitig

$$|N(r, f)| < k_2 g(r)(1 + \varepsilon_1).$$

Beweis. Wir gehen von folgender Bemerkung aus: Es sei für einen bestimmten r -Wert $M(r, f) = k_1$. Nach unserer Bedingung (2) gilt $N(r) \leq k_3$. Zudem besitzt die Funktion $f(z)$ im Kreisring $r \leq |z| \leq 2r$ höchstens k_2 Nullstellen nach unserer Bedingung (1). Deshalb können die

Funktionen $N(r, f)$ und $M(r, f)$ nicht beliebig rasch wachsen. Man erhält unmittelbar folgende Schranken:

$$M(r 2^k, f) \leq k_1 + (k_3 + k_2) \log 2 + (k_3 + 2k_2) \log 2 + \dots + (k_3 + k k_2) \log 2$$

$$M(r 2^k, f) \leq k_1 + \left\{ k k_3 + \frac{k_2 k (k+1)}{2} \right\} \log 2 \leq \frac{k^2 k_2}{2} (1 + \eta_1) \log 2$$

und

$$N(r 2^k, f) \leq k_3 + k k_2 \leq k k_2 (1 + \eta_2)$$

wobei η_1, η_2 beliebig kleine, feste, positive Größen bedeuten. Im Falle des Gleichheitszeichens ist $\lim_{k \rightarrow \infty} \frac{k}{g(r 2^k)} = 1$ wie aus der obigen Definition hervorgeht. Offenbar erhält man in dieser Weise die maximalste Steigung in allen Punkten des Intervalles $(r, r 2^k)$, die überhaupt möglich ist. Durch Spiegelung dieser Figur an einer senkrechten (zum Streifen) in einem beliebigen Punkte dieses Intervalles bemerkt man, daß der Richtungskoeffizient auch nicht kleiner sein kann als die soeben berechnete negative Schranke. Genau gleich kann man schließen, wenn $M(r, f)$ negativ ist.

Bei festem ε_1 darf man annehmen, daß $\bar{k}(\varepsilon_1) = k_1$ ist unter entsprechender Änderung von k_3 . Normiert man nun unsere Zahl g passend, so läßt sich der Hilfssatz a kurz so formulieren:

Bedeutet ε_1 eine beliebig kleine, feste, positive Zahl und $g(r)$ die zu einem Werte r gehörige Grenzzahl, so gilt:

$$|N(r, f)| < k_2 g(r) (1 + \varepsilon_1),$$

d. h. in jeder Höhe hat der absolute Betrag des Richtungskoeffizienten von $M(r, f)$ eine bestimmte, obere Schranke.

Hilfssatz b. *Es sei ε_2 eine beliebig kleine, feste, positive Zahl. Dann gibt es zu jeder endlichen Zahl c eine Zahl $k^*(\varepsilon_2, c)$ mit folgender Eigenschaft:*

Für alle r Werte, welche die Ungleichung

$$|M(r, f)| \geq k^*(\varepsilon_2, c)$$

erfüllen, gilt gleichzeitig

$$(1 - \varepsilon_2) < \left| \frac{M(c r, f)}{M(r, f)} \right| < 1 + \varepsilon_2.$$

Beweis. Es sei z. B. $M(r, f)$ und $N(r, f)$ positiv. Man erhält den maximalsten Wert für $M(c r, f)$, wenn $M(r, f)$ gemäß der beim Beweis des vorhergehenden Hilfssatzes besprochenen Art wächst. Deshalb gewinnen wir folgende Abschätzung:

$$M(c r, f) < M(r, f) + k_2 g(r) (1 + \varepsilon_1) \log c + k_2 \kappa \log c,$$

wo κ eine endliche, von c und k_2 abhängige Konstante bedeutet.

Berücksichtigt man die Definition von $g(r)$, so folgt aus dieser Ungleichung die obere Schranke für den Quotienten $\frac{M(cr, f)}{M(r, f)}$. Mit Hilfe der oben gemachten Bemerkung über die minimalsten Richtungskoeffizienten erhält man die untere Schranke.

Auf Grund unserer Definition ergibt sich zu diesem Hilfssatz das folgende

Korollar b. Es sei ε_3 eine beliebig kleine, feste, positive Zahl. Dann gibt es zu jeder endlichen Zahl c eine Zahl $k^{**}(\varepsilon_3, c)$ mit folgender Eigenschaft:

Für alle r Werte, welche die Ungleichung

$$|M(r, f)| \geq k^{**}(\varepsilon_3, c)$$

erfüllen, gilt gleichzeitig

$$1 - \varepsilon_3 < \left| \frac{g(cr, f)}{g(r, f)} \right| < 1 + \varepsilon_3.$$

Bei festem ε_3 und c kann offenbar $k^{**}(\varepsilon_3, c)$ gleich k_1 gesetzt werden.

Hilfssatz c. *Es sei r ein beliebiger Wert, $g(r)$ seine zugehörige Grenzwahl. Die Anzahl der Pole einer Quasi-Ausnahmefunktion im Kreisring $(r, r2^x)$ resp. im Kreisring $(r2^{-x}, r)$ wobei $x \geq g(r)$, ist $< cx$, c bedeutet eine endliche, positive Konstante.*

Beweis. Wir unterscheiden zwei Fälle.

$$(1) \quad |M(r, f)| \geq k_1,$$

dann gilt

$$g(r) = \sqrt{\frac{2|M(r, f)|}{k_1 \log 2}}.$$

$M(r, f)$ sei positiv. Dann konstruieren wir wieder wie in den vorhergehenden Betrachtungen die maximalste $M(r, f)$ -Funktion, die überhaupt möglich ist. Man erhält auf Grund unserer Bedingung (1)

$$\begin{aligned} M(r2^x) &\leq M(r) + (N(r) + k_2) \log 2 + \dots + (N(r) + x k_2) \log 2 \\ &= M(r) + \left[x N(r) + \frac{k_2 x(x+1)}{2} \right] \log 2. \end{aligned}$$

Nach der Definition von $g(r)$ und dem Hilfssatz a schließen wir

$$\begin{aligned} M(r2^x) &< \frac{(g(r))^2 k_2 \log 2}{2} + \left[x g(r) k_2 (1 + \varepsilon_1) + \frac{k_2 x(x+1)}{2} \right] \log 2 \\ &< 2 k_2 x^2 \log 2 \left[1 + \varepsilon_1 + \frac{1}{2x} \right]. \end{aligned}$$

Daraus schließen wir nach der Definitions-Gleichung für die Grenzwahl

$$g(r2^x) < 2x \sqrt{1 + \varepsilon_1 + \frac{1}{2x}}$$

und nach Hilfssatz a

$$N(r 2^x) < 2k_2 x \left(\sqrt{1 + \varepsilon_1 + \frac{1}{2x}} \right) (1 + \varepsilon_1).$$

Offenbar erhält man infolge des Hilfssatzes a und der Bedingung 1 als obere Schranke für die Anzahl der Pole in unserem Kreisring $(r, r 2^x)$ den Wert

$$4k_2 x \left(\sqrt{1 + \varepsilon_1 + \frac{1}{2x}} \right) (1 + \varepsilon_1).$$

Wäre dies nicht der Fall, so wäre für mindestens einen r -Wert des abgeschlossenen Intervalls $(r, r 2^x)$ das zugehörige $|N(r, f)|$ zu groß, d. h. nicht dem Hilfssatz a genügend. In genau gleicher Weise kann man für den Kreisring $(r, r 2^{-x})$, d. h. rückwärts schließen.

Wenn $M(r, f) < k_1$, so zerlegt man das Intervall $(r, r 2^x)$ in zwei Teilintervalle. Im 1. Intervall sei $M(r, f) \leq k_1$, es kann also unsere Bedingung (1) angewendet werden, und im 2. Intervall die obige Betrachtung.

Speziell folgt aus dieser Betrachtung, daß die Anzahl der Pole im Kreisring $(r 2^x, r 2^{x+1})$ kleiner als cx sein muß, wenn $x \geq g(r)$. Wir benötigen später noch das folgende

Korollar c. Es sei r ein solcher Wert, daß $M(r) > k_1$. Im Kreisring $(r, c_1 r 2^{g(c_1 r)})$ befinden sich höchstens $\alpha c_1 g(r)$ Pole, wobei α und c_1 endliche Konstanten bedeuten.

Beweis. Im Kreisring $(c_1 r, c_1 r 2^{g(c_1 r)})$ befinden sich nach dem Vorhergehenden höchstens $cg(c_1 r)$ Pole. $g(c_1 r)$ läßt sich auf Grund des Korollares b abschätzen. Im Kreisring $(r, c_1 r)$ befinden sich ebenfalls höchstens $c_2 g(r)$ Pole, wobei c_2 eine Konstante bedeutet. Damit ist dieses Korollar bewiesen.

Hilfssatz d. *Es sei r ein beliebiger fester Wert, $g(r)$ seine zugehörige Grenzzahl. Mit β_v bezeichnen wir die Pole unserer Quasi-Ausnahmefunktion. Sie seien nach ihren absoluten Beträgen der Größe nach geordnet. Dann gelten die folgenden Ungleichungen*

$$\sum_{|\beta_v| \geq r 2^{g(r)}} \frac{r}{|\beta_v|} < C, \quad \sum_{|\beta_v| \leq r 2^{-g(r)}} \frac{|\beta_v|}{r} < C.$$

C bedeutet eine endliche Konstante.

