

Werk

Titel: Commentarii Mathematici Helvetici

Verlag: Birkhäuser

Jahr: 1933

Kollektion: Mathematica

Digitalisiert: Niedersächsische Staats- und Universitätsbibliothek Göttingen

Werk Id: PPN358147735_0005

PURL: http://resolver.sub.uni-goettingen.de/purl?PPN358147735_0005

Übergeordnetes Werk

Werk Id: PPN358147735

PURL: <http://resolver.sub.uni-goettingen.de/purl?PPN358147735>

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 there 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.

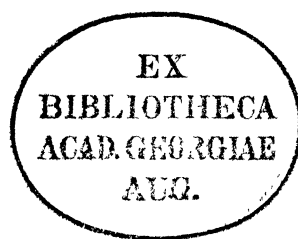
Contact

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

COMMENTARII MATHEMATICI HELVETICI

VOLUMEN **5**
MCMXXXIII

EDITI
SOCIETATE MATHEMATICA
HELVETICA IN AEDIBUS
ORELL FÜSSLI TURICI



21933.4442

INHALTSVERZEICHNIS

TABLE DES MATIÈRES

INDICE

	Seite - page pagina
<i>Ahlfors, Lars</i> , Ueber die Kreise, die von einer Riemannschen Fläche schlicht überdeckt werden	28
<i>Carathéodory, C.</i> , Ueber die Einteilung der Variationsprobleme von Lagrange nach Klassen	I
<i>Finsler, P.</i> , Die Existenz der Zahlenreihe und des Kontinuums ...	88
<i>Fischer, Anna</i> , Grundlagen der elliptischen Geometrie ...	246
<i>Fréchet, Maurice</i> , Les probabilités continues « en chaîne » ...	175
<i>Fueter, Rudolf</i> , Ein Satz über die Ring- und Strahlklassenzahlen in algebraischen Zahlkörpern	319
<i>Gonseth, F.</i> , Sur l'axiomatique de la théorie des ensembles et sur la logique des relations	108
<i>Hadamard, J.</i> , La propagation des ondes et les caustiques ...	137
<i>Hopf, Heinz</i> , Die Klassen der Abbildungen der n -dimensionalen Polyeder auf die n -dimensionale Sphäre	39
<i>Mächler, W.</i> , Laplace'sche Integraltransformation und Integration partieller Differentialgleichungen vom hyperbolischen und parabolischen Typus	256
<i>Nevanlinna, Rolf</i> , Ein Satz über die konforme Abbildung Riemannscher Flächen	95
<i>Ostrowski, Alexander</i> , Asymptotische Abschätzung des absoluten Be- trages einer Funktion, die die Werte 0 und 1 nicht annimmt	55
<i>Rivier, W.</i> , Note sur un type d'équations différentielles du premier ordre	254
<i>Scherrer, W.</i> , Ueber eine Eigenschaft der Raumkurven ...	20
— Geometrische Deutung des Gauß'schen Verschlingungsintegrals	25
<i>Skolem, Th.</i> , Ein allgemeines quadratisches Reziprozitätsgesetz in den- jenigen algebraischen Zahlkörpern, worin 2 voll zerfällt ...	305

Ueber die Einteilung der Variationsprobleme von Lagrange nach Klassen

Von C. CARATHÉODORY, z. Z. Athen

Einleitung

1. Das Ziel dieser Arbeit ist die Aufstellung einer für alle regulären Punkttransformationen geltenden Invariante, die für das Studium der Variationsprobleme von Lagrange grundlegend ist. Es handelt sich um Folgendes: ist im $(n+1)$ -dimensionalen Raum der $(t, x_1, \dots, x_n)$ ein Variationsproblem mit der Funktion $L(t, x_i, \dot{x}_i)$ unter dem Integral und und den p Differentialgleichungen

$$(1.1) \quad G_{k'}(t, x_i, \dot{x}_i) = 0 \quad (k' = 1, \dots, p)$$

als Nebenbedingungen gegeben, so füllen die Extremalen, die aus einem beliebigen aber festen Punkt A des Raumes ausgehen, einen Raumteil aus, dessen Dimension mit $(n+1-q)$ bezeichnet werden soll. Die Zahl q , die für das betrachtete Problem jeweils charakteristisch ist, soll die *Klasse* des Problems heißen; es wird sich zeigen, daß sie von Null bis p inklusive variieren kann, und daß sie aus den Funktionen L und $G_{k'}$ mit Hilfe von Differentiationen (und Eliminationen) bestimmbar ist. Es handelt sich also um eine reine Differentialinvariante.

2. Das Problem von Lagrange ist in den letzten Jahren von verschiedenen Autoren, besonders von *Bliss*¹⁾ und seinen Schülern und

¹⁾ *G. A. Bliss*. A note on the problem of Lagrange in the calculus of variations. Amer. Bull. 22 (1916) pp. 220—225.

D. M. Smith. Jacobi's condition for the problem of Lagrange in the calculus of variations. Trans. Amer. Math. Soc. 17 (1916) p. 459.

G. A. Bliss. The problem of Mayer with variable end points. Amer. Trans. 19 (1918) pp. 305—314.

Miss G. A. Larew. Necessary conditions in the problem of Mayer in the calculus of variations. Amer. Trans. 20 (1919) p. 1.

G. A. Bliss. Some recent developments in the calculus of variations. Amer. Bull. 26 (1920) pp. 343—361.

G. A. Bliss. The transformation of Clebsch in the calculus of variations. (Proceed. intern. mathem. Congress Toronto 1924, Vol. I, pp. 589—603).

Miss G. A. Larew. The Hilbert Integral and Mayer Fields for the problem of Mayer in the calculus of variations. Amer. Trans. 26 pp. 61—67.

G. A. Bliss. The problem of Lagrange in the calculus of variations. Autograph. Vorlesung (1925) pp. 1—75.

G. A. Bliss. The problem of Lagrange in the calculus of variations. Amer. Journ. of Mathem. 52 (1930) pp. 672—748.

außerdem von *Marston Morse*²⁾ behandelt worden. Der Grund aber, weshalb die Lagrangeschen Probleme nicht schon in diesen sehr sorgfältigen Arbeiten nach Klassen eingeteilt worden sind, ist folgender: nach den Arbeiten von *E. v. Escherich*³⁾, der zuerst auf gewisse Singularitäten der Lagrangeschen Probleme hingewiesen hatte (cf. § 18), hat man sich im Anschluß an *H. Hahn* und *Bolza*⁴⁾ gewöhnt, nur solche Extremalen zu betrachten, die sich, wie man sagt, «normal verhalten». Auf diese Weise konnte man allerdings die Schwierigkeiten, auf welche *v. Escherich* hingewiesen hatte, umgehen, aber es wurde gleichzeitig der Weg verbaut, der zu den feineren Unterscheidungen im Verhalten der Extremalen führt. In der Tat ist z. B. die Bedingung des normalen Verhaltens einer Extremalen bei den Lagrangeschen Problemen mit festen Endpunkten im wesentlichen mit der Forderung identisch, daß die Klasse des Problems verschwinden soll. Und ganz ähnlich ist die Theorie des normalen Verhaltens der Extremalen von Lagrangeschen Variationsproblemen mit variablen Endpunkten mit der Klasse des Problems verknüpft. Diese Zusammenhänge werden erst recht verständlich, wenn man den Begriff der Klasse unabhängig von der üblichen Theorie entwickelt, und deshalb werde ich von allen oben zitierten Arbeiten absehen, und mich nur auf die Formeln stützen, die ich in einer früheren Untersuchung angegeben habe⁵⁾. Diese Abhandlung wird im Folgenden mit *Geod. Aeq.* zitiert.

3. Zusammenstellung bekannter Resultate.

Wir nehmen an, daß die Funktionalmatrix

$$(3.1) \quad \left(\frac{\partial G_{k'}}{\partial \dot{x}_i} \right) \quad (k' = 1, \dots, p; i = 1, \dots, n)$$

²⁾ a) The problems of Lagrange and Mayer under general end conditions. Proc. Nat. Acad. of Science. 16 (1930) pp. 229—233.

b) Sufficient conditions in the problem of Lagrange with variable end conditions. Amer. Journ. of Mathem. 53 (1931) pp. 517—546.

c) Sufficient conditions in the problem of Lagrange with fixed end points. Annals of Mathem. 32 (1931) pp. 567—577.

Ferner:

M. Morse und *S. B. Myers*. The problems of Lagrange and Mayer with variable end points. Proceed. Amer. Acad. of Arts and Science (Boston Mass.) 66 (1931) pp. 235—253.

³⁾ *E. v. Escherich* Die zweite Variation der einfachen Integrale. Wiener Sitzungsber. Mathem. Naturw. Klasse 107, 108, 110 (1898, 1899, 1901).

⁴⁾ *H. Hahn*. Math. Ann. 58 (1903) p. 152.

O. Bolza. Vorlesungen über Variationsrechnung. (Leipzig, Teubner, 1909) p. 564.

⁵⁾ *C. Carathéodory*. Die Methode der geodätischen Aequidistanten und das Problem von Lagrange. Acta Mathematica 47 (1926) pp. 199—236.

der linken Seite von (1. 1) den Rang p besitzt. Außerdem soll für die Linienelemente, die man betrachtet, die *Legendresche Bedingung* erfüllt sein (Geod. Aeq. p. 210). Es soll also, wenn man mit $\mu_1, \dots, \mu_p$ Lagrangesche Multiplikatoren bezeichnet, und die Funktion

$$(3. 2) \quad M(t, x_i, \dot{x}_i, \mu_j) = L + \mu_k G_k$$

eingführt, die Gleichung in ρ

$$(3. 3) \quad D(\rho) = \begin{vmatrix} M_{\dot{x}_i \dot{x}_j} - \delta_{ij} \rho, & \frac{\partial G_k}{\partial \dot{x}_i} \\ \frac{\partial G_m}{\partial \dot{x}_j}, & 0 \end{vmatrix} = 0$$

lauter positive, von Null verschiedene Wurzeln besitzen.

Unter diesen Umständen kann man an Stelle der $(2n + p + 1)$ Veränderlichen, die in (3. 2) auftreten, und die außerdem den p Bedingungs-
gleichungen (1. 1) genügen sollen, $(2n + 1)$ von einander unabhängige kanonische Koordinaten (t, x_i, y_i) einführen, die erlauben, alle Einzelheiten des Problems mit Hilfe einer einzigen Hamiltonschen Funktion $H(t, x_i, y_i)$ zu untersuchen (Geod. Aeq. p. 218).

Ist umgekehrt eine beliebige Funktion $H(t, x_i, y_i)$ gegeben, die mindestens zwei Mal stetig differenzierbar ist, so kann sie als Hamiltonsche Funktion eines Variationsproblems angesehen werden, für welches die Legendresche Bedingung erfüllt ist, wenn die Gleichung in σ

$$(3. 4) \quad |H_{y_i y_j} - \delta_{ij} \sigma| = 0$$

keine einzige negative Wurzel besitzt (Geod. Aeq. p. 221).

Die Anzahl ferner der verschwindenden Wurzeln der Gleichung (3. 4) ist immer gleich der Anzahl der Bedingungs-
gleichungen (1. 1) des Lagrangeschen Problems, welchem die Funktion H zugeordnet ist. Ist also insbesondere die Hessesche Determinante $|H_{y_i y_j}| \neq 0$, so ist die Anzahl der Bedingungs-
gleichungen gleich Null, d. h. das betrachtete Variations-
problem ist ein gewöhnliches und kein Problem von Lagrange.

4. Die Lösung eines Variationsproblems mit der Hamiltonschen Funktion $H(t, x_i, y_i)$ hängt bekanntlich von der Konstruktion einer „vollständigen Figur“ ab, die aus einer Schar geodätisch äquidistanter Flächen

$$(4. 1) \quad S(t, x_i) = \lambda$$

und der sie durchsetzenden Extremalenschar besteht. Man erhält die Flächenschar (4. 1) durch Integration der partiellen Differentialgleichung (Geod. Aeq. p. 222)

$$(4. 2) \quad S_t + H(t, x_i, S_{x_i}) = 0$$

Die Cauchyschen Charakteristiken von (4. 2) sind Lösungen des Systems

$$(4. 3) \quad \dot{x}_i = H_{y_i}, \quad \dot{y}_i = -H_{x_i}$$

von kanonischen Differentialgleichungen.

Die Extremalen unseres Variationsproblems werden durch die Projektion $x_i = x_i(t)$ dieser Kurven auf den $(n + 1)$ -dimensionalen Raum der (t, x_i) dargestellt.

Betrachtet man ein Stück des Raumes, das von der Flächenschar (4. 1) einfach überdeckt wird, so bilden die sie transversal schneidenden Extremalen, die man auch als Lösungen der Differentialgleichungen

$$(4. 4) \quad \dot{x}_i = H_{y_i}(t, x_j, S_{x_j}),$$

gewinnen kann, ein Feld, auf welches die Weierstraßsche Theorie anwendbar ist (Geod. Aeq. p. 215)⁶⁾.

5. *Ein Hilfssatz.*

Wir betrachten nun einen Punkt P_0 des Raumes der (t, x_i) und ein Extremalenstück e_0 , das P_0 enthält. Aus der Integrationstheorie der partiellen Differentialgleichung (4. 2) folgt ein Resultat, das für alles Weitere grundlegend ist.

Man kann mit P_0 als Mittelpunkt eine Kugel κ abgrenzen, sodaß jede Extremale e , die in einer gewissen Nachbarschaft von e_0 liegt, in ein Feld eingebettet werden kann, das die Kugel κ vollständig überdeckt.

Wir bezeichnen mit $\mathfrak{E}$ die Gesamtheit der Extremalen e , die nach geeigneter Wahl von κ die vorstehende Eigenschaft besitzen. erinnert man sich daran, daß für alle betrachteten Extremalen die Legendresche

⁶⁾ Man bemerke, daß durch die Benutzung der kanonischen Koordinaten jeder Unterschied zwischen den gewöhnlichen und den sogenannten „Mayerschen“ Feldern von Extremalen in Fortfall kommt. Es ist ein großer Vorteil, daß die sehr komplizierte Begriffsbildung der Mayerschen Felder vermieden werden kann.

Bedingung erfüllt sein soll, so folgt, wenn man bei der Wahl von $\mathfrak{E}$ einige auf der Hand liegende Einschränkungen macht, die die Benützung der Weierstraßschen E -Funktion sicherstellen, der

Satz 1. *Keine zwei Extremalenstücke e_1 und e_2 aus $\mathfrak{E}$, die beide im Innern von $\mathfrak{x}$ liegen, können dieselben Endpunkte besitzen, ohne zusammenzufallen.*

Um dies zu zeigen, nehmen wir an, daß e_1 und e_2 dieselben Punkte A und B verbinden, ohne identisch zu sein, und bezeichnen mit $\mathcal{J}_1$ und $\mathcal{J}_2$ die Werte des Integrals über L längs e_1 und e_2 . Nach Konstruktion eines Feldes von Extremalen, das e_1 enthält, kann man die Differenz $(\mathcal{J}_2 - \mathcal{J}_1)$ mit Hilfe der E -Funktion, die zu diesem Felde gehört, durch die Gleichung

$$(5.1) \quad \mathcal{J}_2 - \mathcal{J}_1 = \int_{e_2} E dt$$

darstellen. Nun ist einerseits $E \geq 0$, aber E ist in der vorstehenden Gleichung nicht durchweg Null, weil nach unserer Voraussetzung Punkte existieren, in welchen e_2 die durch diese Punkte gehende Feldkurve unter einem von Null verschiedenen Winkel schneidet, und weil dann in diesen Punkten $E > 0$ ist. Hieraus schließt man, daß $\mathcal{J}_2 > \mathcal{J}_1$ sein muß. Durch Vertauschung von e_1 mit e_2 hätte man aber $\mathcal{J}_1 > \mathcal{J}_2$ gefunden, was zu einem Widerspruche führt.