Beweis. Im Kreisring $(r 2^x, r 2^{x+1})$ resp. $(r 2^{-x}, r 2^{-(x+1)})$ wobei $x \geq g(r)$, befinden sich nach dem Hilfssatz c höchstens cx Pole. Man findet deshalb

$$\sum_{|\beta_\nu| \geq r 2^{g(r)}} \frac{r}{|\beta_\nu|} < \sum_{x=g(r)}^{\infty} \frac{rcx}{r 2^x} < c \left[1 + \frac{1}{2} + \frac{2}{2^2} + \frac{3}{2^3} + \dots \right] = C,$$

$$\sum_{|\beta_\nu| \leq r 2^{-g(r)}} \frac{|\beta_\nu|}{r} < \sum_{x=g(r)}^{\infty} \frac{rcx}{r 2^x} < c \left[1 + \frac{1}{2} + \frac{2}{2^2} + \frac{3}{2^3} + \dots \right] = C.$$

Spezialfall. Für alle r -Werte mit gleichmäßig beschränktem $g(r)$ sind die Reihen $\sum_{|\beta_\nu| \geq r} \frac{r}{|\beta_\nu|}$ und $\sum_{|\beta_\nu| \leq r} \frac{|\beta_\nu|}{r}$ gleichmäßig beschränkt.

§ 6.

Die Darstellung der Quasi-Ausnahmefunktionen als kanonische Produkte.

Wendet man die Betrachtungen des vorhergehenden Paragraphen speziell für $r = 0$ an, so findet man

$$\sum \frac{1}{|\beta_\nu|^\varepsilon} < K_1 + K_2 \left(\frac{k}{2^{k\varepsilon}} + \frac{k+1}{2^{(k+1)\varepsilon}} + \dots \right).$$

K_1, K_2, k bedeuten endliche positive Konstanten, ε eine beliebig kleine positive Zahl. Diese Reihe ist für einen beliebigen positiven Wert von ε konvergent. Daraus schließen wir unter Anwendung einer bekannten Terminologie den

Satz: *Der Konvergenz-Exponent der Pole einer Quasi-Ausnahmefunktion ist Null.*

Auf Grund der Definition einer Quasi-Ausnahmefunktion und der Invarianz der Konvergenz-Eigenschaft einer Folge gegenüber einer linearen Transformation schließt man sofort, daß durch die Bildung von $\frac{\alpha f(z) + \beta}{\gamma f(z) + \delta}$ aus der Funktion $f(z)$ wieder eine Quasi-Ausnahmefunktion entsteht. Speziell sind also auch $\frac{1}{f(z)}$ und $f(z) - c$, wobei c eine beliebige Konstante bedeutet, Quasi-Ausnahmefunktionen. Deshalb schließen wir auf Grund des obigen Satzes:

Der Konvergenz-Exponent irgendeines Funktionswertes a einer Quasi-Ausnahmefunktion ist Null.

Es sind also insbesondere die Konvergenz-Exponenten von drei Funktionswerten einer Quasi-Ausnahmefunktion gleich Null. Daraus folgt auf Grund eines fundamentalen Satzes aus der Theorie der meromorphen Funktionen⁸⁾ unser

⁸⁾ Vgl. R. Nevanlinna, Zur Theorie der meromorphen Funktionen, Acta mathematica 46 (1925), S. 1—98. — Man vergleiche speziell Satz 1, S. 64.

Hauptsatz 1. *Die Ordnung einer Quasi-Ausnahmefunktion ist gleich Null. Sie läßt sich in folgender Weise darstellen:*

$$f(z) = dz^p \prod \frac{1 - \frac{z}{\alpha_\nu}}{1 - \frac{z}{\beta_\nu}}.$$

p bedeutet eine ganze Zahl, die auch Null sein kann. *d* ist eine endliche Konstante.

Aus der obigen Bemerkung folgt auch, daß unsere Hilfssätze c und d ebenfalls für die Nullstellen gelten, wie man in gleicher Weise wie für die Pole feststellen könnte.

§ 7.

Nachweis der Quasi-Ausnahmeeigenschaft von Quotienten kanonischer Produkte als Folge der Bedingungen (1) und (2).

Es seien nun zwei beliebige Folgen von Zahlen $\alpha_1, \alpha_2 \dots$ und $\beta_1, \beta_2 \dots$ gegeben. Sie sollen die Nullstellen und Pole einer meromorphen Funktion sein und unsern Bedingungen (1) und (2) genügen, unter der weitem Annahme, daß *p* eine ganze positive oder negative Zahl bedeutet. Wir wissen bereits, daß zwei solche Folgen den Konvergenz-Exponenten Null besitzen, daß mithin der mit ihnen in der obigen Weise gebildete Quotient von zwei kanonischen Produkten wirklich in der ganzen *z*-Ebene gleichmäßig konvergent ist. Wir wollen zeigen, daß eine solche Funktion eine Quasi-Ausnahmefunktion in unserm Sinne darstellt.

Zu diesem Zweck genügt es, zu zeigen, daß die Funktionsschar $f(\sigma_n z)$ für jeden Punkt $z \neq 0, \infty$ quasi-normal von endlicher Ordnung ist, wenn σ_n eine wachsende, ins ∞ strebende Folge komplexer Zahlen bedeutet. Wir können auf Grund früherer Ausführungen sogar voraussetzen, daß diese Folge aus lauter reellen Zahlen bestehe, was im nachstehenden der Fall sein soll.

Offenbar genügt es, zu zeigen, daß eine solche Funktionsschar $f(\sigma_n z)$ in einem Kreisring $\frac{1}{a} < |z| < a$ quasi-normal von endlicher Ordnung sei, wobei $a > 2$, aber sonst beliebig vorausgesetzt werden soll. Wir unterscheiden nun zwei Fälle:

Wir bilden die Folge $M\left(\frac{\sigma_n}{2a}, f\right)$. Diese Folge kann beschränkt oder unbeschränkt sein, und diese beiden Fälle wollen wir ins Auge fassen. Wenn die Folge unbeschränkt ist, dann können wir eine Teilfolge so auswählen, daß

$$\lim_{n \rightarrow \infty} M\left(\frac{\sigma_n}{2a}, f\right) = \pm \infty,$$

wobei *n* die Indizes einer Teilfolge durchläuft.

Es sei die obige Folge σ_n bereits unsere Teilfolge und es werde z. B. $+\infty$ vorausgesetzt.

Nun bilden wir unsere Funktionsschar $f(\sigma_n z) = f_n(z)$.

$$f_n(z) = d(\sigma_n z)^p \frac{\prod \left(1 - \frac{\sigma_n z}{\alpha_\nu}\right)}{\prod \left(1 - \frac{\sigma_n z}{\beta_\nu}\right)}.$$

Wir zerlegen eine solche Funktion in vier Faktoren

$$f_n(z) = f_n^{(1)}(z) f_n^{(2)}(z) f_n^{(3)}(z) f_n^{(4)}(z),$$

wobei

$$f_n^{(1)}(z) = \frac{d(\sigma_n z)^p \prod \left(1 - \frac{\sigma_n z}{\alpha_\nu}\right)}{\prod \left(1 - \frac{\sigma_n z}{\beta_\nu}\right)}, \quad \text{wobei } |\alpha_\nu| \leq \frac{\sigma_n}{2a} \quad \text{und} \quad |\beta_\nu| \leq \frac{\sigma_n}{2a} 2^{-g\left(\frac{\sigma_n}{2a}\right)},$$

$$f_n^{(2)}(z) = \prod \left(1 - \frac{\sigma_n z}{\alpha_\nu}\right), \quad \text{,,} \quad \frac{\sigma_n}{2a} < |\alpha_\nu| \leq 2a \sigma_n,$$

$$f_n^{(3)}(z) = \frac{1}{\prod \left(1 - \frac{\sigma_n z}{\beta_\nu}\right)}, \quad \text{,,} \quad \frac{\sigma_n}{2a} 2^{-g\left(\frac{\sigma_n}{2a}\right)} < |\beta_\nu| \leq 2a \sigma_n 2^{g(2a\sigma_n)},$$

$$f_n^{(4)}(z) = \frac{\prod \left(1 - \frac{\sigma_n z}{\alpha_\nu}\right)}{\prod \left(1 - \frac{\sigma_n z}{\beta_\nu}\right)}, \quad \text{,,} \quad |\alpha_\nu| > 2a \sigma_n \quad \text{und} \quad |\beta_\nu| > 2a \sigma_n 2^{g(2a\sigma_n)}.$$

Wir betrachten zunächst die Folge der $f_n^{(4)}(z)$. Infolge unserer Voraussetzungen erhalten wir nachstehende Ungleichungen

$$\left| \frac{\sigma_n z}{\alpha_\nu} \right| < \frac{1}{2}, \quad \left| \frac{\sigma_n z}{\beta_\nu} \right| < \frac{1}{2}.$$

Wegen $|\log(1-z)| = \left| z + \frac{z^2}{2} + \dots \right| \leq 2|z|$ für $|z| \leq \frac{1}{2}$ folgt weiter für eine geeignete Bestimmung des Logarithmus

$$|\log f_n^{(4)}(z)| < \sum_{|\alpha_\nu| > 2a\sigma_n} \frac{2\sigma_n a}{|\alpha_\nu|} + \sum_{|\beta_\nu| > 2a\sigma_n 2^{g(\sigma_n 2a)}} \frac{2\sigma_n a}{|\beta_\nu|}.$$

Wir haben $\lim_{n \rightarrow \infty} M\left(\frac{\sigma_n}{2a}\right) = \infty$ vorausgesetzt. Daraus folgt nach dem Hilfssatz b, daß $\lim_{n \rightarrow \infty} M(2\sigma_n a) = \infty$, indem man ihn für $c = 4a^2$ anwendet.