6. Das accessorische Variationsproblem.

Die kanonischen Koordinaten x_i, y_i sollen längs einer festen Extremalen e_0 den Gleichungen

$$(6.1) \quad x_i = \bar{x}_i(t), \quad y_i = \bar{y}_i(t)$$

genügen. Wir führen zur Abkürzung folgende Funktionen von t ein, die also längs e_0 definiert sind:

$$(6.2) \quad a_{ij}(t) = H_{x_i x_j}(t, \bar{x}_k, \bar{y}_k)$$

$$(6.3) \quad b_{ij}(t) = H_{x_i y_j}(t, \bar{x}_k, \bar{y}_k)$$

$$(6.4) \quad c_{ij}(t) = H_{y_i y_j}(t, \bar{x}_k, \bar{y}_k).$$

Stellt man durch die Gleichungen

$$(6.5) \quad x_i = x_i(t, u), \quad y = y_i(t, u)$$

die kanonischen Koordinaten der Linienelemente einer Extremalenschar dar, die von einem Parameter u abhängt, und reduzieren sich für $u = 0$ die Funktionen (6.5) auf $\bar{x}_i$ bzw. $\bar{y}_i$; setzt man ferner

$$(6.6) \quad \xi_i(t) = \left. \frac{\partial x_i(t, u)}{\partial u} \right|_{u=0}, \quad \eta_i(t) = \left. \frac{\partial y_i(t, u)}{\partial u} \right|_{u=0},$$

so sind bekanntlich die Funktionen $\xi_i(t)$ und $\eta_i(t)$ Lösungen der Differentialgleichungen

$$(6.7) \quad \begin{cases} \dot{\xi}_i = b_{ki} \xi_k + c_{ik} \eta_k \\ \dot{\eta}_i = -a_{ik} \xi_k - b_{ik} \eta_k. \end{cases}$$

7. Man bemerke jetzt, daß die Funktion

$$(7.1) \quad \mathcal{H}(t, \xi_i, \eta_i) = \frac{1}{2} a_{ij} \xi_i \xi_j + b_{ij} \xi_i \eta_j + \frac{1}{2} c_{ij} \eta_i \eta_j$$

als Hamiltonsche Funktion eines Variationsproblems angesehen werden kann, dessen Extremalen durch die kanonischen Gleichungen (6.7) bestimmt werden. Die Determinante (3.4) des ursprünglichen Variationsproblems fällt, wenn man sie für die Linienelemente von e_0 bildet, mit der analogen Determinante des neuen Variationsproblems zusammen, woraus man schließt, daß auch hier die Legendresche Bedingung erfüllt ist. Von einem verwandten Resultat ausgehend, hat *Bliss* in der Arbeit, die er dem Toronto Congreß vorgelegt hat, außerordentlich elegante Schlüsse gezogen. Für uns wird vor allem maßgebend sein, daß man den Satz 1 des § 5 auch auf die Lösungen der Differentialgleichungen (6.7) anwenden kann. Wählen wir für die eine der beiden Extremalen, die in diesem Satze benutzt werden, die Lösung $\xi_i \equiv 0, \eta_i \equiv 0$ der Gleichungen (6.7), so können wir nunmehr folgenden Satz aussprechen:

Satz 2. *Jeder Punkt t_0 der t -Achse, für welchen die früheren Voraussetzungen gelten, ist Mittelpunkt eines Intervalls $\delta(t_0)$, in dem eine beliebige Lösung der Differentialgleichung (6.7) höchstens eine gemeinsame Nullstelle der $\xi_i(t)$ besitzen kann, falls nicht für diese Lösung alle $\xi_i(t)$ identisch Null sind.*

8. Aufstellung der Klasse des Variationsproblems.

Wir betrachten die Gesamtheit der Extremalen unseres Variationsproblems, die durch einen festen Punkt (t^0, x_i^0) hindurchgehen, und in einer Umgebung einer Extremalen e^0 liegen, die diesen Punkt ebenfalls enthält. Diese Extremalen werden durch Lösungen

$$(8.1) \quad x_i = x_i(t, u_j), \quad y_i = y_i(t, u_j)$$

der kanonischen Differentialgleichungen (4.3) dargestellt, die von n Parametern $(u_1, \dots, u_n)$ abhängen, und den Anfangsbedingungen

$$(8.2) \quad x_i(t^0, u_j) = x_i^0, \quad y_i(t^0, u_j) = u_i$$

genügen.

Um die Dimension desjenigen Teiles des Raumes festzustellen, der von der Extremalenschar (8.1) überdeckt wird, muß man den Rang der Funktionaldeterminante

$$(8.3) \quad \left| \frac{\partial x_i}{\partial u_j} \right|$$

untersuchen.

Nun führe man die Bezeichnungen ein:

$$(8.4) \quad \xi_{ij}(t) = \frac{\partial x_i}{\partial u_j}, \quad \eta_{ij}(t) = \frac{\partial y_i}{\partial u_j},$$

wobei man in die rechten Seiten dieser Gleichungen diejenigen Werte der u_k einsetzt, die zur Extremalen e^0 gehören. Dann sind die Funktionen ξ_{ij} , η_{ij} Lösungen der kanonischen Differentialgleichungen (6.7) des accessorischen Problems, die in Folge von (8.2) die Anfangsbedingungen

$$(8.5) \quad \xi_{ij}(t^0) = 0, \quad \eta_{ij}(t^0) = \delta_{ij}$$

befriedigen müssen. Die Determinante (8.3) nimmt jetzt die Gestalt an

$$(8.6) \quad \Delta(t, t^0) = |\xi_{ij}(t)|$$

und alles kommt darauf hinaus ihren Rang $R(t)$ zu bestimmen, falls $t \neq t_0$ ist.

9. Entweder ist jetzt im ganzen Intervall $\delta(t_0)$ des Satzes 2 (§ 7) der Rang $R(t)$ von $A(t, t_0)$ gleich n , oder es gibt im Innern dieses Intervalls einen Punkt t_1 , für welchen dieser Rang gleich $(n-q)$ und $q > 0$ ist.

Im letzten Falle gibt es q linear unabhängige n -dimensionale Vektoren

$$(9.1) \quad \lambda_k^{(\alpha)} \quad (k = 1, \dots, n; \alpha = 1, \dots, q),$$

so daß mit den Bezeichnungen

$$(9.2) \quad \xi_i^{(\alpha)}(t) = \lambda_k^{(\alpha)} \xi_{ik}(t), \quad \eta_i^{(\alpha)}(t) = \lambda_k^{(\alpha)} \eta_{ik}(t)$$

für jedes α und jedes i alle Gleichungen

$$(9.3) \quad \xi_i^{(\alpha)}(t_1) = 0$$

erfüllt sind. Da nun für jedes α die Funktionen $\xi_i^{(\alpha)}, \eta_i^{(\alpha)}$ Lösungen der Differentialgleichungen (6.7) sind, und da die $\xi_i^{(\alpha)}$ alle in den beiden Punkten t_0 und t_1 verschwinden, müssen nach dem Satze 2 des § 7 die Gleichungen

$$(9.4) \quad \xi_i^{(\alpha)}(t) = 0$$

identisch befriedigt sein.

Hieraus folgt nun weiter, daß in jedem Punkt (t) des Intervalles $\delta(t^0)$ mindestens q linear unabhängige Kombinationen der Zeilen der Determinante (8.6) existieren, die in diesem Punkte verschwinden, und man entnimmt hieraus, daß überall $R(t) \leq R(t_1)$ sein muß. Da nun t_1 einen beliebigen von t^0 verschiedenen Punkt des Intervalles $\delta(t^0)$ bedeuten sollte, für welchen $R(t_1) < n$ war, ist dies nur dann möglich, wenn $R(t)$ für alle von t^0 verschiedenen Punkte immer denselben Wert $(n-q)$ besitzt. Diese Zahl q , die auch Null sein kann, soll die *Klasse der Extremale* e^0 im Punkte t^0 genannt werden.

10. Die q Funktionen $\eta_i^{(\alpha)}(t)$, die in der zweiten Gleichung (9.2) vorkommen, sind Lösungen der Differentialgleichungen (6.7), in welchen man die ξ_k identisch gleich Null gesetzt hat. Sie sind also Lösungen der Differentialgleichungen

$$(10.1) \quad \dot{\eta}_i = -b_{ik} \eta_k,$$

die außerdem noch die n linearen Gleichungen

$$(10.2) \quad c_{ik} \eta_k = 0$$

befriedigen. Für $t = t^0$ hat man außerdem, wegen (8. 5) und (9. 2)

$$(10. 3) \quad \eta_i^{(\alpha)}(t^0) = \lambda_i^{(\alpha)},$$

und hieraus entnimmt man, daß die q Funktionen $\eta_i^{(\alpha)}(t)$ linear unabhängig sind.

Wir bezeichnen mit δ^* ein beliebiges Teilintervall von $\delta(t^0)$ und mit q^* die Anzahl der linear unabhängigen Lösungen von (10. 1) die im Intervalle δ^* den Relationen (10. 2) genügen. Es ist selbstverständlich, daß

$$(10. 4) \quad q^* \geq q$$

ist. Nun bemerke man, daß jede dieser Lösungen auch aufgefaßt werden kann als Lösung von (6. 7), für welche alle ξ_i in δ^* identisch verschwinden. Da nun nach Voraussetzung δ^* in $\delta(t^0)$ enthalten ist, müssen nach dem Satze 2 des § 7 alle diese ξ_i überall in $\delta(t_0)$ verschwinden. Hieraus folgt, daß der Rang $R(t)$ der Determinante (8. 6) der Bedingung

$$(10. 5) \quad R(t) \leq n - q^*$$

genügt. Man hat also insbesondere $n - q \leq n - q^*$ und durch Vergleichung mit (10. 4)

$$(10. 6) \quad q^* = q.$$

Andererseits ist aber auch q^* die Klasse der Extremalen e_0 in einem beliebigen Punkte von δ^* ; man sieht also, daß die Klasse von e_0 unabhängig ist von der Wahl des Punktes P_0 auf dieser Extremalen. Es gilt daher der

Satz 3. *Es sei e^0 ein Extremalenstück, auf welchem die Legendresche Bedingung erfüllt ist. Die Gesamtheit der Extremalen, die durch einen beliebigen aber festen Punkt von e^0 hindurchgehen, bildet eine Punktmenge, deren Dimension $(n + 1 - q)$ ist; die Zahl q , die positiv oder Null ist, und die außerdem unabhängig von der Wahl von P_0 ist, heißt die Klasse der Extremalen e^0 . Sie ist gleich der Anzahl derjenigen linear unabhängigen Lösungen von (10. 1), die der Bedingung (10. 2) genügen.*

11. Berechnung der Klasse. Ist das gestellte Variationsproblem ein gewöhnliches, so ist der Rang der Matrix (c_{ik}) gleich n und das System von Gleichungen (10. 2) besitzt dann keine von Null verschiedene Lösung. In diesem Falle ist selbstverständlich die Klasse q gleich Null.

Wir nehmen also an, daß wir ein Lagrangesches Variationsproblem mit p Differentialgleichungen als Nebenbedingungen vor uns haben, was gleichbedeutend damit ist, daß der Rang der Matrix (c_{ik}) gleich $(n - p)$ sein soll.

Wir führen nun die Bezeichnung ein

$$(11.1) \quad v_i(t) = c_{ik}(t) \cdot \eta_k(t).$$

Setzen wir für die η_k beliebige Lösungen der Differentialgleichungen (10.1) ein, so sind die Funktionen (11.1) mindestens p Mal stetig differenzierbar, falls, wie wir von jetzt ab annehmen wollen, die Hamiltonsche Funktion H mindestens $(p+2)$ Mal stetig differenzierbar ist. Aus (11.1) und (10.1) folgt dann

$$\begin{aligned} \frac{dv_i}{dt} &= \dot{c}_{ik} \eta_k + c_{ij} \dot{\eta}_j \\ &= (\dot{c}_{ik} - c_{ij} b_{jk}) \eta_k. \end{aligned}$$

Setzt man also zur Abkürzung

$$(11.2) \quad c_{ik}^{(1)} = \dot{c}_{ik} - c_{ij} b_{jk}$$

und allgemein für $m = 1, 2, \dots$

$$(11.3) \quad c_{ik}^{(m+1)} = \dot{c}_{ik}^{(m)} - c_{ij}^{(m)} b_{jk},$$

so hat man für alle betrachteten Werte von m

$$(11.4) \quad \frac{d^m v_i}{dt^m} = c_{ik}^{(m)} \eta_k.$$

Es handelt sich jetzt darum, die Anzahl der linear unabhängigen Lösungen $\eta_k^{(\alpha)}$ von (10.1) zu bestimmen, für welche alle

$$(11.5) \quad v_i^{(\alpha)} = c_{ik}^{(\alpha)} \eta_k^{(\alpha)}$$

identisch verschwinden.

12. Zu diesem Zweck betrachten wir die Matrix

$$(12.1) \quad A_m = \begin{pmatrix} c_{ik} \\ c_{i_1 k}^{(1)} \\ \dots \\ c_{i_m k}^{(m)} \end{pmatrix},$$

die n Kolonnen und $m \cdot n$ Zeilen besitzt und bezeichnen mit r_m ihren Rang. Das heißt r_m soll die Ordnung der größten Unterdeterminante von (12. 1) darstellen, die auf dem betrachteten Extremalenstück e_0 nicht *identisch* verschwindet. Es ist jedenfalls

$$(12. 2) \quad (n - p) \leq r_1 \leq \dots \leq r_p.$$

Hat man nun $r_p = n$, so gibt es keine einzige nicht identisch verschwindende Lösung η_i von (10. 1), für welche die rechten Seiten von (11. 1) und von (11. 4) für $m = 1, \dots, p$ alle identisch Null sind. In diesem Falle muß die Klasse von e^0 gleich Null sein.

Ist dagegen $r_p < n$, so gibt es wegen (12. 2) einen *kleinsten* Wert für m zwischen Null und $(p - 1)$, sodaß

$$(12. 3) \quad r_{m+1} = r_m$$

ist; außerdem ist

$$(12. 4) \quad n - r_m > 0.$$

Es gibt ferner auf der t -Achse ein Intervall δ , auf welchem eine r_m -reihige Unterdeterminante der Matrix A_m von Null verschieden ist. Man kann dann Koeffizienten

$$g_{ji}^{(1)}(t), g_{ji}^{(2)}(t), \dots, g_{ji}^{(m)}(t)$$

finden, die auf δ stetig sind und für welche die Gleichungen

$$(12. 4) \quad c_{jk}^{(m+1)} = g_{ji} \cdot c_{ik} + g_{ji_1}^{(1)} c_{i_1 k}^{(1)} + \dots + g_{ji_m}^{(m)} c_{i_m k}^{(m)} \quad (j, k = 1, \dots, n)$$

sämtlich erfüllt sind.

Außerdem können wir $(n - r_m)$ linear unabhängige Lösungen $\eta_k^{(\alpha)}$ von (10. 1) aufzufinden machen, die in einem Punkte t^0 von δ den Bedingungen

$$(12. 5) \quad v_i^{(\alpha)}(t^0) = 0, \frac{dv_i^{(\alpha)}(t^0)}{dt} = 0, \dots, \frac{d^{(m)} v_i^{(\alpha)}(t^0)}{dt^m} = 0$$

genügen. Aus (11. 4) und (12. 4) folgt aber andererseits daß diese Funktionen $v_i^{(\alpha)}$ sämtlich Lösungen der linearen homogenen Differentialgleichungen

$$\frac{d^{m+1} v_j}{dt^{m+1}} = g_{ji} v_i + g_{ji_1}^{(1)} \frac{dv_{i_1}}{dt} + \dots + g_{ji_m}^{(m)} \frac{d^m v_{i_m}}{dt^m}$$

sind.

Hieraus folgt, daß die $v_i^{(\alpha)}$ in δ identisch verschwinden müssen, und es muß daher die Klasse q unseres Variationsproblems der Relation

$$(12.7) \quad q \geq n - r_m$$

genügen.

Andererseits ist der Rang r_p der Matrix A_p höchstens gleich $(n - q)$; man hat daher, mit Berücksichtigung von (12.2)

$$(12.8) \quad r_m \leq r_p \leq (n - q).$$

Die Vergleichung von (12.7) mit (12.8) und (12.2) liefert schließlich die Relationen

$$(12.9) \quad r_m = r_{m+1} = \dots = r_p = (n - q).$$

Man sieht außerdem, daß für q die Bedingungen

$$(12.10) \quad 0 \leq q \leq p$$

gelten, und man kann durch Beispiele feststellen, daß q alle Werte zwischen Null (inklusive) und p (inklusive) annehmen kann.

Es ist sehr bemerkenswert, daß man den Rang der Determinante $\Delta(t, t^0)$ des § 8 bestimmen kann, ohne die Differentialgleichungen (6.7) zu integrieren. Es gilt also der

Satz 4. *Die Klasse q der Extremale e ist immer gleich $(n - r_p)$, wenn r_p den Rang der Matrix A_p bezeichnet.*

13. Probleme von der Klasse Null.

Aus unseren früheren Ausführungen folgt, daß alle gewöhnlichen Variationsprobleme von der Klasse Null sind. Aber auch für Lagrangesche Variationsprobleme sind die Probleme von nicht verschwindender Klasse die Ausnahme. Wir nehmen als Beispiel die Hamiltonsche Funktion

$$(13.1) \quad H = \frac{1}{2} y_1^2 + x_1 y_2.$$

Hier ist $n = 2$ und man findet, daß die Matrizen der c_{ij} und $c_{ij}^{(1)}$ folgendermaßen lauten:

$$(13.2) \quad (c_{ij}) = \begin{pmatrix} 1, 0 \\ 0, 0 \end{pmatrix}, \quad (c_{ij}^{(1)}) = \begin{pmatrix} 0, -1 \\ 0, 0 \end{pmatrix};$$

hieraus folgt dann sofort, daß $p = 1$ und $q = 0$ ist. Die kanonischen Differentialgleichungen der Extremalen

$$(13.3) \quad \dot{x}_1 = y_1, \quad \dot{y}_1 = -y_2, \quad \dot{x}_2 = x_1, \quad \dot{y}_2 = 0$$

lassen sich sofort integrieren. Z. B. können die Extremalen, die durch den Anfangspunkt der Koordinaten hindurchgehen, folgendermaßen geschrieben werden

$$x_1 = 3a t^2 + 2bt, \quad x_2 = at^3 + bt^2,$$

und man berechnet, daß die Funktionaldeterminante

$$\frac{\partial(x_1, x_2)}{\partial(a, b)} = t^4,$$

wie es sein soll, nicht identisch Null ist.

Für die Probleme der Klasse Null ist es außerordentlich einfach, die hinreichenden Bedingungen für die Existenz eines Minimums und auch die Theorie der konjugierten Punkte aufzustellen. Diese Dinge sind wiederholt, zuletzt von *Morse* (s. Fußnote ² c)), behandelt worden. Die Behauptung von *Morse*, daß die üblichen Methoden hier versagen, ist aber nur bedingt richtig. Man kann sehr gut die gewöhnliche Methode benutzen, wenn man sich eines Kunstgriffes bedient, den *L. Tonelli* erfunden hat ⁷).

14. *Probleme mit endlichen Gleichungen als Nebenbedingungen.*

Unter den Lagrangeschen Problemen, deren Klasse nicht verschwindet, sind diejenigen am einfachsten zu behandeln, die unter den Nebenbedingungen einige von der Gestalt

$$(14.1) \quad \psi(t, x_j) = \text{const.},$$

oder, was auf dasselbe hinauskommt, Nebenbedingungen von der Form

$$(14.2) \quad \frac{\partial \psi}{\partial t} + \frac{\partial \psi}{\partial x_j} \dot{x}_j = 0$$

⁷) *L. Tonelli*. Sul problema isoperimetrico con un punto terminale mobile (Mem. R. Acc. Bologna (7) Vol. 10, 1922–23). S. auch Geod. Aequid. p. 217.

enthalten. Allgemeiner kann man verlangen, daß nach einer geeigneten Punkttransformation die Hamiltonsche Funktion $H(t, x_i, y_i)$ in den neuen Koordinaten die Veränderlichen $(y_{m+1}, \dots, y_n)$ nicht enthält.

Die Behandlung dieser Probleme kann man offenbar auf diejenige eines Problems in einem Raume niedrigerer Dimension zurückführen, und in diesem reduzierten Raume kann die Klasse des Problems sehr wohl gleich Null sein.

Dies ist z. B. der Fall, wenn man das Lagrangesche Problem mit der Hamiltonschen Funktion

$$(10.3) \quad H = \frac{(y_1 - y_2)^2}{2}$$

betrachtet; man verifiziert sofort auch auf direktem Wege, daß man hier

$$(14.4) \quad n = 2, \quad p = 1, \quad q = 1$$

hat.

15. Die Mayerschen Probleme.

Eine andere Art von Lagrangeschen Problemen, deren Behandlung bekanntlich keine Schwierigkeiten bereitet, sind die sogenannten Mayerschen Probleme, *solange diese von der Klasse Eins sind.*

Die Mayerschen Probleme sind solche, bei denen die Funktion $L(t, \dot{x}_i)$ die Gestalt hat

$$(15.1) \quad L = U_t + U_{x_i} \dot{x}_i.$$

Definiert man dann nach (3.2) die Funktion M , so hat man weiter mit Benutzung von (15.1)

$$(15.2) \quad y_i = M_{\dot{x}_i} = U_{x_i} + \mu_{k'} \frac{\partial G_{k'}}{\partial \dot{x}_i}$$

$$(15.3) \quad H = -L + y_i \dot{x}_i = -U_t + \mu_{k'} \frac{\partial G_{k'}}{\partial \dot{x}_i} \dot{x}_i.$$

Setzt man also

$$(15.4) \quad \bar{y}_i = y_i - U_{x_i}, \quad \bar{H} = H + U_t,$$

so muß man, um $\bar{H}$ zu berechnen, die $\dot{x}_i$, $\mu_{k'}$ als Funktionen von $(t, x_i, \bar{y}_i)$ aus den Gleichungen

$$(15.5) \quad \bar{y}_i = \mu_{k'} \frac{\partial G_{k'}}{\partial \dot{x}_i}, \quad G_{k'} = 0$$

entnehmen, und hierauf

$$(15.6) \quad \bar{H} = \bar{y}_i \dot{x}_i$$

setzen. Aus den Gleichungen (15.5) folgt, daß die $\dot{x}_i$ als homogene Funktionen von der nullten Ordnung in den $\bar{y}_j$ erscheinen, und daher muß $\bar{H}$ homogen erster Ordnung in diesen Größen sein. Man hat also

$$(15.7) \quad \bar{H} = \bar{y}_j \frac{\partial \bar{H}}{\partial \bar{y}_j};$$

nun ist aber auch

$$(15.8) \quad \frac{\partial \bar{H}}{\partial \bar{y}_j} = H_{y_j},$$

und man kann statt der Gleichung (15.7) mit Benutzung von (15.4) und (15.8) schreiben

$$(15.9) \quad H = -U_t + (y_j - U_{x_j}) H_{y_j}.$$

Die Identität (15.9) ist andererseits charakteristisch dafür, daß das vorliegende Problem ein Mayersches sei; denn die Gleichung (15.1) folgt aus

$$(15.10) \quad L = -H + y_i H_{y_i}, \quad \dot{x}_i = H_{y_i},$$

sobald (15.9) besteht.

16. Es ist sehr leicht zu zeigen, daß die Klasse der Mayerschen Probleme, wie zu erwarten ist, immer von Null verschieden ist. Differenziert man nämlich die Identität (15.9) nach x_i und y_j , so erhält man

$$(16.1) \quad H_{x_i} = -U_{tx_i} - U_{x_i x_j} H_{y_j} + (y_j - U_{x_j}) H_{x_i y_j}$$

$$(16.2) \quad 0 = (y_j - U_{x_j}) H_{y_i y_j}.$$

Setzt man also jetzt, um Anschluß an die Bezeichnungen der §§ 8-12 zu bekommen,

$$(16.3) \quad \eta_i = y_i - U_{x_i},$$

so folgt durch Differentiation längs einer Extremalen

$$\begin{aligned} \dot{\eta}_i &= \dot{y}_i - U_{tx_i} - U_{x_i x_j} \dot{x}_j \\ &= -H_{x_i} - U_{tx_i} - U_{x_i x_j} H_{y_j}, \end{aligned}$$

also mit Berücksichtigung von (16.1) und (16.3)

$$(16.4) \quad \dot{\eta}_i = -\eta_j H_{x_i y_j}.$$

Andererseits hat man nach (16.2) und (16.3)

$$(16.5) \quad \eta_j H_{y_i y_j} = 0.$$

Sind also die Größen η_i nicht alle identisch Null, so besagen die beiden letzten Gleichungen, daß die Differentialgleichungen (10.1) Lösungen besitzen, die den Bedingungen (10.2) genügen. Hier ist also immer $q > 0$.

Die einzigen Extremalen auf welchen die vorige Schlußweise nicht anwendbar ist, sind diejenigen, längs welchen alle $y_i = U_{x_i}$ sind; diese werden aber öfters, wie das folgende Beispiel zeigt, auch aus anderen Gründen ausgeschlossen.

17. *Beispiel.*

Es sei eine einzige Differentialgleichung, nämlich

$$(17.1) \quad G = \sum_1^n \dot{x}_i^2 - 1 = 0$$

gegeben. Man hat dann

$$(17.2) \quad \bar{y}_i = 2\mu \dot{x}_i$$

und

$$\bar{H} = \bar{y}_i \dot{x}_i = 2\mu.$$

Andererseits folgt aus (17.2) und (17.1)

$$\Sigma \bar{y}_i^2 = 4u^2$$

und es ist daher

$$(17.3) \quad \bar{H} = \sqrt{\Sigma \bar{y}_i^2},$$

also nach (15.4)

$$(17.4) \quad H = -U_i + \sqrt{\Sigma (y_i - U_{x_i})^2}.$$

Für die Legendresche Bedingung hat man nun nach Geod. Äquid. p. 220.

$$D(\varrho) = D(0) \mid \delta_{ij} - \varrho H_{y_i y_j} \mid$$

und man findet nach einigen Rechnungen

$$(17.5) \quad D(\varrho) = D(0) \left(1 - \frac{\varrho}{\bar{H}}\right)^{n-1}.$$

Die Legendresche Bedingung ist also dann und nur dann erfüllt, wenn $\bar{H} > 0$ ist, woraus folgt, daß nicht alle $y_i = U_{x_i}$ sein dürfen.

18. *Höhere Singularitäten.*

Die Lagrangeschen Probleme der Variationsrechnung, die nicht unter die schon genannten Kategorien fallen, sind fast gar nicht erforscht worden.

Das einzige mir bekannte derartige Beispiel, das in allen Einzelheiten behandelt ist, befindet sich in meiner Dissertation⁸⁾. Es handelt sich um folgendes Problem:

Eine Kugel wird von ihrem Mittelpunkt aus auf eine ihrer Tangentialebenen projiziert; man verlangt zwei Punkte dieser Ebene durch eine Kurve von gegebener Länge zu verbinden, die den Schatten einer möglichst kurzen sphärischen Kurve darstellt.

Die Lösung besteht hier immer aus zwei geradlinigen einen Winkel bildenden Strecken, wobei die Winkelhalbierende stets durch den Berührungspunkt der Kugel mit der Ebene hindurchgeht.

⁸⁾ C. Carathéodory. Ueber die diskontinuierlichen Lösungen in der Variationsrechnung. (Gött. Diss. 1904) p. 50—62.

Als Lagrangesches Problem aufgefasst, ist die Klasse dieses Problems gleich Eins, solange man nur kontinuierliche Extremalen betrachtet; durch die Einführung der diskontinuierlichen Lösungen geht nun das Problem in ein solches von der Klasse Null über.

19. Es gibt aber auch Fälle, in denen auch die Betrachtung diskontinuierlicher Lösungen nichts nützen kann. Nehmen wir z. B. die Hamiltonsche Funktion

$$(19.1) \quad H = \frac{(y_1 - f(t) x_1 y_2)^2}{2(1 - y_2)} - \frac{\dot{f}(t) x_1^2 y_2}{2}$$

an. Man setze zur Abkürzung

$$(19.2) \quad \frac{y_1 - f(t) x_1 y_2}{1 - y_2} = \alpha.$$

Die Differentialgleichungen der Extremalen lauten dann

$$(19.3) \quad \dot{x}_1 = H_{y_1} = \alpha$$

$$(19.4) \quad \dot{x}_2 = H_{y_2} = \frac{\alpha^2}{2} - f(t) x_1 \alpha - \dot{f}(t) \frac{x_1^2}{2}$$

$$(19.5) \quad \dot{y}_1 = -H_{x_1} = f(t) y_2 \alpha + \dot{f}(t) y_2 x_1$$

$$(19.6) \quad \dot{y}_2 = -H_{x_2} = 0.$$

Man hat hier in der Tat $p = 1$, weil aus (19.3) und (19.4) folgt

$$(19.7) \quad \dot{x}_2 = \frac{\dot{x}_1^2}{2} - f(t) x_1 \dot{x}_1 - \dot{f}(t) \frac{x_1^2}{2}.$$

Andererseits entnimmt man aus (19.3) und (19.5)

$$\dot{y}_1 = y_2 \frac{d}{dt} (f x_1);$$

da nun nach (19.6) y_2 konstant ist, folgt aus der letzten Gleichung

$$(19.8) \quad y_1 - f x_1 y_2 = \text{const.}$$

und es ist daher auch α eine Konstante. Jetzt kann man, wegen (19. 3) schreiben

$$(19. 9) \quad x_1 = \alpha t + \beta$$

und aus (19. 7) folgt dann die Berechnung von $x_2(t)$ durch eine Quadratur. Aus dem letzten Resultat sieht man, daß die Gesamtheit der Extremalen durch einen Punkt eine zweidimensionale Fläche bilden, und daß daher $q = 1$ ist.

Diskontinuierliche Lösungen sind hier nicht vorhanden; aus (19. 2) bis (19. 4) folgt nämlich, daß, wenn man y_1 und y_2 vorgibt, die Größen $\dot{x}_1$ und x_2 *eindeutig* bestimmt werden.

Endlich sieht man daß die Legendresche Bedingung immer erfüllt ist, wenn $y_2 < 1$ ist, denn man berechnet ganz ähnlich wie im § 17

$$D(\varrho) = D(0) \left(1 - \frac{(1 - y_2)^2 + (y_1 - f x_1)^2}{(1 - y_2)^3} \varrho \right).$$

Eine eingehendere Untersuchung scheint also hier notwendig zu sein.

(Eingegangen den 27. Februar 1932)

Ueber eine Eigenschaft der Raumkurven

Von W. SCHERRER, Bern

Die Eigenschaft der Schraubenlinie, auf einem ausgezeichneten Zylinder zu verlaufen, scheint auf den ersten Blick von sehr individueller Natur zu sein. Es besteht aber die überraschende Tatsache, daß jeder beliebigen Raumkurve in eindeutiger Weise eine ausgezeichnete abwickelbare Regelfläche zugeordnet werden kann, welche im Falle der Schraubenlinie in den Schraubenzylinder übergeht.

Bei der Herleitung dieser Beziehung leistet die Vektoralgebra vorzügliche Dienste. Ich bezeichne in der üblichen Weise Vektoren mit deutschen Buchstaben $a, b, c, \dots, x, y$ und speziell das skalare Produkt mit $a b$, das vektorielle mit $[a b]$.

Wir betrachten nun eine beliebige Raumkurve C , wählen auf ihr die Bogenlänge s als Parameter und stellen ihren Verlauf in bezug auf ein rechtwinkliges Koordinatensystem dar durch den vom Ursprung O zum laufenden Punkt $P(s)$ führenden Ortsvektor $r(s)$.