Nun können wir den Hilfssatz d anwenden (für die erste Summe kommt der Spezialfall des Hilfssatzes d bezüglich der Nullstellen in Betracht) und finden

$$|\log f_n^{(4)}(z)| < 2C \quad \text{oder} \quad e^{-2C} < |f_n^{(4)}(z)| < e^{2C}.$$

Auf Grund eines bekannten Satzes von Montel betr. die normalen Funktionenscharen, besitzt daher die Folge der $f_n^{(4)}(z)$ eine in unserm Kreisring $\frac{1}{a} < |z| < a$ gleichmäßig konvergente Teilfolge mit einer holomorphen, im Kreisring nirgends verschwindenden Grenzfunktion. In den folgenden Betrachtungen sollen auch für die Folgen $f_n^{(1)}(z)$, $f_n^{(2)}(z)$, $f_n^{(3)}(z)$ nur noch die Funktionen mit den Indizes der soeben ausgewählten Teilfolge berücksichtigt werden.

Nun betrachten wir die Folge der $f_n^{(2)}(z)$. Infolge unserer Voraussetzung über die Folge $M\left(\frac{\sigma_n}{2a}\right)$ besitzen die Funktionen im Kreisring $\frac{\sigma_n}{2a} \leq |z| \leq 2a\sigma_n$ eine beschränkte Anzahl von Nullstellen gemäß unserer Bedingung (1). Es handelt sich also um eine Folge von Polynomen mit beschränktem Grade von der Form $\prod\left(1 - \frac{z}{\alpha_\nu}\right)$, die in unserm Kreisring gleichmäßig beschränkt sind. Deshalb kann man eine gleichmäßig beschränkte Teilfolge auswählen. Die Grenzfunktion $h(z)$ dieser Teilfolge ist in unserm Kreisring $\frac{1}{a} < |z| < a$ gleichmäßig beschränkt, verschwindet nicht identisch und besitzt eine endliche Anzahl von Nullstellen.

Für die Betrachtung der Folge $f_n^{(1)}(z)$ zerlegen wir jede Funktion in ein Produkt von zwei Faktoren, indem wir setzen $f_n^{(1)}(z) = \psi_n(z)\varphi_n(z)$, wobei

$$\psi_n(z) = \frac{d(\sigma_n z)^p \prod_{|\alpha_\nu| \leq \frac{\sigma_n}{2a}} \frac{\sigma_n z}{\alpha_\nu}}{\prod_{|\beta_\nu| \leq \frac{\sigma_n}{2a}} \frac{\sigma_n z}{\beta_\nu}} \prod_{\frac{\sigma_n}{2a} \leq \left(\frac{\sigma_n}{2a}\right)^2 < |\beta_\nu| \leq \frac{\sigma_n}{2a}} \frac{\sigma_n z}{\beta_\nu},$$

$$\varphi_n(z) = \pm \frac{\prod_{|\alpha_\nu| \leq \frac{\sigma_n}{2a}} \left(1 - \frac{\alpha_\nu}{\sigma_n z}\right)}{\prod_{|\beta_\nu| \leq \left(\frac{\sigma_n}{2a}\right)^2} \left(1 - \frac{\beta_\nu}{\sigma_n z}\right)}.$$

Für die Folge der $\varphi_n(z)$ kann man genau gleich schließen, wie für die Folge $f_n^{(4)}(z)$. Man findet

$$\log |\varphi_n(z)| < \sum_{|\alpha_\nu| \leq \frac{\sigma_n}{2a}} \frac{2a|\alpha_\nu|}{\sigma_n} + \sum_{|\beta_\nu| \leq \left(\frac{\sigma_n}{2a}\right)^2} \frac{2|\beta_\nu|a}{\sigma_n} < 2C,$$

$$e^{-2C} < |\varphi_n(z)| < e^{2C}.$$

Man kann wieder auf Grund des schon angewendeten Auswahlssatzes von Montel eine in unserm Kreisring $\frac{1}{a} < |z| < a$ gleichmäßig konvergente Teilfolge auswählen, deren Grenzfunktion in diesem Kreisring nirgends verschwindet und sich holomorph verhält.

Es bleibt noch die Folge $f_n^{(3)}(z)\psi_n(z)$ zu diskutieren. Wir erhalten

$$\begin{aligned} \log |f_n^{(3)}(z)\psi_n(z)| &= N\left(\frac{\sigma_n}{2a}\right) \log |2az| + M\left(\frac{\sigma_n}{2a}\right) \\ &- \log \prod_{\substack{\frac{\sigma_n}{2a} < |\beta_\nu| \leq \frac{\sigma_n}{2a} \\ -g\left(\frac{\sigma_n}{2a}\right)}} \left|1 - \frac{\beta_\nu}{\sigma_n z}\right| - \log \prod_{\substack{\frac{\sigma_n}{2a} < |\beta_\nu| < 2a\sigma_n}} \left|1 - \frac{\sigma_n z}{\beta_\nu}\right| \\ &- \log \prod_{2a\sigma_n \leq |\beta_\nu| \leq 2a\sigma_n 2^{g(2a\sigma_n)}} \left|1 - \frac{\sigma_n z}{\beta_\nu}\right|. \end{aligned}$$

Jeder Faktor dieser Produkte ist $\leq 1 + 2a^2$. Die Anzahl dieser Faktoren ist auf Grund des Hilfssatzes c und seines Korollares $< Kg\left(\frac{\sigma_n}{2a}\right)$, wobei K eine endliche, positive Konstante bedeutet. Deshalb finden wir unter Anwendung des Hilfssatzes a

$$\log |f_n^{(3)}(z)\psi_n(z)| > -k_2 g\left(\frac{\sigma_n}{2a}\right) \log(2a^2) + M\left(\frac{\sigma_n}{2a}\right) - Kg\left(\frac{\sigma_n}{2a}\right) \log(1 + 2a^2).$$

Berücksichtigt man die Definition von $g(r)$ und unsere Voraussetzung über die Folge $M\left(\frac{\sigma_n}{2a}\right)$, so findet man, daß diese Folge in unserm Kreisring gleichmäßig gegen ∞ strebt.

Faßt man diese Ergebnisse zusammen, so gelangt man zu folgendem Resultat: Es ist möglich, aus unserer Folge eine im ganzen Kreisring, abgesehen von endlich vielen irregulären Punkten, gleichmäßig konvergente Teilfolge herauszugreifen. Diese irregulären Punkte sind die endlich vielen Nullstellen der Funktion $h(z)$. Sie sind von endlicher Ordnung, da die Funktionen $f_n(z)$ der konvergenten Folge in der Umgebung eines solchen irregulären Punktes eine beschränkte Anzahl von Nullstellen besitzen und die Grenzfunktion identisch ∞ beträgt.

Wir haben bei diesem Beweis angenommen, daß $\lim_{n \rightarrow \infty} M\left(\frac{\sigma_n}{2a}\right) = \pm \infty$.

Läßt sich eine beschränkte Teilfolge aus dieser Folge herausgreifen, so kann man direkt die Methode von Ostrowski anwenden. Man zerlegt in folgender Weise:

$$f(z) = f_n^{(1)}(z) f_n^{(2)}(z) f_n^{(3)}(z),$$

wobei

$$f_n^{(1)}(z) = d(\sigma_n z)^p \frac{\prod_{|\alpha_\nu| \leq \frac{\sigma_n}{2a}} \left(1 - \frac{\sigma_n z}{\alpha_\nu}\right)}{\prod_{|\beta_\nu| \leq \frac{\sigma_n}{2a}} \left(1 - \frac{\sigma_n z}{\beta_\nu}\right)},$$

$$f_n^{(2)}(z) = \frac{\prod_{\frac{\sigma_n}{2a} < |\alpha_\nu| \leq 2a\sigma_n} \left(1 - \frac{\sigma_n z}{\alpha_\nu}\right)}{\prod_{\frac{\sigma_n}{2a} < |\beta_\nu| \leq 2a\sigma_n} \left(1 - \frac{\sigma_n z}{\beta_\nu}\right)},$$

$$f_n^{(3)}(z) = \frac{\prod_{|\alpha_\nu| > 2a\sigma_n} \left(1 - \frac{\sigma_n z}{\alpha_\nu}\right)}{\prod_{|\beta_\nu| > 2a\sigma_n} \left(1 - \frac{\sigma_n z}{\beta_\nu}\right)}.$$

Die Folgen $f_n^{(1)}(z)$ und $f_n^{(3)}(z)$ lassen sich wie unsere vorhergehende Folge $f_n^{(4)}(z)$ behandeln.

Die Funktionen $f_n^{(2)}(z)$ sind dank unserer Voraussetzung gemäß Bedingung (1) rationale Funktionen, deren Zähler und Nenner Polynome mit beschränktem Grade darstellen. Es läßt sich deshalb aus dieser Folge eine im Kreisring $\frac{1}{a} < |z| < a$, abgesehen von endlich vielen irregulären Punkten, gleichmäßig konvergente Teilfolge mit einer Grenzfunktion $\not\equiv 0, \infty$ herausgreifen. Die irregulären Punkte dieser konvergenten Teilfolge können dadurch entstehen, daß gewisse Nullpunkte und Pole unserer Funktionen gegen denselben Punkt konvergieren. Gerade um diese Eventualität zu vermeiden, müssen die Argumente der Pole und Nullstellen einer J -Ausnahmefunktion nach Ostrowski folgende Bedingung erfüllen:

Es gibt ein nur von $f(z)$ abhängiges $c_4 > 0$, so daß für alle $\nu > 0$, $\mu > 0$

$$\left| \frac{b_\nu}{a_\mu} - 1 \right| > c_4 \quad \text{ist.}$$

Unsere Bemerkungen zeigen also, daß solche Funktionen, welche nur die drei Bedingungen von Ostrowski, betreffend die absoluten Beträge der Nullstellen und Pole erfüllen, eine spezielle Klasse der Quasi-Ausnahmefunktionen bilden.