Zu dem beliebigen aber bestimmten $P_0 = P(s_0)$ auf C konstruieren wir jetzt auf folgende Weise eine neue Raumkurve $\bar{C}(s_0)$: Wir ziehen auf der Tangentenfläche von C eine zu den Tangenten orthogonale Trajektorie, welche durch P_0 läuft. Sie sei bezeichnet mit C^* . Hierauf errichten wir in jedem Punkte P^* von C^* die Normalebene m der durch P^* laufenden Erzeugenden der Tangentenfläche. Wir erhalten so eine einparametrische Schar von Ebenen und die von ihr eingehüllte Raumkurve $\bar{C}(s_0)$ soll die neue Raumkurve sein, welche den Mittelpunkt der nachfolgenden Betrachtungen bildet.

Es wird sich zeigen, daß $\bar{C}(s_0)$ ebenfalls durch P_0 läuft, Ich könnte also die eben beschriebene Konstruktion wiederholen, indem ich nun in bezug auf $\bar{C}(s_0)$ als Ausgangskurve durch P_0 eine analoge Kurve $\bar{\bar{C}}(s_0)$ bestimmen würde. Es wird sich aber herausstellen, daß $\bar{\bar{C}}(s_0)$ wieder mit der ursprünglichen Kurve C *zusammenfällt*. Ich nenne daher $\bar{C}(s_0)$ die zu C *reziproke Kurve* durch P_0 .

Variiert man nun den Punkt P_0 , also s_0 in $\bar{C}(s_0)$, so erhält man die Gesamtheit der zu C reziproken Raumkurven, die offenbar eine Fläche F bildet. Greife ich eine bestimmte Kurve $\bar{C}(s_0)$ heraus, so kann ich in analoger Weise die Gesamtheit der zu ihr reziproken Raumkurven zu

einer Fläche $\bar{F}(s_0)$ zusammenfassen. Mit dem Parameter s_0 deuten wir an, daß bei Abänderung von s_0 eine Veränderung von $\bar{F}(s_0)$ zu erwarten ist, da ja $\bar{C}(s_0)$ in eine andere Kurve übergeht. Wir werden aber im Gegensatz zu dieser Vermutung als Hauptergebnisse unserer Untersuchung folgende Aussagen gewinnen:

1. Die Fläche $\bar{F}(s_0)$ hängt nicht vom Parameter s_0 ab, sondern fällt für jeden Wert von s_0 mit der ursprünglichen Fläche F zusammen.
2. Die somit einzig vorhandene Fläche F ist eine abwickelbare Regelfläche.

Um den Beweis für diese Behauptungen zu führen, suchen wir den expliziten Ausdruck für den Ortsvektor des laufenden Punktes der reziproken Kurve $\bar{C}(s_0)$. Zu dem Zwecke führen wir das begleitende Dreibein des laufenden Punktes $P = P(s)$ auf der ursprünglichen Kurve C ein. Tangente, Normale, Binormale, Krümmung und Torsion in diesem Punkte seien bezeichnet mit $\mathbf{t}$, $\mathbf{n}$, $\mathbf{b}$, κ und τ . Neben der Gleichung

$$\mathbf{r}'(s) = \mathbf{t} \quad (1)$$

gelten dann die Formeln von Frenet:

$$\begin{aligned} \mathbf{t}' &= -\kappa \mathbf{n} \\ \mathbf{n}' &= \kappa \mathbf{t} + \tau \mathbf{b} \\ \mathbf{b}' &= -\tau \mathbf{n} \end{aligned} \quad (2)$$

Außerdem führen wir den Darboux'schen Vektor δ der Momentandrehung des begleitenden Dreibeins ein:

$$\delta = \tau \mathbf{t} + \kappa \mathbf{b}. \quad (3)$$

Die durch den Punkt $P_0 = P(s_0)$ laufende Trajektorie C^* ist zugleich eine Fadenevolvente von C . Der Ortsvektor $\mathbf{r}^*$ ihres laufenden Punktes P^* ist also als Funktion von s gegeben durch

$$\mathbf{r}^* = \mathbf{r}(s) - (s - s_0) \mathbf{t}(s).$$

Die Gleichung der Normalebene zu der durch P^* laufenden Erzeugenden der Tangentenfläche lautet:

$$\mathbf{t}(s) \cdot \mathbf{r} = \mathbf{t}(s) \cdot \mathbf{r}(s) - (s - s_0).$$

Hier ist $\mathbf{r}$ der laufende Punkt dieser Ebene.

Wir wollen nun im folgenden den Parameter s als Argument nur noch schreiben, wenn es für das bessere Verständnis notwendig erscheint.

Die Punkte γ der gesuchten Hüllkurve $\bar{C}(s_0)$ müssen nun folgenden Bedingungen genügen:

$$t\gamma = t\gamma - (s - s_0)$$

$$n\gamma = n\gamma$$

$$b\gamma = b\gamma - x\tau^{-1}(s - s_0).$$

Die beiden letzten dieser Gleichungen erhält man in bekannter Weise aus der ersten durch zweimalige Ableitung nach s bei festem γ unter Berücksichtigung der Formeln (2) von Frenet. Liegen allgemein drei derartige Gleichungen vor:

$$a\gamma = a$$

$$b\gamma = b$$

$$c\gamma = c,$$

so lautet ihre Auflösung in Vektorform:

$$\gamma = \frac{a[b, c] + b[c, a] + c[a, b]}{a[b, c]}.$$

Die Anwendung auf unsern Fall liefert bei Beachtung der Relationen

$$[n, b] = t; [b, t] = n; [t, n] = b$$

und der Definition (3) die Lösung:

$$\gamma = \gamma(s) - \tau^{-1}(s) \cdot (s - s_0) \delta(s).$$

Hält man in dieser Darstellung s_0 fest, so liefert sie die durch $P_0 = P(s_0)$ laufende Reziproke $\bar{C}(s_0)$, als Funktion der Bogenlänge von C . Variiert man aber auch s_0 , so erhält man die am Anfang beschriebene Fläche F . Man sieht unmittelbar, daß F eine Regelfläche darstellt. Hält man nämlich s fest und variiert s_0 , so resultiert eine Gerade durch den Punkt $P(s)$ (deren Ortsvektor ja $\gamma(s)$ ist), deren Richtung durch den Darboux'schen Vektor $\delta(s)$ in diesem Punkte gegeben ist.

Wir wollen nun zeigen, daß man dieselbe Fläche erhält, wenn man eine der reziproken $\bar{C}(s_0)$ herausgreift und zu ihr in entsprechender Weise die Fläche $\bar{F}(s_0)$ konstruiert. Zur Vereinfachung der Rechnung können wir annehmen, daß P_0 dem Anfangspunkt der Parameterzählung, also dem Werte $s_0 = 0$ entspreche und wollen diesen Punkt von nun an mit O bezeichnen. Die zugehörige Reziproke $\bar{C}(0)$ bezeichnen wir kurz mit C . Ihre Gleichung lautet, wenn wir mit $\bar{\mathfrak{r}}$ ihren Ortsvektor bezeichnen:

$$\bar{\mathfrak{r}}(s) = \mathfrak{r}(s) - s \cdot \tau^{-1} \delta(s).$$

Ihr Dreibein, ihre Bogenlänge, Krümmung und Torsion bezeichnen wir entsprechend mit $\bar{\mathfrak{t}}$, $\bar{\mathfrak{n}}$, $\bar{\mathfrak{b}}$, $\bar{x}$ und $\bar{\tau}$. Um nun die analoge Konstruktion durchzuführen, müssen wir in erster Linie die Bogenlänge $\bar{s}$ ermitteln. Wenn man in (5) für δ den Ausdruck $\tau t + x b$ wieder einführt, erhält man leicht:

$$\frac{d\bar{\mathfrak{r}}}{ds} = -\{s \cdot \tau^{-1}(s) \ x(s)\}' b.$$

Nun folgt aus $\left(\frac{d\bar{s}}{ds}\right)^2 = \left(\frac{d\bar{\mathfrak{r}}}{ds}\right)^2$ die gesuchte Bogenlänge:

$$\bar{s} = s \cdot \tau^{-1}(s) \ x(s).$$

Jetzt erhält man $\bar{\mathfrak{t}}$ aus

$$\bar{\mathfrak{t}} = \frac{d\bar{\mathfrak{r}}}{d\bar{s}} = \frac{d\bar{\mathfrak{r}}}{ds} : \frac{d\bar{s}}{ds}$$

und in analoger Weise alle übrigen zu $\bar{\mathfrak{t}}$ gehörigen Größen. Wir stellen die sämtlichen so sich ergebenden Ausdrücke in einer Tabelle zusammen:

$\bar{s}$	$= s \tau^{-1} x$
$\bar{\mathfrak{t}}$	$= -b$
$\bar{\mathfrak{n}}$	$= n$
$\bar{\mathfrak{b}}$	$= t$
$\bar{x}$	$= \tau \{(s \tau^{-1} x)'\}^{-1}$
$\bar{\tau}$	$= -x \cdot \{(s \tau^{-1} x)'\}^{-1}$
$\bar{\delta}$	$= \{(s \tau^{-1} x)'\}^{-1} \delta.$

In diesen Gleichungen sind also alle rechts vorkommenden Größen Funktionen von s und der Strich bedeutet die Ableitung nach s .

Nun können wir leicht erkennen, daß die Fläche $\bar{F}$ mit F zusammenfällt. Die Gleichung (5) gestattet nämlich folgende Interpretation: Der dem Punkt $\mathfrak{r}(s)$ von C entsprechende Punkt $\bar{\mathfrak{r}}(s)$ von $\bar{C}$ liegt auf der durch $\mathfrak{r}(s)$ laufenden Erzeugenden von F . Durch den Punkt $\bar{\mathfrak{r}}$ geht andererseits eine Erzeugende von $\bar{F}$. Nach der letzten Gleichung von (6) fällt aber ihre Richtung mit derjenigen der ersten Erzeugenden zusammen, F und $\bar{F}$ haben somit dieselben Erzeugenden, sind also identisch.

Die Wahl des Ausgangspunktes O war willkürlich, entsprechend der Festsetzung des Nullpunktes für s . Also gilt allgemein $\bar{F}(s_0) = F$. Damit ist die erste Behauptung bewiesen.

Nun können wir leicht beweisen, daß die gefundene Regelfläche F abwickelbar ist. Wir brauchen nur zu zeigen, daß die Flächennormale längs einer Erzeugenden eine konstante Richtung aufweist. Wir bestimmen zunächst die Flächennormale im Punkte O ($s = 0$). Hier gilt:

$$\mathfrak{N} = [\mathfrak{t}, \mathfrak{i}] = [\mathfrak{t}, -\mathfrak{b}] = \mathfrak{n}.$$

Also fällt für jeden Punkt von C die Flächennormale $\mathfrak{N}$ mit der Kurvennormale $\mathfrak{n}$ zusammen. Zuzufolge der gefundenen Reziprozität gilt dasselbe auch längs $\bar{C}$ und somit schließlich für alle Reziproken. Betrachten wir nun speziell den Punkt $\mathfrak{r}(s)$ auf C und den ihm auf den zugehörigen Erzeugenden von F entsprechenden Punkt $\bar{\mathfrak{r}}(s)$ auf $\bar{C}$ (siehe (5)). Die dritte der Gleichungen (6) ergibt $\bar{\mathfrak{n}} = \mathfrak{n}$ und nach dem obigen also $\bar{\mathfrak{N}} = \mathfrak{N}$. Halte ich nun den Punkt $\bar{\mathfrak{r}}(s)$ fest, verschiebe hingegen O auf C , so durchläuft der Punkt $\bar{\mathfrak{r}}(s)$ die ganze zu $\mathfrak{r}(s)$ gehörige Erzeugende und immer ist die jeweilige Flächennormale $\bar{\mathfrak{N}}$ gleich der festen Normalen $\mathfrak{N}$ in $\mathfrak{r}(s)$. Damit ist die aufgestellte Behauptung bewiesen.

Man kann noch bemerken, daß zufolge der Relation $\mathfrak{t}\bar{\mathfrak{t}} = -\mathfrak{t}\mathfrak{b} = 0$ die Reziproken sich gegenseitig rechtwinklig durchsetzen.

Im Falle der gewöhnlichen Schraubenlinie sind x und τ konstant und man bestätigt leicht, daß die gefundene Fläche F den Schraubenzyylinder darstellt.

(Eingegangen den 7. März 1932)

Geometrische Deutung des Gauß'schen Verschlingungsintegrals

Von W. SCHERRER, Bern

Sind $x_1(u)$, $x_2(u)$, $x_3(u)$ und $y_1(v)$, $y_2(v)$, $y_3(v)$ die Parameterdarstellungen zweier geschlossener Raumkurven C und D , so stellt bekanntlich das von Gauß eingeführte Integral

$$\iint_{C,D} \frac{(y_1 - x_1)(dx_2 dy_3 - dx_3 dy_2) + (y_2 - x_2)(dx_3 dy_1 - dx_1 dy_3) + (y_3 - x_3)(dx_1 dy_2 - dx_2 dy_1)}{[(y_1 - x_1)^2 + (y_2 - x_2)^2 + (y_3 - x_3)^2]^{\frac{3}{2}}} du dv$$

eine Invariante gegenüber stetigen Transformationen von C und D dar, deren Nichtverschwinden mit Sicherheit auf das Vorhandensein einer Verschlingung schließen läßt, während das Umgekehrte nicht zu gelten braucht.

Dieses Integral läßt sich in einfacher Weise als Raumwinkel einer Fläche in bezug auf einen Punkt deuten. Um die Verhältnisse bequemer zu beschreiben, bediene ich mich der Vektorsymbolik und setze

$$\begin{aligned} \mathfrak{x}(u) &= (x_1(u), x_2(u), x_3(u)) \\ \mathfrak{y}(v) &= (y_1(v), y_2(v), y_3(v)) \end{aligned}$$

Das Integral läßt sich jetzt auch schreiben in der Form:

$$\iint_{C,D} \frac{[\mathfrak{y}(v) - \mathfrak{x}(u)] \cdot [\mathfrak{x}'(u), \mathfrak{y}'(v)]}{\{(\mathfrak{y}(v) - \mathfrak{x}(u))^2\}^{\frac{3}{2}}} du dv.$$

Diesem Integral stelle ich nun gegenüber den Raumwinkel Ω einer beliebigen Fläche in bezug auf den Koordinatenursprung. Bedeutet $\mathfrak{x}(u, v)$ den vom Ursprung aus zum laufenden Punkt auf der Fläche führenden Ortsvektor als Funktion der Flächenparameter u und v , so ist der gesuchte Raumwinkel Ω gegeben durch die Gleichung:

$$\Omega = \iint_F \frac{\mathfrak{x}(u, v) [\mathfrak{x}_u(u, v), \mathfrak{x}_v(u, v)]}{\{(\mathfrak{x}(u, v))^2\}^{\frac{3}{2}}} du dv.$$

Setze ich in diesen Ausdruck für $\mathfrak{x}(u, v)$ den Wert $\mathfrak{y}(v) - \mathfrak{x}(u)$ ein, so erhalte ich abgesehen von einem festen und an sich willkürlichen Vorzeichen genau das Gauß'sche Verschlingungsintegral.

Wir haben uns nur noch Rechenschaft zu geben über die Bedeutung der durch

$$\mathfrak{x}(u, v) = \mathfrak{y}(v) - \mathfrak{x}(u)$$

dargestellten Fläche. Offenbar kann sie aufgefaßt werden als Schiebefläche:

$$\mathfrak{x}(u, v) = \mathfrak{y}(v) + (-\mathfrak{x}(u)).$$

Sie entsteht also folgendermaßen: Man spiegle die durch die Gleichung $\mathfrak{x} = -\mathfrak{x}(u)$ definierte Kurve C am Koordinatenursprung O . Man erhält so die der Gleichung $\mathfrak{x} = \mathfrak{x}(u)$ entsprechende Kurve $\bar{C}$. Nun verschiebe man diese Kurve parallel mit sich selbst und starr verbunden mit dem zu Beginn der Bewegung loszulösenden Nullpunkt soweit, bis der mitgeführte Nullpunkt $\bar{O}$ auf die Kurve D zu liegen kommt. Schließlich führe man den Punkt $\bar{O}$ der Kurve D entlang und nehme dabei die Kurve $\bar{C}$ starr in bezug auf $\bar{O}$ und parallel zu sich selbst mit. Diese Schiebefläche bildet gewissermaßen einen die Kurve D begleitenden Schlauch. Willkürlich bei dieser Konstruktion sind die Lage des Ursprungs und die Reihenfolge der Kurven.

Am einfachsten liegen die Verhältnisse, wenn eine der Kurven ein Zentrum hat, welches man dann als Nullpunkt wählt. Man betrachte etwa einen horizontalen Kreis C und einen zweiten D , der durch das Zentrum von C geht und senkrecht zur Ebene von C steht.

Daß der Wert des Gauß'schen Integrals ein Vielfaches von 4π ist, erkennt man nun ohne weiteres, ebenso die Invarianz gegenüber stetigen Transformationen.

Auch die in dem Enzyklopädieartikel von *Dehn* und *Heegaard* über Analysis situs angedeutete Verallgemeinerung von *Dyck* ergibt sich auf diesem Wege unmittelbar. Sind nämlich.

$$\mathfrak{x}(u_1, u_2, \dots, u_p) = (x_1(u_1, \dots, u_p), \dots, x_n(u_1, \dots, u_p))$$

$$\mathfrak{y}(v_1, v_2, \dots, v_q) = (y_1(v_1, \dots, v_q), \dots, y_n(v_1, v_2, \dots, v_q))$$

die Parameterdarstellungen zweier Mannigfaltigkeiten C und D im R_n mit den Dimensionszahlen p und q , die der Bedingung $p + q = n - 1$ genügen, so bilde man die Determinante

$$\Delta = \begin{vmatrix} y_1 - x_1, y_2 - x_2, \dots, y_n - x_n \\ \frac{\partial x_1}{\partial u_1}, \frac{\partial x_2}{\partial u_1}, \dots, \frac{\partial x_n}{\partial u_1} \\ \vdots \\ \frac{\partial x_1}{\partial u_p}, \frac{\partial x_2}{\partial u_p}, \dots, \frac{\partial x_n}{\partial u_p} \\ \frac{\partial y_1}{\partial v_1}, \frac{\partial y_2}{\partial v_1}, \dots, \frac{\partial y_n}{\partial v_1} \\ \vdots \\ \frac{\partial y_1}{\partial v_q}, \frac{\partial y_2}{\partial v_q}, \dots, \frac{\partial y_n}{\partial v_q} \end{vmatrix}$$

und die Distanz

$$r = \sqrt{(y_1 - x_1)^2 + (y_2 - x_2)^2 + \dots + (y_n - x_n)^2}.$$

Auf das Integral

$$\iint_{C \cdot D} \frac{\Delta}{r^n} du_1 du_2 \dots du_p \cdot dv_1 dv_2 \dots dv_q$$

läßt sich dann alles oben Gesagte übertragen.

(Eingegangen den 7. März 1932)

Ueber die Kreise die von einer Riemannschen Fläche schlicht überdeckt werden

Von LARS AHLFORS, Åbo (Finnland)

1. Einleitung

In den *Comptes Rendus*¹⁾ habe ich neulich als Verallgemeinerung des bekannten Blochschen Satzes bewiesen, daß von *fünf* punktfremden Kreisen wenigstens einer die Eigenschaft besitzt von einem Blatt der von den Werten einer in der ganzen Ebene meromorphen Funktion gebildeten Riemannschen Fläche schlicht überdeckt zu werden. Wendet man diesen Satz auf eine ganze Funktion an, so sieht man daß dieselbe Behauptung schon mit *vier* im endlichen gelegenen Kreisen richtig ist.

In dieser Arbeit werde ich zeigen, daß die Anzahl der fraglichen Kreise für ganze Funktionen in der Tat auf *drei* reduziert werden kann, wie schon Bloch selbst vermutet hat²⁾. Allgemeiner werde ich eine im Einheitskreise E reguläre Funktion $f(z)$ mit $f(0) = 0$, $|f'(0)| = 1$ betrachten und eine von $f(z)$ unabhängige Ungleichung herleiten, welche eine hinreichende Bedingung darstellt, damit von drei gegebenen Kreisen wenigstens einer innerhalb E schlicht angenommen werde. Aus dieser Ungleichung folgt besonders der Blochsche Satz mit einem numerischen (übrigens sehr ungenauen) Wert der Blochschen Konstante. Es geht ferner hervor, daß der Blochsche Kreis innerhalb eines um den Nullpunkt geschlagenen Kreises von festem Radius gewählt werden kann.

Auf die möglichen Erweiterungen meiner Resultate gehe ich in dieser Arbeit nicht ein, um so mehr da ich die Absicht habe, in einer umfassenderen Abhandlung auf diese und ähnliche Fragen zurückzukommen.

2. Vorbereitende Betrachtungen

Wir betrachten in der w -Ebene die drei außerhalb einander gelegenen Kreise C_1 , C_2 , C_3 und ziehen ihren gemeinsamen Orthogonalkreis $\bar{C}$ mit dem Mittelpunkt w_0 und dem Radius R . Von diesem Orthogonalkreis schneidet der Kreis C_i ($i = 1, 2, 3$) einen Bogen ab, dessen Zentriwinkel mit α_i bezeichnet werde. Zwischen C_1 und C_2 fällt ein Bogen $\delta_{1,2}$ von

¹⁾ t. 194, p. 1145.

²⁾ La conception actuelle de la théorie des fonctions entières et méromorphes. L'Enseignement Mathématique, 1926, p. 87.

$\bar{C}$, zwischen C_2 und C_3 liegt $\delta_{2,3}$ und zwischen C_3 und C_1 der Bogen $\delta_{3,1}$. Mit $\delta_{i,j}$ soll gleichzeitig der Zentriwinkel des entsprechenden Bogens bezeichnet werden. Schließlich wird noch die Bezeichnung

$$d = \text{Min} \left(\delta_{i,j}, \alpha_i, \frac{\pi - \alpha_i}{3} \right)$$

eingeführt für die kleinste der neun eingeklammerten Größen. Es ist offenbar $0 < d \leq \frac{\pi}{4}$.

Die Funktion $f(z) = z + \dots$ ³⁾ sei regulär in E . Wir bezeichnen mit Δ_i die Gebiete, wo die Funktion $w = f(z)$ zum Kreise C_i gehörige Werte annimmt. Falls ein Gebiet Δ_i ganz im inneren von E gelegen ist, so liegt in diesem Δ_i für jeden zu C_i gehörigen Wert w dieselbe endliche Anzahl von Wurzeln der Gleichung $f(z) = w$. Ich nehme an, daß diese Anzahl für alle ganz innerhalb E gelegenen Gebiete Δ_1 , Δ_2 und Δ_3 größer als Eins ist, und werde aus dieser Annahme eine Ungleichung zwischen $|w_0|$, R und d herleiten. Falls diese Ungleichung nicht erfüllt ist, so ist dies folglich ein Zeichen, daß es im inneren von $|z| < 1$ ein Gebiet Δ_i ($i = 1, 2$ oder 3) gibt, wo $f(z)$ jeden zum entsprechenden Kreise C_i gehörigen Wert genau einmal annimmt.

Die Punkte der w -Ebene projizieren wir stereographisch auf die Oberfläche einer die w -Ebene im Punkte w_0 berührenden Kugel vom Durchmesser R . Der Kreis $\bar{C}$ geht dabei in den Äquatorialkreis über, und den Kreisen C_i entsprechen auf der Kugelfläche drei in Bezug auf die Äquatorialebene symmetrische Kreise. Um den Punkt ∞ zeichnen wir noch einen vierten Kreis C_∞ , dessen Durchmesser ein Bogen vom Zentriwinkel d ist. Aus der Definition von d geht hervor, daß der Abstand von C_∞ zu den Kreisen C_i wenigstens gleich d ist. Dem Kreise C_∞ entsprechen in der z -Ebene die Gebiete Δ_∞ , wo die Werte von $f(z)$ innerhalb C_∞ fallen. Diese Gebiete strecken sich alle bis zum Rande von E .

Die verschiedenen Gebiete Δ_1 , Δ_2 und Δ_3 sind durch Kurven $\gamma_{i,j}$ verbunden, auf welchen die Werte von $f(z)$ zu dem zwischen C_i und C_j liegenden Bogen $\delta_{i,j}$ von $\bar{C}$ gehören. Auf dem Rande eines vollständig innerhalb E gelegenen Gebietes Δ_i nimmt $f(z)$ jeden auf der Peripherie von C_i gelegenen Wert wenigstens zweimal an. Folglich gehen von dem Gebiet Δ_i wenigstens zwei Kurven $\gamma_{i,j}$ und zwei Kurven $\gamma_{k,i}$ ($i \neq j \neq k$) aus, welche zu einem Gebiet Δ_j bzw. Δ_k führen⁴⁾. Von einem beliebigen

³⁾ Die Annahme $f(0) = 0$ ist natürlich unwesentlich und bedeutet nur eine Verschiebung der w -Ebene.

⁴⁾ Die Kurven $\gamma_{i,j}$ können beliebig oft verzweigt sein; es gibt aber immer einen Zweig, der zu einem Gebiet Δ_j führt.

inneren Gebiete Δ_i ausgehend läßt sich folglich eine Kette $\Gamma_{i,j}$ bilden, die aus abwechselnden Gebieten Δ_i und Δ_j verbunden durch Kurven $\gamma_{i,j}$ besteht. Diese Kette ist entweder endlos, oder endet mit einem Gebiet, das sich bis zum Rande von E streckt. Ebenso geht durch Δ_i eine Kette $\Gamma_{i,k}$.

Die Ketten $\Gamma_{i,j}$ können nicht geschlossen sein, d. h. man kommt nie zu einem Gebiet Δ_i oder Δ_j , das mit einem vorigen identisch ist. Man zeigt nämlich mit Hilfe des Argumentprinzips, daß die Funktion $f(z)$ in dem von einer geschlossenen Kette $\Gamma_{i,j}$ begrenzten Gebiet nur solche Werte annehmen könnte, die entweder sämtlich zu C_i oder sämtlich zu C_j gehören. In derselben Weise wird gezeigt, daß zwei Ketten $\Gamma_{i,j}$ und $\Gamma_{i,k}$ höchstens *ein* gemeinsames Gebiet Δ_i haben.

3. Zurückführung des Beweises auf den Beweis eines Hilfssatzes

Wenn $f(z)$ im Kreise $|z| < r_0 < 1$ keinen zu C_∞ gehörigen Wert annimmt, d. h. wenn in diesem Kreise $|f(z) - w_0| \leq \frac{R}{\operatorname{tg} \frac{4}{d}}$ gilt, so folgt

in bekannter Weise $r_0 \leq \frac{R}{\operatorname{tg} \frac{4}{d}} < \frac{4R}{d}$. Ist $f(z)$ in demselben Kreis von

allen zu C_i gehörigen Werten verschieden, so gilt, wenn der Mittelpunkt und Radius von C_i mit a_i und ϱ_i bezeichnet werden, $\left| \frac{1}{f(z) - a_i} \right| \leq \frac{1}{\varrho_i}$

für $|z| < r_0$. Die Ableitung der Funktion $\frac{1}{f(z) - a_i}$ im Nullpunkt ist gleich $-\frac{1}{a_i^2}$. Daraus folgt

$$\begin{aligned} r_0 \leq \frac{|a_i|^2}{\varrho_i} &\leq \frac{\left(|w_0| + \frac{R}{\cos \frac{\alpha_i}{2}} \right)^2}{R \operatorname{tg} \frac{\alpha_i}{2}} = \frac{2 \left(|w_0| \cos \frac{\alpha_i}{2} + R \right)^2}{R \sin \alpha_i} \\ &\leq \frac{2(|w_0| + R)^2}{R \sin d} < \frac{4(|w_0| + R)^2}{R d}. \end{aligned}$$

Im folgenden wählen wir $r_0 = \frac{4(|w_0| + R)^2}{R d}$, und nehmen an, daß

durch diese Wahl $r_0 < 1$ ausfällt. Da die Ungleichung $r_0 \geq \frac{4R}{d}$ *a fortiori* erfüllt ist, so sind wir sicher, daß es Gebiete aller vier Arten Δ_1 , Δ_2 , Δ_3 und Δ_∞ gibt, welche ganz oder zum Teil im Kreise $|z| < r_0$ liegen.

Ich fixiere nun ein Gebiet $\Delta_\infty^{(0)}$ der vierten Art, das vom Kreise $|z| = r_0$ getroffen wird, gehe von einem auf diesem Kreis gelegenen Punkt des betrachteten Gebietes aus und beschreibe die Kreisperipherie im positiven Sinn bis ich zum erstenmal auf eine Kette $\Gamma_{1,2}$, $\Gamma_{2,3}$ oder $\Gamma_{3,1}$ treffe. Es sei dies die Kette $\Gamma_{i,j}^{(0)}$. Zwischen $\Delta_\infty^{(0)}$ und $\Gamma_{i,j}^{(0)}$ liegt ein eindeutig bestimmtes Teilgebiet D_0 des Kreisringes $r_0 < |z| < 1$, das auf seinem Rande einen Teil des soeben im positiven Sinne durchlaufenen Kreisbogens enthält. Falls in D_0 kein Gebiet Δ_k ($i \neq j \neq k$) liegt, so setzen wir $\Omega_0 = D_0$ und das Verfahren bricht hiermit ab. Im entgegengesetzten Falle sei $\Delta_k^{(1)}$ dasjenige Gebiet Δ_k , dessen Abstand r_1 ($> r_0$) vom Nullpunkt am kleinsten ist. Wir ziehen den $\Delta_k^{(1)}$ berührenden Bogen des Kreises $|z| = r_1$ und bezeichnen mit Ω_0 das zwischen diesem Bogen und dem Kreise $|z| = r_0$ fallende Teilgebiet von D_0 .

Durch das Gebiet $\Delta_k^{(1)}$ geht eine Kette $\Gamma_{k,i}^{(1)}$, welche in der einen Richtung von $\Delta_k^{(1)}$ gerechnet ganz innerhalb D_0 verläuft und zur Peripherie des Einheitskreises führt. Das außerhalb Ω_0 gelegene Teilgebiet von D_0 , welches zwischen $\Gamma_{k,i}^{(1)}$ und $\Delta_\infty^{(0)}$ liegt, ist von derselben Art wie D_0 . Wir können also von diesem Gebiet ausgehend das auf D_0 angewandte Verfahren wiederholen.

Andererseits betrachten wir die von dem Gebiet $\Delta_k^{(1)}$ ausgehenden Kurven λ_k , welche der kürzesten Strecke zwischen C_k und C_∞ entsprechen. Es gibt eine dieser Kurven, welche innerhalb D_0 durch die Kette $\Gamma_{k,i}^{(1)}$ vom Kreise $|z| = r_0$ getrennt wird. Sie führt entweder zum Rande von E , oder zu einem neuen Gebiet Δ_∞ . Wir bezeichnen mit D_1 das außerhalb Ω_0 gelegene Teilgebiet von D_0 , das auf der einen Seite von $\Gamma_{i,j}^{(0)}$, auf der anderen Seite von $\Delta_k^{(1)}$ und den damit zusammenhängenden λ_k und Δ_∞ begrenzt wird. Wenn nun weder im inneren noch auf dem Rande von D_1 ein zu einem Δ_∞ gelegener Punkt liegt⁵⁾, so setzt man $\Omega_1 = D_1$ und das Verfahren bricht ab. Anderenfalls sei $\Delta_\infty^{(1)}$ das in D_1 gelegene oder an D_1 grenzende Gebiet Δ_∞ , dessen Abstand r_2 vom Nullpunkt am kleinsten ist. Durch einen Bogen des Kreises $|z| = r_2$, der das Gebiet $\Delta_\infty^{(1)}$ berührt, schneiden wir dann das Teilgebiet Ω_1 von D_1 ab⁶⁾.

⁵⁾ Dieser Fall kann nur eintreten, wenn λ_k zum Rande von E führt.