Es ist klar, daß die so entstehenden irregulären Punkte für unsere Folge $f_n(z) = f(\sigma_n z)$ von endlicher Ordnung sind, da sie in der Umgebung eines solchen eine beschränkte Anzahl von Nullstellen und Polen besitzen.

§ 8.

Die beiden Hauptsätze.

Zunächst noch die folgende Bemerkung: Bildet man die Funktionsschar $f(\sigma_n z)$, wo σ_n irgendeine ins ∞ strebende Folge komplexer Zahlen bedeutet, $f(z)$ eine Quasi-Ausnahmefunktion und $z \neq 0, \infty$, so ist sie in der Umgebung eines solchen Punktes quasi-normal endlicher Ordnung. Bildet man dieselbe Schar für $z = 0$, so wird sie quasi-normal, aber im allgemeinen nicht mehr von endlicher Ordnung sein. Verlangt man, daß auch der Punkt $z = 0$ nur ein irregulärer Punkt endlicher Ordnung sei, dann muß die Funktion eine rationale Funktion sein. Denn auf Grund der Definition eines irregulären Punktes endlicher Ordnung einer konvergenten Folge schließen wir, daß die Funktion $f(z)$ in der ganzen z -Ebene jeden Funktionswert mit eventueller Ausnahme eines einzigen in einer beschränkten Anzahl von Punkten annehmen darf. Daraus folgt unsere Behauptung.

Nach genau der gleichen Methode wie bei der Bestimmung der J -Ausnahmefunktion⁹⁾ kann man allgemeiner jene Quasi-Ausnahmefunktionen bestimmen, die sich meromorph und eindeutig für $|z| \geq R$ verhalten. Wir erhalten deshalb unsern folgenden

Hauptsatz 1. Jede für $|z| \geq R$ meromorphe Quasi-Ausnahmefunktion ist gleich dem Produkt einer in der ganzen Ebene meromorphen Quasi-Ausnahmefunktion mit einer für $|z| > R$ regulären und für $z = \infty$ den Wert 1 annehmenden Funktion. Umgekehrt ist jedes solche Produkt eine Quasi-Ausnahmefunktion.

Jede in der ganzen z -Ebene meromorphe Quasi-Ausnahmefunktion ist gleich einem Quotienten von zwei ganzen Funktionen der Ordnung 0

$$f(z) = dz^p \prod \frac{\left(1 - \frac{z}{\alpha_\nu}\right)}{\left(1 - \frac{z}{\beta_\nu}\right)} \quad (p \text{ ganz } \geq 0, d \neq 0),$$

deren Nullstellen α_ν und Pole β_ν den folgenden zwei Bedingungen genügen:

a) Setzt man

$$M(r) = r^p \frac{\prod_{|\alpha_\nu| \leq r} \frac{r}{|\alpha_\nu|}}{\prod_{|\beta_\nu| \leq r} \frac{r}{|\beta_\nu|}},$$

so existieren zwei nur von $f(z)$ abhängige positive Konstanten k_1 und k_2 mit folgender Eigenschaft: Gilt für einen r -Wert

⁹⁾ Vgl. O., S. 256—257.

$$M(r) \leq k_1 \quad \text{bzw.} \quad M(r) \geq \frac{1}{k_1},$$

so beträgt die Anzahl der Pole bzw. der Nullstellen im Kreisring $r \leq |z| \leq 2r$ höchstens k_2 .

b) Es sei $N(r)$ die Differenz zwischen der Anzahl der Nullstellen und der Pole im Kreise $|z| \leq r$. Dann existiert eine nur von $f(z)$ abhängige positive Konstante k_3 mit folgender Eigenschaft: Wenn für einen r -Wert

$$\frac{1}{k_1} \leq M(r) \leq k_1,$$

so gilt gleichzeitig

$$|N(r)| \leq k_3.$$

Beispiele.
$$f(z) = \frac{\prod_{n=1}^{\infty} \left(1 + \frac{z e^{\frac{2\pi i}{n}}}{2^n}\right)}{\prod_{n=1}^{\infty} \left(1 + \frac{z e^{-\frac{2\pi i}{n}}}{2^n}\right)}, \quad f(z) = \frac{\prod_{n=1}^{\infty} \left(1 - \frac{z}{2^n}\right)}{\prod_{n=1}^{\infty} \left(1 + \frac{z}{2^{n^2}}\right)}.$$

Aus diesem Satz ergibt sich der folgende

Spezialfall. Jede ganze Quasi-Ausnahmefunktion ist von der Ordnung Null und besitzt in einem Kreisring von beschränkter logarithmischer Breite eine beschränkte Anzahl von Nullstellen. Umgekehrt ist jede solche Funktion eine Quasi-Ausnahmefunktion.

Beispiel.
$$g(z) = \prod_{n=1}^{\infty} \left(1 - \frac{z}{2^n}\right).$$

Aus dem Montelschen Satz¹⁰⁾, daß eine Funktionsschar von in einem bestimmten Gebiet G meromorphen Funktionen, welche in G drei feste Funktionswerte in einer beschränkten Anzahl annehmen, eine quasi-normale Schar endlicher Ordnung bilden, folgt unser

Hauptsatz 2. Jede für $|z| > R$ meromorphe Funktion, die keine Quasi-Ausnahmefunktion darstellt, besitzt für jeden beliebig kleinen Winkel φ mindestens eine Folge von Kreisen, die man vom Nullpunkt aus unter diesem Winkel sieht, in denen die Funktion jeden Funktionswert, ausgenommen höchstens zwei feste Ausnahmewerte, unbeschränkt oft annimmt.

Göttingen, den 30. August 1927.

¹⁰⁾ Vgl. loc. cit. 3 b, S. 21, Kap. III.

Nachtrag bei der 1. Korrektur. Inzwischen ist das Buch von Herrn Montel über normale Funktionsscharen erschienen, das insbesondere einen Teil der Resultate von Herrn Ostrowski wiedergibt. Man vgl. P. Montel: *Leçons sur les familles normales de fonctions analytiques et leurs applications*, Paris 1927, Verlag Gauthier-Villars. Vor wenigen Tagen hat mir Herr Valiron Korrekturbogen von zwei Arbeiten gesandt, die dieses Jahr in den *Acta Mathematica* und im *Journal de Mathématiques* erscheinen sollen und die sich ebenfalls mit der Verschärfung der Julia-Ostrowskischen Sätze befassen. Man vergleiche: G. Valiron, *Recherches sur le théorème de M. Borel dans la théorie des fonctions méromorphes*, *Acta Mathematica* 52, 1928 und G. Valiron, *Le théorème de M. Picard et le complément de M. Julia*, *Journal de Mathématiques* 7, 1928. Der Satz III in der Acta-Arbeit kreuzt sich mit meinem Hauptsatz II.

Zürich, den 2. April 1928.

Geodesics in non-holonomic geometry.

Von

J. L. Synge in Dublin (Ireland).

Synopsis:

1. Introduction.
2. The constraint.
3. Equations of constrained geodesics derived from a variational principle.
4. Second form of the equations of constrained geodesics.
5. Parallel (Γ) propagation.
6. Geodesic stability for a general correspondence.
7. Isometric and normal correspondences.
8. Dynamical significance of the paper.

1. Introduction.

The stability of geodesics in Riemannian space was first discussed simultaneously and independently by Levi-Civita¹⁾ and myself²⁾. The tensorial equation obtained may be written

$$(1.1) \quad \bar{\eta}^r + G^r_{.msn} x^{m'} \eta^s x^{n'} = 0,$$

where η^r is the infinitesimal vector joining a point on the fundamental geodesic to the corresponding point of a neighbouring geodesic, $x^{r'} = \frac{dx^r}{ds}$, where x^r is the coordinate system, $G^r_{.msn}$ is the mixed curvature tensor

¹⁾ „Sur l'écart géodésique,“ *Math. Annalen* 97 (1926), p. 291—320. Cf. also Levi-Civita, „The Absolute Differential Calculus,“ English Translation (1927).

²⁾ „On the Geometry of Dynamics,“ *Phil. Trans. Roy. Soc., A*, 226 (1926), p. 31—106. That paper will be referred to as GD. The dynamical problem of „Stability in the action sense,“ discussed in GD., Chap. IX, is precisely the geometrical problem of the stability of geodesics in Riemannian space.

of the manifold, and the operation denoted by superimposed bars is defined (for any vector given along a curve) by

$$(1.2) \quad \begin{cases} \bar{X}^r = \frac{dX^r}{ds} + \left\{ \begin{matrix} r \\ mn \end{matrix} \right\} X^m \frac{dx^n}{ds}, \\ \bar{\bar{X}}^r = \frac{d\bar{X}^r}{ds} + \left\{ \begin{matrix} r \\ mn \end{matrix} \right\} \bar{X}^m \frac{dx^n}{ds}. \end{cases}$$

The question of stability for systems subject to non-holonomic constraints was not discussed in these papers. Subsequently a number of papers have appeared dealing with geodesic stability and non-holonomic constraints³⁾.