⁶⁾ Falls schon das auf $|z| = r_1$ gelegene Randstück von D_1 von einem Δ_∞ getroffen wird, so fällt der ganze Schritt aus.

Schließlich betrachten wir noch das außerhalb $\mathcal{Q}_1$ fallende Teilgebiet von D_1 , welches zwischen $\mathcal{A}_\infty^{(1)}$ und $\Gamma_{i,j}^{(0)}$ liegt, und wiederholen von diesem Gebiet ausgehend nochmals das auf D_0 angewandte Verfahren. Wir haben hierdurch einen endlichen oder unendlichen Algorithmus definiert, der zur Konstruktion eines Systems aneinandergelegter Gebiete $\mathcal{Q}$ führt. Jedes Gebiet $\mathcal{Q}$ wird von zwei Kreisbogen $|z| = r'$ und $|z| = r''$ und zwei die Endpunkte dieser Bogen verbindenden Kurven L_1 und L_2 begrenzt. Wiederholt man den ganzen Algorithmus, indem man von $\mathcal{A}_\infty^{(0)}$ ausgehend die Peripherie $|z| = r_0$ jetzt in *negativer* Richtung durchläuft, so erhält man ein neues, ähnliches System von Gebieten $\mathcal{Q}$.

Die Funktion $f(z)$ besitzt in jedem Gebiet $\mathcal{Q}$ eine der folgenden zwei Eigenschaften:

1. Die auf L_1 angenommenen Randwerte gehören zu C_∞ , und die Randwerte auf L_2 liegen auf C_i , C_j oder auf dem zwischenliegenden Bogen $\delta_{i,j}$, während $f(z)$ im inneren von $\mathcal{Q}$ keinen zu C_k ($i \neq j \neq k$) gehörigen Wert annimmt.

2. Die Randwerte auf L_1 gehören zu C_i , C_j oder $\delta_{i,j}$, und die Werte auf L_2 gehören zu C_k oder zu der kürzesten Strecke zwischen C_k und C_∞ . Im inneren nimmt $f(z)$ keinen zu C_∞ gehörigen Wert an.

Der erste Fall tritt für $\mathcal{Q}_0$, der zweite für $\mathcal{Q}_1$ ein.

Im nächsten Abschnitt werden wir einen Hilfssatz beweisen, aus dem sofort folgt, daß die Gebiete $\mathcal{Q}$ in beiden Fällen einer Ungleichung

$$(1) \quad \int_{r'}^{r''} \frac{dr}{r \theta(r)} < K$$

genügen müssen, wobei K eine Konstante und $r \theta(r)$ die Gesamtlänge der in $\mathcal{Q}$ liegenden Bogen des Kreises $|z| = r$ bezeichnen. Für jedes r ist offenbar $\Sigma \theta(r) \leq 2\pi$, wenn die Summe über alle vom Kreise $|z| = r$ getroffenen Gebiete $\mathcal{Q}$ erstreckt wird.

Aus (1) läßt sich die gesuchte Ungleichung leicht herleiten. Es sei in der Tat $n(r)$ die Anzahl der vom Kreise $|z| = r$ getroffenen Gebiete $\mathcal{Q}$. Dann ist, wie aus dem Aufbau des Algorithmus sofort hervorgeht, die Anzahl der ganz oder zum Teil in $|z| < r$ gelegenen Gebiete $\mathcal{Q}$ höchstens gleich $3n(r) - 4$. Wir erhalten folglich, wenn wir die zu diesen Gebieten gehörigen Ungleichungen (1) addieren

$$(2) \quad \int_{r_0}^r \left(\Sigma \frac{1}{\theta(t)} \right) \frac{dt}{t} < (3n(r) - 4) K,$$

wo die Summe wieder über alle vom Kreise $|z| = t$ getroffenen Gebiete $\mathcal{Q}$ zu erstrecken ist. Da die Anzahl dieser Gebiete gleich $n(t)$ ist, so erhält man nach dem Satz vom arithmetischen und harmonischen Mittel

$$\sum \frac{1}{\theta(t)} \geq \frac{n(t)^2}{\sum \theta(t)} \geq \frac{n(t)^2}{2\pi}.$$

Wird dies in (2) eingeführt, so findet man

$$\int_{r_0}^r n(t)^2 \frac{dt}{t} < (3n(r) - 4) 2\pi K$$

oder, wenn das linksstehende Integral mit $\alpha(r)$ bezeichnet wird,

$$(\alpha(r) + 8\pi K)^2 < 36\pi^2 K^2 \frac{d\alpha(r)}{d \log r},$$

welche Ungleichung für $r_0 < r < 1$ gültig ist. Durch Integration ergibt sich dann

$$\begin{aligned} \log \frac{1}{r_0} = \int_{r_0}^1 d \log r &< 36\pi^2 K^2 \int_{r_0}^1 \frac{d\alpha(r)}{(\alpha(r) + 8\pi K)^2} \\ &\leq 36\pi^2 K^2 \int_{\alpha=0}^{\infty} \frac{d\alpha}{(\alpha + 8\pi K)^2} = \frac{9\pi K}{2}, \end{aligned}$$

oder

$$r_0 > e^{-\frac{9}{2}\pi K}.$$

Führt man noch den Wert von r_0 ein (S. 356), so erhält man die endgültige Ungleichung

$$(3) \quad \frac{4(|w_0| + R)^2}{Rd} > e^{-\frac{9}{2}\pi K},$$

wo noch der Wert von K zu ermitteln ist.

Bei der Herleitung von (3) wurde vorausgesetzt, daß der auf der linken Seite stehende Ausdruck kleiner als 1 ist. Im entgegengesetzten Falle ist die Ungleichung trivial.

4. Der Hilfssatz

Auf der Riemannschen Kugel seien E_1 und E_2 zwei punktfremde Kontinuen und E_3 ein drittes Kontinuum, das sowohl mit E_1 als mit E_2 wenigstens einen gemeinsamen Punkt besitzt. In einem Gebiet Ω der vorhin betrachteten Art sei eine meromorphe Funktion $f(z)$ gegeben, welche keinen zu E_3 gehörigen Wert annimmt und deren Randwerte auf L_1 zu E_1 und auf L_2 zu E_2 gehören. Dann gilt

$$(4) \quad \int_{\gamma}^r \frac{dr}{r \theta(r)} \leq \frac{8\pi}{\delta^2},$$

wo δ den Zentriwinkel des kleinsten, auf der Riemannschen Kugel gemessenen Abstandes zwischen E_1 und E_2 bezeichnet:

Vorbemerkung: Auch wenn E_3 keinen zu E_1 und E_2 gehörigen Punkt enthält, bekommt man durch diesen Satz eine Schranke für das Integral (4), vorausgesetzt daß E_1 und E_3 ohne Ueberschreitung von E_2 und ebenso E_2 und E_3 ohne Ueberschreitung von E_1 verbunden werden können. Nur bedeutet δ dann nicht mehr den Abstand zwischen E_1 und E_2 , sondern den Abstand der Kontinuen, die durch Hinzufügung der Verbindungskurven entstehen. Jedenfalls ist δ höchstens gleich dem Durchmesser von E_3 .

Beweis: Es sei G das von E_1 , E_2 und E_3 bestimmte Restgebiet der Riemannschen Kugel, das zu allen drei Kontinuen gehörige Randpunkte besitzt. Wir wählen einen beliebigen zu E_3 gehörigen Randpunkt von G und bezeichnen ihn mit W .

Der in Ω fallende Teil des Kreises $|z| = r$ ($r' < r < r''$) wird durch die Funktion $f(z)$ auf eine Kurve abgebildet, welche das Gebiet G in zwei oder mehrere Teilgebiete zerlegt. Es sei g_r dasjenige Teilgebiet, das den Randpunkt W besitzt, und γ_r der Teil des Randes von g_r , der die Kontinuen E_1 und E_2 verbindet. Schließlich bezeichnen wir noch mit ϑ_r den Teil des Kreises $|z| = r$, der auf γ_r abgebildet wird.

Bezeichnet R den Durchmesser der Riemannschen Kugel, so ist die Länge von γ_r wenigstens gleich $\frac{\delta R}{2}$. Man hat folglich

$$\int_{\mathfrak{F}_r} \frac{|f'(z)|}{1 + |f(z)|^2} |dz| \geq \frac{\delta R}{2},$$

woraus durch Anwendung der Schwarzschen Ungleichung folgt

$$\frac{\delta^2 R^2}{4} \leq \int_{\mathfrak{F}_r} |dz| \int_{\mathfrak{F}_r} \frac{|f'|^2}{(1 + |f|^2)^2} |dz| \leq r \theta(r) \int_{\mathfrak{F}_r} \frac{|f'|^2}{(1 + |f|^2)^2} |dz|.$$

Dividiert man durch $r \theta(r)$ und integriert zwischen den Grenzen r' und r'' , so wird

$$(5) \quad \frac{\delta^2 R^2}{4} \int_{r'}^{r''} \frac{dr}{r \theta(r)} \leq \int_{r'}^{r''} dr \int_{\mathfrak{F}_r} \frac{|f'|^2}{(1 + |f|^2)^2} |dz|.$$

Das rechtsstehende Doppelintegral stellt die Größe der von den Kurven γ_r erzeugten Fläche dar. Wir werden zeigen, daß durch jeden Punkt höchstens zwei Kurven γ_r hindurchgehen können, und daß die betrachtete Fläche folglich höchstens gleich $2\pi R^2$ ist⁷⁾.

Wir nehmen an, daß durch einen Punkt w die zwei Kurven γ_{r_1} und γ_{r_2} ($r' < r_1 < r_2 < r''$) hindurchgehen. Wir verbinden w mit W durch eine Kurve Γ , die innerhalb des gemeinsamen Gebietes von g_{r_1} und g_{r_2} verläuft, und betrachten den vom Punkte $f(z)$ beschriebenen Weg, wenn z die Randkurve des zwischen $|z| = r_1$ und $|z| = r_2$ gelegenen Teilgebietes von $\mathcal{Q}$ umläuft, wobei die Stellen wo $f(z) = w$ durch kleine, nach außen gerichtete Halbkreise zu umgehen sind. Es ist nun klar, daß der Punkt $f(z)$ die Kurve Γ nur dann überschreiten kann, wenn z einen dieser Halbkreise beschreibt. Also ist die Variation von $\arg \frac{f(z) - w}{f(z) - W}$ höchstens gleich 2π mal die Gesamtanzahl der auf $|z| = r_1$ und $|z| = r_2$ gelegenen Wurzeln der Gleichung $f(z) = w$. Hieraus folgt, daß zwischen diesen Kreisen keine Wurzel der genannten Gleichung liegen kann, womit die Richtigkeit unserer Behauptung erwiesen ist.

Aus (5) folgt nunmehr

$$\frac{\delta^2 R^2}{4} \int_{r'}^{r''} \frac{dr}{r \theta(r)} \leq 2\pi R^2,$$

d. h. die Ungleichung (4).

⁷⁾ Sie wird tatsächlich kleiner als $2\pi R^2$, wenn E_3 vom positiven Flächenmaß ist.

5. Zuendeführung des Beweises

Wir wenden jetzt den Hilfssatz auf die zwei auf S. 358 betrachteten Fälle an.

Im ersten Falle ist E_3 gleich C_k . Für E_1 kann man C_∞ und für E_2 das aus C_i , C_j und $\delta_{i,j}$ bestehende Kontinuum wählen. Um aber zu erreichen, daß E_3 einen Punkt von E_1 und E_2 enthält, rechnen wir zu E_1 noch den kürzesten Bogen zwischen C_k und C_∞ , und zu E_2 den zu demselben größten Kreis gehörigen Bogen zwischen C_∞ und dem Äquatorialkreis $\bar{C}$, nötigenfalls unter Hinzunahme eines Bogens von $\bar{C}$, der den auf diesem Kreis gelegenen Endpunkt mit C_i oder C_j verbindet. Für den kleinsten Abstand δ zwischen E_1 und E_2 kommen folgende Bogen in Betracht: 1. Der Durchmesser α_k von C_k ; 2. Die Abstände $\frac{\pi - d - \alpha_i}{2}$ und $\frac{\pi - d - \alpha_j}{2}$ zwischen C_∞ und C_i bzw. C_j ; 3. Der kleinste Abstand von C_i oder C_j zu dem Bogen zwischen C_k und C_∞ ; dieser kleinste Abstand wird an einem der Endpunkte erreicht und ist mithin größer als $\delta_{k,i}$ bzw. $\delta_{j,k}$ oder $\frac{\pi - d - \alpha_i}{2}$ bzw. $\frac{\pi - d - \alpha_j}{2}$. Es wird also jedenfalls $\delta \geq d$.

Im zweiten Falle hat man E_3 gleich C_∞ . E_2 ist das von C_k und dem kürzesten Bogen zwischen C_k und C_∞ gebildete Kontinuum, und für E_1 wählt man die Kreise C_i und C_j mit dem zwischenliegenden Bogen $\delta_{i,j}$, verbunden mit C_∞ durch einen Bogen des durch die Mittelpunkte von C_k und C_∞ gehenden größten Kreises und nötigenfalls einen Bogen des Äquatorialkreises $\bar{C}$. Dann sieht man ganz wie im ersten Falle ein, daß $\delta \geq d$ ist.

Wir sind also berechtigt, in die Formel (3) den Wert $K = \frac{8\pi}{d^2}$ einzutragen, und haben damit unseren Hauptsatz bewiesen.

Satz. Wenn die drei Kreise C_1 , C_2 und C_3 der Bedingung

$$(6) \quad \frac{(|w_0| + R)^2}{R} \leq \frac{d}{4} \cdot e^{-\frac{36\pi^2}{d^2}}$$

genügen, wobei die Bedeutung von w_0 , R und d dem vorhergehenden Texte zu entnehmen ist, so wird der Einheitskreis $|z| < 1$ durch jede in ihm reguläre Funktion $f(z) = z + \dots$ auf eine Riemannsche Fläche abgebildet, welche wenigstens einen der Kreise C_i schlicht überdeckt.

Es ist bemerkenswert, daß die rechte Seite von (6) nur die Größe d enthält und also ausschließlich von der *relativen* Größe und Lage der Kreise C_i abhängt. Demnach können wir immer durch eine auf drei beliebige Kreise ausgeübte Translation und Ähnlichkeitstransformation erreichen, daß die Bedingung (6) erfüllt wird.

6. Verschiedene Folgerungen

Will man aus unserem Satze die ursprüngliche *Blochsche* Formulierung herleiten, so setzt man $w_0 = 0$ und betrachtet einen symmetrischen Fall, wo die Kreise alle gleich groß sind und durch eine Drehung um 120° um den Nullpunkt auseinander hervorgehen. Um ein möglichst großes d zu erhalten wählt man $\alpha_i = \frac{\pi - \alpha_i}{3}$, d. h. $\alpha_i = \frac{\pi}{4}$, woraus $\delta_{i,j} = \frac{5\pi}{12}$ und folglich $d = \frac{\pi}{4}$ berechnet wird. Die Ungleichung (6) ist erfüllt für

$$R = \frac{\pi}{16} e^{-576},$$

und die Radien der entsprechenden Kreise C_i sind

$$B = \frac{\pi}{16} \operatorname{tg} \frac{\pi}{8} e^{-576}.$$

Hierin ist das folgende Resultat enthalten:

Es gibt eine Konstante B von der Eigenschaft, daß die von den Werten jeder im Einheitskreise regulären Funktion $f(z) = z + \dots$ erzeugte Riemannsche Fläche einen Kreis vom Radius B schlicht überdeckt, dessen Mittelpunkt im Abstände $\frac{B}{\sin \frac{\pi}{8}}$ vom Nullpunkt liegt.

Es kann noch hinzugefügt werden, daß die Strahle, die man vom Nullpunkt aus durch die Mittelpunkte der Kreise mit der genannten Eigenschaft zieht, wenigstens ein Drittel aller aus dem Nullpunkt ausgehenden Strahle umfassen.

Schließlich betrachten wir noch eine in der ganzen endlichen Ebene reguläre Funktion $f(z)$. Die Funktion $\frac{f(Az) - f(0)}{A f'(0)}$ ist für jedes noch

so große A regulär in E und genügt außerdem unseren Normierungsbedingungen. Uebt man auf drei beliebige, außerhalb einander gelegene Kreise C_1, C_2, C_3 die entsprechende Transformation $\frac{w - f(0)}{Af'(0)}$ aus, so sieht man, daß für ein genügend großes A die Bedingung (6) erfüllt wird. Also wird wenigstens einer der Kreise C_i von der Riemannschen Fläche der Umkehrfunktion von $f(z)$ schlicht überdeckt. Wir haben hiermit den folgenden Satz gefunden, der zur Bestimmung des Typus einer einfach zusammenhängenden Riemannschen Fläche angewandt werden kann.

Von drei außerhalb einander gelegenen Kreisen wird wenigstens einer von der Riemannschen Fläche der Umkehrfunktion einer ganzen Funktion schlicht überdeckt.

Geht man auf den Beweis zurück so erkennt man, daß die Kreise durch drei beliebige, einfach zusammenhängende Gebiete ersetzt werden können.

(Eingegangen den 10. März 1932)

Zusatz während der Korrektur

Nachdem diese Abhandlung zum Druck gelangt ist, habe ich gemerkt, daß die topologischen Betrachtungen auf S. 31—32 nicht ganz einwandfrei sind. Da ich in einer bald erscheinenden größeren Arbeit die Gelegenheit habe, den Fehler richtigzustellen, bitte ich die Leser, sich mit dieser Erklärung zu befriedigen.

Die Klassen der Abbildungen der n -dimensionalen Polyeder auf die n -dimensionale Sphäre

Von HEINZ HOPF, Zürich

1. Eine Behauptung von Brouwer und ihre Modifikation

Der Grad einer Abbildung f einer n -dimensionalen geschlossenen orientierten Mannigfaltigkeit μ auf eine ebensolche Mannigfaltigkeit μ' besitzt die wichtige Eigenschaft, bei stetiger Abänderung von f ungeändert zu bleiben¹⁾; mit andern Worten: zwei Abbildungen f und g von μ auf μ' , welche zu einer „Abbildungsklasse“ gehören, haben denselben Grad. Brouwer hat auf dem Internationalen Mathematikerkongreß in Cambridge 1912 die Behauptung ausgesprochen, daß „in vielen Fällen“ die Umkehrung dieses Satzes gelte, also aus der Gleichheit der Grade zweier Abbildungen ihre Zugehörigkeit zu einer Klasse folge²⁾. Er hat gleichzeitig einen Beweis seiner Behauptung für den Fall angegeben, in dem μ und μ' Kugelflächen sind; dann hat er ihre Gültigkeit für den allgemeineren Fall erwiesen, in dem zwar μ' eine Kugel, μ aber eine beliebige Fläche ist³⁾, und später habe ich gezeigt, daß dieser letzte Satz für beliebige Dimensionenzahl richtig ist, daß also $\mu = M^n$ eine n -dimensionale Mannigfaltigkeit, $\mu' = S^n$ die n -dimensionale Sphäre sein darf⁴⁾.

In der vorliegenden Arbeit soll nun gezeigt werden, daß der Gültigkeitsbereich der Brouwerschen Behauptung noch weiter ist, falls man sich nicht genau an ihren Wortlaut hält, sondern sie einer Modifikation unterzieht, die mir überdies, worüber nachher (Nr. 2) noch einige Worte gesagt werden sollen, die prinzipielle Bedeutung der Behauptung und der an sie anschließenden Sätze in ein klareres Licht zu setzen scheint.

In der neuen Erweiterung soll wieder $\mu' = S^n$ die n -dimensionale Sphäre, $\mu = P^n$ aber soll ein beliebiges n -dimensionales Polyeder sein. Dann hat eine Abbildung f von P^n auf S^n keinen Grad im ursprünglichen

¹⁾ Brouwer, Ueber Abbildung von Mannigfaltigkeiten, Math. Annalen 71 (1912).

²⁾ Brouwer, Sur la notion de «classe» de transformations d'une multiplicité, Proc. V. Intern. Congress of Math. (Cambridge 1912), vol. II.

³⁾ Brouwer, Over één-éénduidige continue transformaties..., Akad. Amsterdam, Versl. 21 (1913); Aufzählung der Abbildungsklassen endlichfach zusammenhängender Flächen, Math. Annalen 82 (1921).

⁴⁾ Abbildungsklassen n -dimensionaler Mannigfaltigkeiten, Math. Annalen 96 (1926).

Sinn. Die Modifikation, die man hier vorzunehmen hat, ist durch die Begriffsbildungen der algebraisch-kombinatorischen Topologie in natürlicher Weise gegeben. Ist Z^n ein n -dimensionaler Zyklus⁵⁾ in P^n , so ist sein Bild $f(Z^n)$ ein n -dimensionaler Zyklus in S^n ; aber die einzigen n -dimensionalen Zyklen in S^n sind die S^n selbst und ihre Vielfachen; daher gibt es eine ganze Zahl c so, daß $f(Z^n) = c \cdot S^n$ ist. Falls P^n eine Mannigfaltigkeit ist, ist c der Brouwersche Grad; wir nennen c auch jetzt den Grad von $f(Z^n)$. Die zu den verschiedenen Zyklen Z^n in P^n gehörigen Grade sind innerhalb der durch f bestimmten Abbildungsklasse konstant. Ihre Betrachtung reicht jedoch für unsern Zweck, die Aufstellung eines vollen Invariantensystems der Abbildungsklasse, nicht aus; das sieht man schon im Falle $n = 2$, wenn man für P^2 die projektive Ebene nimmt: dann ist in P^2 überhaupt kein Z^2 vorhanden, und es gibt trotzdem zwei Abbildungsklassen; diese kann man aber durch ihre „Parität“ oder den „Abbildungsgrad mod. 2“ voneinander unterscheiden, und daran erkennt man, wie man im allgemeinen Fall fortzufahren hat: es sei Z_m^n ein Zyklus mod. m mit irgend einem ganzen $m > 1$ ⁶⁾; dann ist sein Bild $f(Z_m^n)$ ein Zyklus mod. m in S^n , und daraus folgt, ähnlich wie oben, daß es eine, mod. m eindeutig bestimmte Zahl c so gibt, daß $f(Z_m^n) \equiv c S^n$ ist. Diese Zahl c , der „Grad mod. m “ von $f(Z_m^n)$, bleibt ebenfalls in der Klasse konstant. Alle diese Grade und Grade mod. m mit beliebigen $m > 1$, die zu den, in P^n in endlicher Anzahl vorhandenen, n -dimensionalen Zyklen und Zyklen mod. m gehören, bilden nun aber — das ist die Erweiterung der Brouwerschen Behauptung, die hier bewiesen werden soll, — ein volles Invariantensystem der Abbildungsklasse; es gilt also

Satz I. *Notwendig und hinreichend dafür, daß zwei Abbildungen f und g von P^n auf S^n zu einer Klasse gehören, ist die Bedingung, daß jeder n -dimensionale Zyklus bzw. Zyklus mod. m (mit beliebigem $m > 1$) aus P^n durch f mit demselben Grade bzw. Grade mod. m abgebildet wird wie durch g .*

Der Beweis dieses Satzes wird in Nr. 3—5 geliefert werden⁶⁾.

⁵⁾ Ein Zyklus ist ein unberandeter Komplex, ein n -dimensionaler Zyklus mod. m ein Komplex, in dessen Rande jedes $(n - 1)$ -dimensionale Simplex mit einer durch m teilbaren Vielfachheit vorkommt. Die Grundtatsachen aus der kombinatorischen Topologie und aus der Topologie der stetigen Abbildungen werden als bekannt vorausgesetzt.

⁶⁾ Den Spezialfall, in dem g eine Abbildung auf einen einzigen Punkt von S^n ist, habe ich bereits früher bewiesen: Ueber wesentliche und unwesentliche Abbildungen von Komplexen, Moskauer Mathematische Sammlung, 1930 (Satz II). Die dortige Methode reicht auch zum Beweis des obigen Satzes I aus, jedoch scheint mir für diesen Zweck die in der vorliegenden Arbeit verwendete Zurückführung auf einen „Erweiterungssatz“ (Nr. 3) den Vorzug zu verdienen.

2. Verallgemeinerung der Fragestellung; Klassen und algebraische Typen von Abbildungen

Die Verwendung von Begriffen der algebraisch-kombinatorischen Topologie, wie sie für die Formulierung des Satzes I notwendig war, führt, wenn man sie konsequent weiter treibt, zu der allgemeinen Problemstellung, in deren Rahmen erst die tiefere Bedeutung der Brouwerschen Behauptung sichtbar wird. Wenn man nämlich zwei beliebige Polyeder P, Q und die Gesamtheit der stetigen Abbildungen von P auf Q betrachtet, so gibt es zwei, ihrem Wesen nach voneinander verschiedene, Gesichtspunkte, unter denen man versuchen kann, in diese Gesamtheit Ordnung zu bringen, die Abbildungen also zu klassifizieren: erstens eben den rein topologischen Begriff der „Abbildungsklasse“, wonach zwei Abbildungen f und g zusammengehören, wenn man die eine stetig in die andere überführen kann; zweitens den, auf den Grundbegriffen der algebraischen Topologie, den Begriffen der Berandung und der Homologie, beruhenden Begriff des „algebraischen Abbildungstypus“, den wir folgendermaßen definieren: f und g gehören zu einem algebraischen Typus, wenn von jedem Zyklus $Z \subset P$ die beiden Bilder $f(Z)$ und $g(Z)$, die ja als Zyklen in Q aufzufassen sind, einander homolog sind, und wenn das Gleiche für die Zyklen mod. m gilt, wobei man natürlich den Begriff der gewöhnlichen Homologie durch den der „Homologie mod. m “ zu ersetzen hat. Man kann noch ein drittes Klassifikationsprinzip hinzufügen, indem man anstelle der Homologiegruppen die Fundamentalgruppe betrachtet, doch soll darauf hier nicht eingegangen werden⁷⁾. Die Dimensionen von P und Q sind für diese Begriffe ganz unwesentlich, sie brauchen nicht einander gleich zu sein. Ist Q n -dimensional, so fällt für die n -dimensionalen Zyklen und Zyklen mod. m in Q der Begriff der Homologie bzw. Homologie mod. m mit dem der Gleichheit bzw. Kongruenz mod. m zusammen; in diesem Fall wird daher der algebraische Typus einer Abbildung, soweit er die n -dimensionalen Zyklen in P und Q betrifft, vollständig durch die Angabe der Grade und Grade mod. m beschrieben; ist speziell $Q = S^n$, so ist für $0 < r < n$ jeder r -dimensionale Zyklus oder Zyklus mod. m in S^n homolog 0, so daß diese Zyklen kein Unterscheidungsmerkmal für die Abbildungstypen liefern; mithin sind dann die Grade und Grade mod. m die einzigen Merkmale der Typen. Daher kann man den Satz I auch so aussprechen:

Satz I'. *Ist P ein n -dimensionales Polyeder, Q die n -dimensionale*

⁷⁾ Man vergl. etwa den § 2 meiner Arbeit: Zur Topologie der Abbildungen von Mannigfaltigkeiten, II. Teil, Math. Annalen 102 (1929).

Sphäre, so gehören zwei Abbildungen von P auf Q dann und nur dann zu einer Klasse, wenn sie denselben algebraischen Typus haben.

Der eine Teil dieses Satzes ist insofern trivial, als bei beliebigen P und Q zwei Abbildungen, die zu einer Klasse gehören, stets denselben algebraischen Typus besitzen, da ein Zyklus $f(Z)$ in Q , wenn man ihn stetig abändert, immer in derselben Homologiekategorie bleibt. Die Einteilung aller Abbildungen in Klassen ist also im allgemeinen, jedenfalls *begrifflich*, *feiner* als die nach algebraischen Typen; dafür, daß sie auch *tatsächlich* feiner sein kann, gibt es Beispiele, von denen nachher noch die Rede sein soll; im allgemeinen reichen somit die *Homologie*-Begriffe nicht aus, um die Klassifikation der Abbildungen nach dem rein topologischen Standpunkt der „*Homotopie*“, d. h. der stetigen Überführbarkeit, durchzuführen. Das ist auch gar nicht zu erwarten, denn der Begriff der Homologie hat kaum etwas mit stetiger Abänderung zu tun; andererseits spielt der Homologiebegriff — und zwar gerade infolge der Entwicklung während der letzten Jahre — eine so beherrschende Rolle in fast allen Gebieten der Topologie, daß die Frage nach den „Ausnahmefällen“ gerechtfertigt ist, in denen er doch dasselbe leistet wie die Homotopie; das sind, für unser Problem, die Fälle, in denen für zwei Abbildungen aus der Gleichheit des algebraischen Typus folgt, daß sie sich stetig ineinander überführen, daß sie sich also auch unter dem Gesichtspunkt der Homotopie nicht voneinander unterscheiden lassen. Wenn man nun die eingangs zitierte Behauptung Brouwers weiter — allerdings recht kräftig — modifiziert, so kann man sie so aussprechen: es gibt eine große Menge von Ausnahmen der eben genannten Art; und der Satz I gibt eine wichtige Klasse aus dieser Menge an. Behauptung und Satz gehören also in den allgemeinen Problemkreis, in dem es sich um die Zusammenhänge zwischen Homologie und Homotopie, genauer: um den *Einfluß von Berandungs- und Homologieeigenschaften auf Homotopieeigenschaften*, handelt⁸⁾.

Es sei nun noch etwas über die „allgemeinen“ Fälle gesagt, in denen P und Q so beschaffen sind, daß die Einteilung in Klassen wirklich feiner ist als die Einteilung nach algebraischen Typen. Bleiben wir zunächst dabei, daß $Q = S^n$ ist; (das ist für alle Anwendungen der wichtigste Fall;) ist dann P ein r -dimensionales Polyeder und $r < n$, so bleibt der Satz I trivialerweise noch richtig, denn dann gibt es nur eine Klasse, — da man das Bild $f(P)$ stetig auf einen Punkt zusammenziehen kann,

⁸⁾ Daß der in Satz I, in seiner in Nr. 1 gegebenen Formulierung, benutzte Begriff des Grades zu den Berandungseigenschaften gehört, ist klar: er benutzt ja den auf dem Begriff des Randes beruhenden Begriff des Zyklus (man vergl.⁵⁾).

— und a fortiori nur einen Typus; wir befinden uns also noch bei einem „Ausnahmefall“; ist dagegen $r > n$, so zeigt das Beispiel $P = S^3$, $Q = S^2$, daß der Satz I nicht für alle P gilt: es gibt dann offenbar nur einen einzigen algebraischen Typus, da jeder 1- oder 2-dimensionale Zyklus ~ 0 in S^3 , sein Bild daher ~ 0 in S^2 ist, dagegen, wie ich gezeigt habe, unendlich viele Klassen⁹⁾. Ist Q keine Sphäre, so ist es leichter, Beispiele zu finden, in denen ein Typus mehrere Klassen enthält; solche erhält man bereits, wenn P und Q geschlossene orientierbare Flächen von Geschlechtern > 0 sind; jedoch reicht in diesem Fall zur Bestimmung der Abbildungsklassen die oben kurz erwähnte Betrachtung der Fundamentalgruppe aus¹⁰⁾. Aber auch diese versagt z. B. in folgendem Fall: P sei eine Kugelfläche, Q eine projektive Ebene, f die Abbildung von P auf Q , die sich ergibt, wenn man P als zweiblättrige unverzweigte Überlagerungsfläche von Q auffaßt, g die Abbildung, die P auf einen einzigen Punkt von Q abbildet; dann sieht man leicht, daß f und g zwar denselben algebraischen Typus besitzen, aber zu verschiedenen Klassen gehören¹¹⁾.

Demnach scheint sich der Satz I nicht auf eine wesentlich größere Gesamtheit von Paaren P, Q ausdehnen zu lassen, es sei denn, daß man neben Polyedern auch andere abgeschlossene Mengen in Betracht zieht¹²⁾.