The method of Vranceanu, in dealing with non-holonomic systems, differs essentially from mine, as developed in G.D. and in the present paper. He has employed a set of vectors *in* the element in which motion subject to the constraint is possible, whereas I have employed a set of vectors *normal* to that element. We have therefore been led to different analytical expressions for the equations of motion of a non-holonomic conservative system (or, geometrically, the differential equations of constrained geodesics), his equations consisting⁴⁾ for a system with N coordinates and M constraints, of $(2N - M)$ equations of the first order for the N coordinates and the $(N - M)$ components of the unit tangent vector along his fundamental vectors of constraint, whereas my equations⁵⁾ consist of a set of N equations of the second order for the N coordinates, M particular first integrals of these equations being known, viz., the equations of constraint. The expressions given by him are invariant with respect to coordinate transformations and tensorial with respect to transformations of the constraint vectors, while the expressions which I shall develop are tensorial with respect to coordinate transformations and invariant with respect to transformations of the constraint vectors.

In the paper last cited, Vranceanu discusses the stability of geodesics. It appears to me, however, that his argument is incorrect, for he assumes that the infinitesimal displacement from the fundamental geodesic to the

³⁾ G. Vranceanu, „Sur les espaces non holonomes,“ *Comptes Rendus* 183 (1926), p. 852—854; „Sur le calcul différentiel absolu pour les variétés non holonomes,“ *ibid.*, p. 1083—1085; „Sopra una classe di sistemi anolonomi,“ *Rend. Acc. Lincei* (6) 3 (1926), p. 548—553; „Sopra le equazioni del motò di un sistema anolonomo,“ *ibid.* 4 (1926), p. 508—511; „Sopra la stabilità geodetica,“ *ibid.* 5 (1927), p. 107—110. T. Boggio, „Sullo scostamento geodetico,“ *Rend. Acc. Lincei* (6) 4 (1926), p. 255—261. U. Crudeli, „Su lo scostamento geodetico elementare; procedimento di estensione della equazione di Jacobi ad una qualsiasi varietà riemanniana,“ *Rend. Acc. Lincei* (6) 5 (1927), p. 248—251. E. Cartan, „Sur l'écart géodésique et quelques notions connexes,“ *Rend. Acc. Lincei* (6) 5 (1927), p. 609—613.

⁴⁾ „Sopra la stabilità geodetica,“ p. 107.

⁵⁾ G.D., p. 58, Theorem XVII (A).

neighbouring geodesic is consistent with the constraints. In general, this cannot continue to be true along the geodesic⁶⁾.

In the present paper the equations of constrained geodesics are derived from a variational principle, and equations for geodesic separation are obtained. The mode of development is geometrical, the dynamical significance being indicated in § 8. The line-element is assumed to be positive definite.

2. The constraint.

Let N be the number of dimensions of the manifold and M the number of constraints.

The following conventions will be observed:

1. Small italic indices un-repeated imply a range of values from 1 to N , and, when repeated in a single term, summation over that range.

2. Greek indices un-repeated imply a range of values from 1 to M , and, when repeated⁷⁾ in a single term, summation over that range.

3. Those indices (other than powers), which do not imply tensorial character with respect to coordinate transformations, are enclosed in brackets⁸⁾.

4. The Christoffel symbol of the second kind is denoted by $\left\{ \begin{smallmatrix} r \\ mn \end{smallmatrix} \right\}$, instead of $\left\{ \begin{smallmatrix} mn \\ r \end{smallmatrix} \right\}$ as in GD.⁹⁾

Let x^r be the coordinate system, and let the line-element be given by

$$(2.1) \quad ds^2 = g_{mn} dx^m dx^n,$$

which form we suppose to be positive definite. Let the constraint be given by

$$(2.2) \quad A_{(\varrho)m} dx^m = 0,$$

which we suppose, in general, to possess no integrable combination. Equations (2.2) can always be replaced by

$$(2.3) \quad B_{(\varrho)m} dx^m = 0,$$

where $B_{(\varrho)}^r$ are M mutually orthogonal unit vectors¹⁰⁾. These vectors will be called the *constraint vectors* and the linear M -space defined by them the *M -space of constraint*. Any curve satisfying (2.3) will be called a

⁶⁾ Cf. Cartan, loc. cit.

⁷⁾ It is not necessary that one should occur as a superscript and the other as a subscript.

⁸⁾ This rule is not observed in the case of Γ_{mn}^r defined in (4.31).

⁹⁾ The present notation is used by Eisenhart, *Riemannian Geometry* (1926).

¹⁰⁾ GD., p. 54.

constrained curve. A curve is a constrained curve if and only if it is perpendicular at every point to every vector lying in the M -space of constraint.

It is of course well known that the non-integrable character of the equations (2.3) implies that the congruences $B_{(\varrho)}^r$ are not normal congruences; the fact that there exists no integrable combination implies that there is no normal congruence having at every point a direction contained in the M -space of constraint. If, on the other hand, there did exist an integrable combination, then the totality of curves passing through any assigned point of the manifold and satisfying that equation at every point would build up a subspace of $(N-1)$ dimensions, from which it would be impossible to depart if one always followed a constrained curve. But under the given conditions of non-holonomicity it will be possible in general to reach any point of the manifold by travelling along a constrained curve from an arbitrarily assigned point.

It is possible to write down a single function whose vanishing is necessary and sufficient for the satisfaction of the constraint by a given curve. Let λ^r denote the unit vector tangent to the curve, and let ω be defined as the positive quantity given by

$$(2.4) \quad \omega^2 = B_{(\varrho)m} B_{(\varrho)n} \lambda^m \lambda^n = (B_{(1)m} \lambda^m)^2 + (B_{(2)m} \lambda^m)^2 + \dots + (B_{(M)m} \lambda^m)^2.$$

Thus ω vanishes if and only if all the equations of constraint are satisfied. It is not difficult to see the geometrical meaning of ω . Let B^r denote any unit vector in the M -space of constraint, so that

$$(2.41) \quad B^r = \theta_{(\varrho)} B_{(\varrho)}^r, \quad \theta_{(\varrho)} \theta_{(\varrho)} = 1,$$

and let φ denote the angle between λ^r and B^r . Then

$$(2.42) \quad \cos \varphi = B_r \lambda^r = \theta_{(\varrho)} B_{(\varrho)r} \lambda^r.$$

If, by variation of $\theta_{(\varrho)}$, we vary the direction of B^r , we have

$$(2.43) \quad \delta \cos \varphi = B_{(\varrho)r} \lambda^r \delta \theta_{(\varrho)},$$

$$(2.44) \quad \theta_{(\varrho)} \delta \theta_{(\varrho)} = 0;$$

if $\cos \varphi$ has a stationary value for all such variations, we must have

$$(2.5) \quad B_{(\varrho)r} \lambda^r = k \theta_{(\varrho)},$$

where k is undetermined. Multiplying by $\theta_{(\varrho)}$ and summing as indicated, we find, by (2.42),

$$(2.6) \quad k = \cos \varphi.$$

If we square (2.5) and sum for ϱ from 1 to M , we find

$$(2.7) \quad \omega^2 = k^2 = \cos^2 \varphi.$$

Thus the stationary values of $\cos \varphi$ are $\pm \omega$, the two signs corresponding to opposed unit vectors B^r in the M -space of constraint. In fact, ω is the cosine of the angle which the vector λ^r makes with the M -space of constraint, this angle being defined as the smallest angle made by λ^r with any vector in the M -space of constraint. We shall call ω the *constraint function*. Its importance is due to the fact that its value is independent of the choice of orthogonal unit vectors $B_{(e)}^r$ in the M -space of constraint.

If we write

$$(2.8) \quad C_{mn} = B_{(e)m} B_{(e)n}$$

then C_{mn} is a symmetrical covariant tensor of the second order with respect to coordinate transformations, and is independent of the choice of the constraint vectors. We shall call C_{mn} the *constraint tensor*. Any vector λ^r , which satisfies the constraints, satisfies

$$(2.9) \quad C_{mn} \lambda^n = 0.$$

3. Equations of constrained geodesics derived from a variational principle.

We shall now define a *constrained geodesic*¹¹⁾ as a constrained curve whose length is stationary for all variations which vanish at the end points and satisfy the equations of constraint.

It is to be noted that it is the *variation*, and not the *varied curve*, that satisfies the equations of constraint. This definition is adopted because constrained geodesics, so defined, have a dynamical significance. From a purely geometrical point of view there are two possible generalisations of the concept of surface geodesics when the constraint is non-holonomic, viz., curves defined as above, and curves of stationary length with respect to variations which make the varied curve satisfy the equations of constraint. We shall reserve the name *constrained geodesics* for curves defined in the former sense.

For any variation δx^r with fixed end points we have¹²⁾

$$(3.1) \quad \delta \int ds = - \int g_{mn} \kappa^m \delta x^n ds,$$

where κ^r is the first curvature vector,

$$(3.11) \quad \kappa^r = \frac{d^2 x^r}{ds^2} + \left\{ \begin{matrix} r \\ mn \end{matrix} \right\} \frac{dx^m}{ds} \frac{dx^n}{ds}.$$

¹¹⁾ Geodesics in non-holonomic geometry have been defined by Vranceanu, *Comptes Rendus* 183 (1926), p. 854, but he has not connected his definition with a variational principle. However it is not difficult to derive his equations from the variational principle here adopted. The curves defined by him are therefore identical with those of the present paper, but the notation is entirely different.

¹²⁾ Proc. London Math. Soc. (2) 25 (1926), p. 253.

In order that a constrained curve may be geodesic, it is necessary and sufficient that $\delta \int ds$ should vanish for all variations consistent with

$$(3.12) \quad B_{(\varrho)n} \delta x^n = 0,$$

and vanishing at the end points. Hence we must have

$$(3.2) \quad \kappa_n \delta x^n \equiv k_{(\varrho)} B_{(\varrho)n} \delta x^n,$$

where $k_{(\varrho)}$ are undetermined. Hence

$$(3.21) \quad \kappa_r = k_{(\varrho)} B_{(\varrho)r}, \quad \kappa^r = k_{(\varrho)} B_{(\varrho)}^r.$$

Now by differentiation of the equations of constraint

$$(3.3) \quad B_{(\varrho)m} \frac{dx^m}{ds} = 0,$$

which are satisfied along the curve, we obtain

$$(3.31) \quad B_{(\varrho)m} \kappa^m + B_{(\varrho)mn} \frac{dx^m}{ds} \frac{dx^n}{ds} = 0,$$

where $B_{(\varrho)mn}$ is the covariant derivative of $B_{(\varrho)m}$. Hence, substituting for κ^m from (3.21), and remembering that the constraint vectors, being unit vectors and mutually orthogonal, satisfy

$$(3.32) \quad B_{(\varrho)m} B_{(\varrho)}^m = \delta_{\varrho}^{\sigma} = \begin{cases} 1, & \text{for } \sigma = \varrho, \\ 0, & \text{for } \sigma \neq \varrho, \end{cases}$$

we find

$$(3.33) \quad k_{(\varrho)} = -B_{(\varrho)mn} \frac{dx^m}{ds} \frac{dx^n}{ds}.$$

Substitution in (3.21) gives

$$(3.4) \quad \kappa^r = -B_{(\varrho)}^r B_{(\varrho)mn} \frac{dx^m}{ds} \frac{dx^n}{ds}.$$

We may state the following result:

Theorem I. *Every constrained geodesic satisfies the equations*

$$(3.5) \quad \frac{d^2 x^r}{ds^2} + \left\{ \begin{matrix} r \\ m n \end{matrix} \right\} + B_{(\varrho)}^r B_{(\varrho)mn} \frac{dx^m}{ds} \frac{dx^n}{ds} = 0.$$

Geometrically, the first curvature vector of a constrained geodesic lies in the M-space of constraint, and has components $-B_{(\varrho)mn} \frac{dx^m}{ds} \frac{dx^n}{ds}$ in the directions of the vectors $B_{(\varrho)}^r$ respectively.

We shall call (3.5) the *first form* of the equations of constrained geodesics.

Along any curve we have

$$(3.51) \quad \frac{d}{ds} \left(B_{(\varrho)m} \frac{dx^m}{ds} \right) = B_{(\varrho)m} \kappa^m + B_{(\varrho)mn} \frac{dx^m}{ds} \frac{dx^n}{ds},$$

which vanishes if the curve satisfies (3.5). Thus we have the result:

Theorem II. *The differential equations (3.5) have the first integrals*

$$(3.52) \quad B_{(\varrho)m} \frac{dx^m}{ds} = \text{constant}.$$

The differential equations (3.5) define a unique curve passing through an arbitrarily assigned point in an arbitrarily assigned direction. From the theorem just established we see that all solutions of (3.5), whose initial directions satisfy the equations of constraint, are constrained geodesics.

It may be remarked that the equations

$$(3.6) \quad \frac{d^2 x^r}{du^2} + \left\{ \begin{matrix} r \\ mn \end{matrix} \right\} + B_{(\varrho)}^r B_{(\varrho)mn} \frac{dx^m}{du} \frac{dx^n}{du} = 0$$

do not in general possess the first integral

$$(3.61) \quad g_{mn} \frac{dx^m}{du} \frac{dx^n}{du} = \text{constant}.$$

For we find

$$(3.62) \quad \frac{d}{du} \left(g_{rs} \frac{dx^r}{du} \frac{dx^s}{du} \right) = 2g_{rs} \left(\frac{d^2 x^r}{du^2} + \left\{ \begin{matrix} r \\ mn \end{matrix} \right\} \frac{dx^m}{du} \frac{dx^n}{du} \right) \frac{dx^s}{du},$$

so that for a curve satisfying (3.6) we have

$$(3.63) \quad \frac{d}{du} \left(g_{rs} \frac{dx^r}{du} \frac{dx^s}{du} \right) = -2B_{(\varrho)s} \frac{dx^s}{du} B_{(\varrho)m n} \frac{dx^m}{du} \frac{dx^n}{du}.$$

This will vanish, giving the first integral (3.61), if the equations of constraint are satisfied, but not in general.

4. Second form of the equations of constrained geodesics.

The equations (3.5) for constrained geodesics suffer from the defect that they are not independent of the choice of the orthogonal unit constraint vectors in the M -space of constraint. We shall now obtain equations which have the requisite independence.

Since all along any constrained curve we have

$$(4.1) \quad B_{(\varrho)m} \frac{dx^m}{ds} = 0,$$

it follows that a constrained geodesic (which satisfies (3.5)) also satisfies

$$(4.2) \quad \frac{d^2 x^r}{ds^2} + \left\{ \begin{matrix} r \\ mn \end{matrix} \right\} + B_{(\varrho)}^r B_{(\varrho)mn} + B_{(\varrho).n}^r B_{(\varrho)m} \frac{dx^m}{ds} \frac{dx^n}{ds} = 0,$$

where

$$(4.21) \quad B_{(\varrho).n}^r = g^{rs} B_{(\varrho)sn}.$$

If we take the covariant derivative of the constraint tensor C_{mn} defined by (2.8) and then raise the first subscript, we have (changing the indices)

$$(4.25) \quad C_{.mn}^r = B_{(\varrho)}^r B_{(\varrho)mn} + B_{(\varrho).n}^r B_{(\varrho)m}.$$

Thus we have the result:

Theorem III. *Every constrained geodesic satisfies*

$$(4.3) \quad \frac{d^2 x^r}{ds^2} + \Gamma^r_{.mn} \frac{dx^m}{ds} \frac{dx^n}{ds} = 0,$$

where

$$(4.31) \quad \Gamma^r_{.mn} = \left\{ \begin{matrix} r \\ mn \end{matrix} \right\} + C^r_{.mn}.$$

Since the constraint tensor is independent of the particular choice of constraint vectors, it follows that we have in (4.3) a form having the requisite independence. We shall call (4.3) the *second form* of the equations of constrained geodesics.

Unlike the first form (3.5), the second form (4.3) does not possess the first integrals (3.52). We find, for any curve satisfying (4.3),

$$(4.4) \quad \begin{aligned} \frac{d}{ds} \left(B_{(\varrho)r} \frac{dx^r}{ds} \right) &= -B_{(\varrho)r} C^r_{.mn} \frac{dx^m}{ds} \frac{dx^n}{ds} + B_{(\varrho)rn} \frac{dx^r}{ds} \frac{dx^n}{ds}, \\ &= [B_{(\varrho)mn} - B_{(\varrho)r} (B^r_{(\sigma)} B_{(\sigma)mn} + B^r_{(\sigma).n} B_{(\sigma)m})] \frac{dx^m}{ds} \frac{dx^n}{ds}, \\ &= -B_{(\varrho)r} B^r_{(\sigma).n} B_{(\sigma)m} \frac{dx^m}{ds} \frac{dx^n}{ds}. \end{aligned}$$

This will not vanish in general, but will vanish at a point if the equations of constraint are satisfied at that point. We may state the result:

Theorem IV. *The differential equations (4.3) possess the particular first integrals*

$$(4.5) \quad B_{(\varrho)m} \frac{dx^m}{ds} = 0.$$

Constrained geodesics are a particular class of curves belonging to the more general system defined by

$$(4.6) \quad \frac{d^2 x^r}{du^2} + \Gamma^r_{.mn} \frac{dx^m}{du} \frac{dx^n}{du} = 0.$$

I have not discovered the geometrical meaning of this system.

5. Parallel (Γ) propagation.

The quantities $\Gamma^r_{.mn}$ are not the components of a tensor; if we transform from the coordinate system x^r to the coordinate system $'x^r$, the formulae of transformation are easily seen to be

$$(5.1) \quad ' \Gamma^r_{.mn} = \Gamma^c_{.ab} \frac{\partial 'x^r}{\partial x^c} \frac{\partial x^a}{\partial 'x^m} \frac{\partial x^b}{\partial 'x^n} + \frac{\partial^2 x^c}{\partial 'x^m \partial 'x^n} \frac{\partial 'x^r}{\partial x^c}.$$

Thus, if a contravariant vector X^r is given along any curve, the quantities

$$(5.2) \quad \tilde{X}^r = \frac{dX^r}{ds} + \Gamma^r_{.mn} X^m \frac{dx^n}{ds}$$

are the components of a contravariant vector. The coefficients $\Gamma^r_{.mn}$ are

competent to define a type of parallel propagation¹³), which we shall call *parallel (Γ) propagation*, defined by the equations

$$(5.21) \quad \tilde{X}^r = 0.$$

This is to be distinguished from the ordinary parallel propagation,

$$(5.22) \quad \bar{X}^r \equiv \frac{dX^r}{ds} + \left\{ \begin{matrix} r \\ mn \end{matrix} \right\} X^m \frac{dx^n}{ds} = 0.$$

We have, in fact,

$$(5.23) \quad \tilde{X}^r = \bar{X}^r + C^r_{.mn} X^m \frac{dx^n}{ds}.$$

It is obvious that the unit tangent vector to a constrained geodesic undergoes parallel (Γ) propagation. If λ^r denotes the unit tangent vector, the equations of a constrained geodesic may be written in either of the forms

$$(5.3) \quad \tilde{\lambda}^r = 0,$$

$$(5.31) \quad \bar{\lambda}^r = -C^r_{.mn} \lambda^m \lambda^n.$$

An interesting feature of the present work is the natural appearance in a discussion immediately suggested by classical dynamical theory of a type of geometry whose physical interpretation is usually connected with electromagnetic theory in the general theory of relativity. Equations (4.3), however, are not as general as the paths of Eisenhart and Veblen¹⁴), because our $\Gamma^r_{.mn}$ are defined by (4.31), in which $C^r_{.mn}$ is not an arbitrary mixed tensor of the type indicated.

It is to be remembered that $C^r_{.mn}$ is not, in general, symmetrical in m and n ; hence $\Gamma^r_{.mn}$ is not, in general, symmetrical in these indices.

6. Geodesic stability for a general correspondence.

In the present section there is first discussed a problem slightly more general than that of the stability of constrained geodesics. The work up to and including equation (6.37) is a discussion of the stability of the curves defined by the equations

$$(6.1) \quad \frac{d^2 x^r}{du^2} + \Gamma^r_{.mn} \frac{dx^m}{du} \frac{dx^n}{du} = 0 \quad (\Gamma^r_{.mn} \neq \Gamma^r_{.nm}),$$

with the particular first integral

$$(6.11) \quad g_{mn} \frac{dx^m}{du} \frac{dx^n}{du} = 1.$$

¹³) This definition differs from the definition of parallel propagation given by Vranceanu, *Comptes Rendus* 183 (1926), p. 853, which applies only to vectors satisfying the constraints.

¹⁴) *Proc. Nat. Acad. Sci.* 8 (1922), p. 19.

The facts that $\Gamma^r_{.mn}$ are defined by (4.31) and that the curves under discussion are constrained curves are not introduced until after (6.37). In the work prior to that equation $\Gamma^r_{.mn}$ may be any functions of position, subject only to the condition that (6.11) is true. We shall refer to any curve satisfying (6.1) and (6.11) as a *path*. On account of the existence of (6.11), we may put $u = s$ in (6.1).

Let C and C^* (Fig. 1) be two neighbouring paths, and let a correspondence be established in some manner between the points of C and the points of C^* . Let O^* correspond to O , and let P^* correspond to P . Regarding O, O^* as fixed and P, P^* as variable, let us write $s^* = O^*P^*$, $s = OP$. The correspondence between the points of the two paths may be expressed by a relation of the form

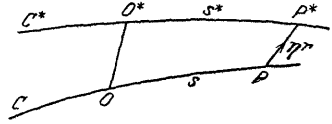


Fig. 1.

$$(6.2) \quad s^* = s + f(s).$$

Now, if λ^r denotes the unit vector tangent to C and η^r the infinitesimal displacement PP^* , we have¹⁵⁾

$$(6.21) \quad f(s) = s^* - s = \int_0^P g_{mn} \lambda^m \bar{\eta}^n ds$$

Hence

$$(6.22) \quad f'(s) = g_{mn} \lambda^m \bar{\eta}^n,$$

and

$$(6.23) \quad f''(s) = g_{mn} \lambda^m \bar{\eta}^n + g_{mn} \lambda^m \bar{\eta}^n.$$

Let x^r denote the coordinates of P . Then the coordinates of P^* are $(x^r + \eta^r)$. Since C^* is a path, we have

$$(6.3) \quad \frac{d^2}{ds^{*2}}(x^r + \eta^r) + \left(\Gamma^r_{.mn} + \eta^s \frac{\partial}{\partial x^s} \Gamma^r_{.mn} \right) \left(\frac{dx^m}{ds^*} + \frac{d\eta^m}{ds^*} \right) \left(\frac{dx^n}{ds^*} + \frac{d\eta^n}{ds^*} \right) = 0.$$

Assuming the function $f(s)$ and its derivatives to be infinitesimals of the order of η^r , and neglecting squares and higher powers of such infinitesimals, we have

$$(6.31) \quad \frac{d}{ds^*} = [1 - f'(s)] \frac{d}{ds},$$

$$(6.32) \quad \frac{d^2}{ds^{*2}} = [1 - 2f'(s)] \frac{d^2}{ds^2} - f''(s) \frac{d}{ds}.$$

Thus (6.3) may be written

¹⁵⁾ Cf. Proc. London Math. Soc. (2) 25 (1926), p. 252.

$$(6.33) \quad [1 - 2f'(s)] \left[\frac{d^2}{ds^2} (x^r + \eta^r) + \left(\Gamma^r_{.mn} + \eta^s \frac{\partial}{\partial x^s} \Gamma^r_{.mn} \right) \left(\frac{dx^m}{ds} + \frac{d\eta^m}{ds} \right) \left(\frac{dx^n}{ds} + \frac{d\eta^n}{ds} \right) \right] - f''(s) \frac{dx^r}{ds} = 0,$$

or, since C is a path,

$$(6.34) \quad \frac{d^2 \eta^r}{ds^2} + (\Gamma^r_{.mn} + \Gamma^r_{.nm}) \frac{d\eta^m}{ds} \lambda^n + \eta^s \lambda^m \lambda^n \frac{\partial}{\partial x^s} \Gamma^r_{.mn} - f''(s) \lambda^r = 0.$$

Now

$$(6.35) \quad \tilde{\eta}^r = \frac{d\eta^r}{ds} + \Gamma^r_{.mn} \eta^m \lambda^n,$$

and

$$(6.36) \quad \begin{aligned} \tilde{\tilde{\eta}}^r = & \frac{d^2 \eta^r}{ds^2} + \eta^m \lambda^n \lambda^s \frac{\partial}{\partial x^s} \Gamma^r_{.mn} + \Gamma^r_{.mn} \lambda^n (\tilde{\eta}^m - \Gamma^m_{.ab} \eta^a \lambda^b) \\ & + \Gamma^r_{.mn} \eta^m (\tilde{\lambda}^n - \Gamma^n_{.ab} \lambda^a \lambda^b) + \Gamma^r_{.mn} \tilde{\eta}^m \lambda^n, \end{aligned}$$

or, since $\tilde{\lambda}^n = 0$,

$$(6.365) \quad \begin{aligned} \tilde{\tilde{\eta}}^r = & \frac{d^2 \eta^r}{ds^2} + 2\Gamma^r_{.mn} \tilde{\eta}^m \lambda^n \\ & + \eta^m \lambda^n \lambda^s \left(\frac{\partial}{\partial x^s} \Gamma^r_{.mn} - \Gamma^r_{.an} \Gamma^a_{.ms} - \Gamma^r_{.ma} \Gamma^a_{.ns} \right). \end{aligned}$$

Substituting for $d^2 \eta^r / ds^2$ in (6.34), we obtain

$$(6.37) \quad \boxed{\tilde{\tilde{\eta}}^r - 2S^r_{.mn} \tilde{\eta}^m \lambda^n + F^r_{.msn} \lambda^m \eta^s \lambda^n - f''(s) \lambda^r = 0,}$$

where

$$(6.38) \quad S^r_{.mn} = \frac{1}{2} (\Gamma^r_{.mn} - \Gamma^r_{.nm}) = \frac{1}{2} (C^r_{.mn} - C^r_{.nm}),$$

the torsion tensor of Cartan, and

$$(6.39) \quad F^r_{.msn} = \frac{\partial}{\partial x^s} \Gamma^r_{.nm} - \frac{\partial}{\partial x^n} \Gamma^r_{.sm} + \Gamma^a_{.nm} \Gamma^r_{.sa} - \Gamma^a_{.sn} \Gamma^r_{.na},$$

a curvature tensor of the manifold with respect to $\Gamma^r_{.mn}$.

Equation (6.37) is the *first tensorial equation for the disturbance vector* η^r . These equations are linear differential equations of the second order.

When we think of C and C^* as constrained geodesics, instead of general paths, we must add the conditions that they should be constrained curves. These conditions are

$$(6.4) \quad B_{(\varrho)n} \lambda^n = 0, \quad \left(B_{(\varrho)n} + \eta^m \frac{\partial}{\partial x^m} B_{(\varrho)n} \right) \left(\lambda^n + \frac{d\eta^n}{ds} \right) = 0,$$

which give

$$(6.41) \quad B_{(\varrho)n} \tilde{\eta}^n + B_{(\varrho)nm} \eta^m \lambda^n = 0,$$

or

$$(6.42) \quad B_{(\varrho)n} \tilde{\eta}^n + (B_{(\varrho)nm} - B_{(\varrho)s} C^s_{.mn}) \eta^m \lambda^n = 0.$$

To investigate the stability of the constrained geodesic C , it is necessary to solve (6.37) and consider only those solutions which satisfy (6.42).

We shall now express (6.37) in another form. We have, analogous to (6.36),

$$(6.5) \quad \bar{\eta}^r = \frac{d^2 \eta^r}{ds^2} + \eta^m \lambda^n \lambda^s \frac{\partial}{\partial x^r} \left\{ \begin{matrix} r \\ mn \end{matrix} \right\} + \left\{ \begin{matrix} r \\ mn \end{matrix} \right\} \lambda^n \left(\bar{\eta}^m - \left\{ \begin{matrix} m \\ ab \end{matrix} \right\} \eta^a \lambda^b \right) \\ + \left\{ \begin{matrix} r \\ mn \end{matrix} \right\} \eta^m \left(\bar{\lambda}^n - \left\{ \begin{matrix} n \\ ab \end{matrix} \right\} \lambda^a \lambda^b \right) + \left\{ \begin{matrix} r \\ mn \end{matrix} \right\} \bar{\eta}^m \lambda^n,$$

or, by virtue of (5.31),

$$(6.51) \quad \bar{\eta}^r = \frac{d^2 \eta^r}{ds^2} + 2 \left\{ \begin{matrix} r \\ mn \end{matrix} \right\} \bar{\eta}^m \lambda^n \\ + \eta^m \lambda^n \lambda^s \left(\frac{\partial}{\partial x^s} \left\{ \begin{matrix} r \\ mn \end{matrix} \right\} - \left\{ \begin{matrix} r \\ an \end{matrix} \right\} \left\{ \begin{matrix} a \\ ms \end{matrix} \right\} - \left\{ \begin{matrix} r \\ ma \end{matrix} \right\} \left\{ \begin{matrix} a \\ ns \end{matrix} \right\} \right) - \eta^m \lambda^s \lambda^t \left\{ \begin{matrix} r \\ mn \end{matrix} \right\} C^{n.st}.$$

Substitution for $d^2 \eta^r / ds^2$ in (6.34) gives

$$(6.52) \quad \boxed{\bar{\eta}^r + (C^r_{.mn} + C^r_{.nm}) \bar{\eta}^m \lambda^n + (G^r_{.mns} + C^r_{.nms}) \lambda^m \eta^s \lambda^n - f''(s) \lambda^r = 0,}$$

where

$$(6.53) \quad G^r_{.mns} = \frac{\partial}{\partial x^s} \left\{ \begin{matrix} r \\ mn \end{matrix} \right\} - \frac{\partial}{\partial x^n} \left\{ \begin{matrix} r \\ sm \end{matrix} \right\} + \left\{ \begin{matrix} a \\ nm \end{matrix} \right\} \left\{ \begin{matrix} r \\ sa \end{matrix} \right\} - \left\{ \begin{matrix} a \\ sm \end{matrix} \right\} \left\{ \begin{matrix} r \\ na \end{matrix} \right\},$$

the curvature tensor of the manifold, and

$$(6.54) \quad C^r_{.nms} = \frac{\partial}{\partial x^s} C^r_{.nm} - C^r_{.am} \left\{ \begin{matrix} a \\ ns \end{matrix} \right\} - C^r_{.na} \left\{ \begin{matrix} a \\ ms \end{matrix} \right\} + C^a_{.nm} \left\{ \begin{matrix} r \\ sa \end{matrix} \right\},$$

the covariant derivative of $C^r_{.nm}$.

Equation (6.52) is the *second tensorial equation for the disturbance vector* η^r . Its left hand side is precisely the left hand side of (6.37) expressed in a different manner. We note that, in the case of no constraint, (6.52) reduces immediately to the equation of Levi-Civita¹⁶.

We shall now obtain from (6.52) an invariant equation for the magnitude of the disturbance vector. Let μ^r be a unit vector codirectional with η^r , so that

$$(6.6) \quad \eta^r = \eta \mu^r, \quad g_{mn} \mu^m \mu^n = 1.$$

Then

$$(6.61) \quad \bar{\eta}^r = \frac{d\eta}{ds} \mu^r + \eta \bar{\mu}^r,$$

and¹⁷

$$(6.62) \quad \bar{\eta}^r \mu_r = \frac{d^2 \eta}{ds^2} - \eta \bar{\mu}^a.$$

¹⁶ Math. Annalen 97 (1926), p. 315, equation (42).

¹⁷ Cf. GD., equation (9.21).

Hence, multiplying (6.52) by μ_r and summing as indicated, we obtain

$$(6.63) \quad \frac{d^2 \eta}{ds^2} - \eta \bar{\mu}^2 + (C^r_{.mn} + C^r_{.nm}) \left(\frac{d\eta}{ds} \mu^m + \eta \bar{\mu}^m \right) \lambda^n \mu_r \\ + (G^r_{.msn} + C^r_{.nms}) \eta \lambda^m \mu^s \lambda^n \mu_r - f''(s) \lambda^r \mu_r = 0,$$

or

$$(6.64) \quad \boxed{\frac{d^2 \eta}{ds^2} + \frac{d\eta}{ds} \mu^r \lambda^m \mu^n (C_{rnm} + C_{rnm}) \\ + \eta [(G_{rmsn} + C_{rnm}) \mu^r \lambda^m \mu^s \lambda^n + (C_{rnm} + C_{rnm}) \mu^r \lambda^m \bar{\mu}^n - \bar{\mu}^2] \\ - f''(s) \lambda^r \mu_r = 0.}$$

This is the *invariant equation for the magnitude of the disturbance vector*.

7. Isometric and normal correspondences.

So far we have been considering a general correspondence, given by the function $f(s)$, between the points of C and C^* . When there are no constraints, C and C^* being then free geodesics, there exists a correspondence of obvious simplicity. This correspondence is obtained by putting $f(s) = 0$; it makes $s^* = s$, and also makes PP^* normal to C , provided that OO^* has been chosen normal to C . In the case of paths (or constrained geodesics) it is no longer possible to obtain a correspondence which combines these two features, and we are led to consider two types of correspondence:

1. The *isometric correspondence*, defined by $s^* = s$. This gives

$$(7.1) \quad f(s) = 0.$$

2. The *normal correspondence*, defined by the condition that PP^* should be normal to C . The analytical expression for this condition is

$$(7.2) \quad g_{mn} \eta^m \lambda^n = 0.$$

The reductions in (6.37), (6.52) and (6.64) in the case of the isometric correspondence are very simple. The last term is to be omitted from each of these equations. We also know, by (6.22), that the particular first integral

$$(7.3) \quad g_{mn} \bar{\eta}^m \lambda^n = 0$$

exists.

In the case of the normal correspondence, differentiation of (7.2) gives

$$(7.4) \quad g_{mn} \bar{\eta}^m \lambda^n + g_{mn} \eta^m \bar{\lambda}^n = 0,$$

or, by (5.31),

$$(7.41) \quad g_{mn} \bar{\eta}^m \lambda^n - g_{mn} C^n_{ab} \eta^m \lambda^a \lambda^b = 0.$$

Differentiating again, we obtain

$$(7.42) \quad g_{mn} \bar{\eta}^m \lambda^n + g_{mn} \bar{\eta}^m \dot{\lambda}^n - g_{mn} C^n_{abc} \eta^m \lambda^a \lambda^b \dot{\lambda}^c \\ - g_{mn} C^n_{ab} \eta^m (-C^a_{cd} \lambda^b \lambda^c \dot{\lambda}^d - C^b_{cd} \lambda^a \lambda^c \dot{\lambda}^d) = 0.$$

Hence, by (6.23),

$$(7.5) \quad f''(s) = \eta^m \lambda^a \lambda^b \lambda^c (C_{mabc} - C_{mpb} C^p_{ca} - C_{map} C^p_{cb}).$$

This is the value for $f''(s)$ which must be substituted in (6.37) and (6.52) in the case of the normal correspondence. The reduction in (6.64) consists in the omission of the last term, since $\lambda^r \mu_r = 0$.

8. Dynamical significance of the paper.

Being given a conservative dynamical system with N coordinates x^r , subject to M constraints given by

$$(8.1) \quad B_{(e)m} dx^m = 0,$$

where $B_{(e)}^r$ are mutually orthogonal unit vectors, the equations of motion may be written¹⁸⁾

$$(8.2) \quad \frac{d^2 x^r}{ds^2} + \left\{ \begin{matrix} r \\ mn \end{matrix} \right\} \frac{dx^m}{ds} \frac{dx^n}{ds} = - B_{(e)}^r B_{(e)mn} \frac{dx^m}{ds} \frac{dx^n}{ds},$$

where ds is the action line-element, given by

$$(8.21) \quad ds^2 = g_{mn} dx^m dx^n = (h - V) a_{mn} dx^m dx^n,$$

h being the constant total energy of the motion, V the potential energy, and $\frac{1}{2} a_{mn} \dot{x}^m \dot{x}^n$ the kinetic energy.

Since (8.2) is identical with (3.5), we have the result:

Theorem V. *The constrained geodesics of the present paper are the curves of motion of a conservative dynamical system, subject to non-holonomic constraints, the line-element being the action line-element.*

Thus the geometrical considerations as to stability apply to conservative dynamical systems subject to non-holonomic constraints, only those disturbances being considered which do not change the total energy.

¹⁸⁾ GD, p. 58, Theorem XVII (A).

Berichtigung

zu dem Aufsatz von St. Cohn-Vossen: „Die parabolische Kurve“,
Math. Annalen **99**, S. 273—308.

Herr Professor Gambier war so freundlich, mich auf seine Abhandlungen über Flächenverbindungen [Bulletin des Sciences Mathématiques (2) **44** (1920) und (2) **45** (1921)] aufmerksam zu machen, die in anderm Zusammenhang ähnliche Fragen und Ergebnisse enthalten wie der dritte Abschnitt meiner Arbeit. Ich bedaure, diese Abhandlungen erst nachträglich kennen gelernt zu haben; der Literaturbericht zu Beginn meines Aufsatzes ist demgemäß zu berichtigen.

Göttingen, den 18. 5. 1928.

Stefan Cohn-Vossen.