Abgesehen von diesen prinzipiellen Gesichtspunkten hat der Satz I auch praktischen Wert insofern, als man mit seiner Hilfe alle Abbildungsklassen von P^n auf S^n wirklich aufzählen kann, wenn man die kombinatorisch-topologische Struktur von P^n kennt; denn der Satz besagt ja, daß man nur die algebraischen Typen aufzuzählen hat, und das ist eine leichte, im wesentlichen algebraische, Aufgabe, die in Nr. 6 gelöst wird.

3. Zurückführung des Hauptsatzes (Satz I) auf einen „Erweiterungssatz“ (Satz II).

Daß die im Satz I genannte Bedingung für die Zugehörigkeit von f und g zu einer Klasse notwendig ist, ist, wie schon mehrfach erwähnt,

⁹⁾ Über die Abbildungen der dreidimensionalen Sphäre auf die Kugelfläche, Math. Annalen 104 (1931).

¹⁰⁾ Brouwer, wie unter³⁾ (Aufzählung..., „Vierter Hauptfall“); Hopf, Beiträge zur Klassifizierung der Flächenabbildungen, Crelles Journal 165 (1931).

¹¹⁾ In der Terminologie meiner unter⁷⁾ zitierten Arbeit hat der „Absolutgrad“ von f den Wert 2, von g den Wert 0; da er (a. a. O. §2) in der Klasse konstant ist, gehören f und g zu verschiedenen Klassen.

¹²⁾ Die Antwort auf die Frage, ob die Abbildungen einer abgeschlossenen Menge F auf die S^n mehr als eine Klasse bilden, ist für wichtige geometrische Eigenschaften von F ausschlaggebend: Alexandroff, Dimensionstheorie, Math. Annalen 106 (1932), Nr. 81 („5. Hauptsatz“); Borsuk, Über Schnitte der n -dimensionalen Euklidischen Räume, Math. Annalen 106 (1932).

bekannt, da der Grad einer Abbildung $f(Z^n)$ und ebenso ein Grad mod. m sich bei stetiger Abänderung von f nicht ändert. Zu beweisen ist, daß die Bedingung hinreicht, daß es also, wenn sie erfüllt ist, eine Schar f_t von Abbildungen von P^n auf S^n gibt, die für $0 \leq t \leq 1$ stetig von t abhängt und in der $f_0 = f$, $f_1 = g$ ist. Zum Zweck des Beweises deuten wir eine solche Schar folgendermaßen. P^{n+1} sei das „Produkt“ von P^n mit einer Strecke der Länge 1; dieses Produkt können wir so konstruieren: wir denken uns den euklidischen Raum R^N , in dem P^n liegt, im R^{N+1} gelegen und errichten nach einer bestimmten der beiden Seiten von R^N die Senkrechten auf R^N in allen Punkten p von P^n ; p_t sei der Punkt, der auf der in p errichteten Senkrechten im Abstand t von p liegt; die Menge aller p_t mit $0 \leq t \leq 1$ ist das Produkt. Es ist ein $(n+1)$ -dimensionales Polyeder P^{n+1} .

Die Punkte $p = p_0$ bilden das Polyeder $P^n = P_0^n$, die Punkte p_1 ein mit P^n kongruentes Polyeder P_1^n ; unter $\bar{P}$ verstehen wir das Polyeder $P_0^n + P_1^n$. Üben wir die Abbildung f auf P_0^n , die Abbildung g mittels der Festsetzung $g(p_1) = g(p)$ auf P_1^n aus, so liegt eine Abbildung F von $\bar{P}$ auf S^n vor. Wenn wir F zu einer Abbildung des ganzen Polyeders P^{n+1} auf S^n erweitern können, so sind wir fertig; denn dann brauchen wir nur $f_t(p) = F(p_t)$ zu setzen, um eine Schar der gewünschten Art zu erhalten. Die *Behauptung* lautet also: *die Abbildung $F(\bar{P})$ läßt sich zu einer Abbildung $F(P^{n+1})$ erweitern.*

Wie lautet jetzt, unter Verwendung von P^{n+1} , $\bar{P}$ und F die *Voraussetzung* des Satzes I? Ich behaupte, daß sie folgendermaßen lautet: *jeder in $\bar{P}$ gelegene n -dimensionale Zyklus oder Zyklus mod. m , der ~ 0 bezw. $\sim 0 \bmod m$ in P^{n+1} ist, wird durch F mit dem Grade 0 abgebildet.*

In der Tat: ist $Z^n \subset \bar{P}$, so zerfällt Z^n in zwei zu einander fremde Teile $X_0^n \subset P_0^n$, $Y_1^n \subset P_1^n$; da Z^n unberandet ist, haben sie selbst keine Ränder, sind also Zyklen. Ist Y_0^n der Y_1^n entsprechende Zyklus in P_0^n , so ist $Y_0^n \sim Y_1^n$ in P^{n+1} , da offenbar $Y_1^n - Y_0^n$ der Rand des $(n+1)$ -dimensionalen Produktes von Y_0^n mit der t -Strecke ist. Folglich ist $Z^n = X_0^n + Y_1^n \sim X_0^n + Y_0^n$ in P^{n+1} , und da wir voraussetzen, daß $Z^n \sim 0$ ist, ist daher auch der in P_0^n gelegene Zyklus $X_0^n + Y_0^n \sim 0$ in P^{n+1} . K sei ein von $X_0^n + Y_0^n$ berandeter Komplex, K_0 seine Projektion auf P_0^n , (die man erhält, indem man für jeden Punkt $p_t \in K$ t durch 0 ersetzt); da bei dieser Projektion (wie bei jeder simplizialen Abbildung) der Rand von K in den Rand des Bildes K_0 übergeht, der Rand $X_0^n + Y_0^n$ von K aber fest bleibt, ist $X_0^n + Y_0^n$ der Rand von K_0 ; also ist $X_0^n + Y_0^n \sim 0$ in P_0^n , und da P_0^n ebenso wie X_0^n und Y_0^n n -dimensional ist, bedeutet das: $X_0^n + Y_0^n = 0$, also $X_0^n = -Y_0^n$. Da mithin

$Z^n = Y_1^n - Y_0^n$ ist, gilt bei der Abbildung: $F(Z^n) = F(Y_1^n) - F(Y_0^n) = g(Y_0^n) - f(Y_0^n)$, und da nach Voraussetzung $f(Y_0^n)$ und $g(Y_0^n)$ den gleichen Grad, etwa c , haben: $F(Z^n) = cS^n - cS^n = 0$; das bedeutet, daß $F(Z^n)$ den Grad 0 hat. Diese Betrachtung gilt in gleicher Weise für gewöhnliche Zyklen und Homologien wie mod. m . Damit ist bewiesen, daß die Voraussetzung des Satzes I jetzt in der angegebenen Form ausgesprochen werden kann.

Somit ist der Satz I auf den folgenden allgemeineren „Erweiterungssatz“ zurückgeführt¹³⁾, in dem P^{n+1} irgend ein $(n+1)$ -dimensionales Polyeder ist:

Satz II. *In einem Teilpolyeder¹⁴⁾ $\bar{P}$ des $(n+1)$ -dimensionalen Polyeders P^{n+1} sei eine Abbildung F auf die S^n gegeben; für jeden n -dimensionalen Zyklus (und Zyklus mod. m) $Z^n \subset \bar{P}$, welcher ~ 0 (bezw. ~ 0 mod. m) in P^{n+1} ist, sei der Grad (bezw. Grad mod. m) gleich (bezw. kongruent) 0. Dann läßt sich F zu einer Abbildung des ganzen P^{n+1} auf die S^n erweitern¹⁵⁾.*

Daß die in der Voraussetzung des Satzes ausgedrückte Bedingung für die Erweiterbarkeit von F zu einer Abbildung von P^{n+1} nicht nur hinreichend, sondern auch notwendig ist, ist klar: wenn $F(P^{n+1})$ existiert und wenn $Z^n \sim 0$ in P^{n+1} , also der Rand eines $K \subset P^{n+1}$ ist, so ist $F(Z^n)$ der Rand von $F(K)$, also ~ 0 auf S^n , also $= 0$; das Analoge gilt mod. m .

Der einfachste Spezialfall des Satzes II ist

Satz II'. *Ist auf dem Rande eines $(n+1)$ -dimensionalen Simplexes eine Abbildung F vom Grade 0 auf die S^n gegeben, so läßt sich F zu einer Abbildung des ganzen Simplexes auf die S^n erweitern.*

Dieser Satz, den ich früher bewiesen habe¹⁶⁾, bildet den wesentlichen topologischen Bestandteil beim Beweise des Satzes II; es müssen aber, wie schon die im Satz II vorkommenden Begriffe der Zyklen und Zyklen mod. m vermuten lassen, noch algebraische Bestandteile hinzukommen; auch diese werden sich auf die Erweiterungen gewisser Abbildungen, nämlich homomorpher Gruppenabbildungen, beziehen.

¹³⁾ Der Zusammenhang zwischen Sätzen über Abänderungen von Abbildungen mit Sätzen über Erweiterungen spielt in der unter ¹²⁾ zitierten Arbeit von Borsuk eine wesentliche Rolle; man vergl. auch die §§ 5, 6 meiner unter 4) genannten Arbeit.

¹⁴⁾ Ein „Teilpolyeder“ $\bar{P}$ eines Polyeders P soll stets aus Simplexen einer gegebenen Simplexzerlegung von P bestehen; die Dimension von $\bar{P}$ ist beliebig.

¹⁵⁾ Man überzeugt sich leicht davon, daß man die auf die gewöhnlichen Zyklen bezügliche Voraussetzung sparen kann, da sie in der auf die Zyklen mod. m bezüglichen enthalten ist.

¹⁶⁾ Wie unter 4); ein Beweis von II' ist dort im letzten Abschnitt der S. 224 enthalten.

4. Algebraische Hilfssätze

Die hier vorkommenden Gruppen sind Abelsch, werden von endlich vielen ihrer Elemente erzeugt und enthalten keine Elemente endlicher Ordnung; die Gruppenoperation bezeichnen wir als Addition. Die Gruppe G heißt, wie üblich, direkte Summe ihrer Untergruppen U, V — geschrieben: $G = U + V$ — wenn sich jedes von 0 verschiedene Element auf eine und nur eine Weise in der Form $u + v$ mit $u \in U, v \in V$ darstellen läßt, oder, was dasselbe ist, wenn 1) jedes Element wenigstens eine Darstellung $u + v$ besitzt, und wenn 2) U und V nur das Nullelement gemeinsam haben. Analog ist die direkte Summe von mehr als zwei Gruppen definiert. Jede der hier betrachteten Gruppen ist bekanntlich direkte Summe von endlich vielen unendlichen zyklischen Gruppen; d. h. jedes Element läßt sich auf eine und nur eine Weise in der Form $\sum a_i x_i$ darstellen, wenn die x_i erzeugende Elemente dieser Zyklen, die a_i ganze Zahlen sind. — Die Untergruppe U von G heiße „abgeschlossen“, wenn sie folgende Eigenschaft hat: ist m eine von 0 verschiedene ganze Zahl, x ein Element von G und $mx \in U$, so ist auch $x \in U$. — Unter einem „Charakter“ von G verstehen wir eine homomorphe Abbildung von G in die additive Gruppe der ganzen Zahlen.

a) Ist U abgeschlossene Untergruppe von G , so ist G direkte Summe von U und einer anderen Untergruppe V .

Beweis: Die Restklassengruppe (Faktorgruppe) R von G nach U ist Abelsch; sie wird von endlich vielen ihrer Elemente erzeugt; (als solche kann man die Restklassen wählen, die die Elemente eines Erzeugendensystems von G enthalten); sie enthält ferner infolge der Abgeschlossenheit von U kein Element endlicher Ordnung. Sie ist daher direkte Summe unendlicher Zyklen; X_i seien Restklassen, die diese Zyklen erzeugen, x_i irgendwelche Elemente aus den X_i ($i = 1, \dots, r$), V sei die von diesen x_i erzeugte Gruppe. Ist y irgend ein Element von G , Y die y enthaltende Restklasse, so ist Y in R von der Form $Y = \sum a_i X_i$, also ist $y = \sum a_i x_i + u$ mit $u \in U$, also $y = u + v$ mit $u \in U, v \in V$. Ist $u_0 \in U$ und $u_0 \in V$, so ist $u_0 = \sum c_i x_i$, also in $R: 0 = \sum c_i X_i$; folglich ist $c_i = 0, u_0 = 0$. Mithin ist $G = U + V$.

b) Ein in einer abgeschlossenen Untergruppe U von G gegebener Charakter läßt sich stets zu einem Charakter von G erweitern.

Beweis: Man stelle G gemäß a) in der Form $U + V$ dar und setze fest, daß der Charakter für alle Elemente von V den Wert 0 hat.

c) Dafür, daß ein in einer beliebigen Untergruppe U von G gege-

bener Charakter χ zu einem Charakter von G erweitert werden kann, ist die folgende Bedingung notwendig und hinreichend: ist x ein Element von G , m eine ganze Zahl, und ist $mx = u \in U$, so ist $\chi(u)$ durch m teilbar.

Beweis: Die Bedingung ist notwendig, da, wenn χ auf G erweitert ist, $\chi(x)$ definiert ist und $\chi(u) = m \cdot \chi(x)$ wird. — Die Bedingung sei erfüllt. Unter $\bar{U}$ verstehen wir die Gesamtheit der Elemente x , welche Vielfache mx in U besitzen; sie bilden eine Gruppe, da mit x auch $-x$ in $\bar{U}$ ist, und da aus $mx \in U$, $ny \in U$ folgt: $mn(x+y) \in U$. Die Gruppe $\bar{U}$ ist ex definitione abgeschlossen. Daher läßt sich nach b) der Charakter, falls er sich auf $\bar{U}$ erweitern läßt, auch auf G erweitern. Wir haben also χ auf $\bar{U}$ auszudehnen. Ist $mx = u \in U$, so setzen wir $\chi(x) = \frac{1}{m} \chi(u)$; das ist nach Voraussetzung eine ganze Zahl. $\chi(x)$ ist auf diese Weise eindeutig bestimmt; denn ist außerdem $m'x = u' \in U$, so ist $mu' = m'u$, also $m \cdot \chi(u') = m' \cdot \chi(u)$, also $\frac{1}{m'} \chi(u') = \frac{1}{m} \chi(u)$. Diese somit in $\bar{U}$ eindeutige Funktion ist ein Charakter; denn ist $mx = u_1$, $ny = u_2$, so ist $mn(x+y) = nu_1 + mu_2 \in U$, also $\chi(x+y) = \frac{1}{m} \chi(u_1) + \frac{1}{n} \chi(u_2) = \chi(x) + \chi(y)$.

d) U und V seien Untergruppen von G ; in U sei ein Charakter χ gegeben. Dafür, daß sich χ derart auf G erweitern läßt, daß er in allen Elementen von V den Wert 0 erhält, ist die folgende Bedingung notwendig und hinreichend: ist x ein Element von G , m eine ganze Zahl, v ein Element von V , und ist $mx + v = u \in U$, so ist $\chi(u)$ durch m teilbar.

Beweis: Die Notwendigkeit der Bedingung ist wieder ohne weiteres klar: wenn ein Charakter χ mit den genannten Eigenschaften in G existiert, so ist $\chi(mx + v) = m \cdot \chi(x) + \chi(v) = m \cdot \chi(x)$. — Die Bedingung sei erfüllt. Ist $z = u = v$ ein Element aus dem Durchschnitt D von U und V , so ist, wenn x_0 das Nullelement von G bezeichnet, $u = mx_0 + v$ mit beliebigem m , also nach Voraussetzung $\chi(u)$ durch jedes m teilbar, also $\chi(u) = \chi(z) = 0$ für jedes $z \in D$. W sei die von U und V erzeugte Gruppe, also die Gesamtheit aller Elemente $u + v$. Wir erweitern χ zunächst auf W , indem wir festsetzen: $\chi(u + v) = \chi(u)$; diese Festsetzung ist eindeutig; denn ist $u + v = u' + v'$, so ist $u - u' = v' - v = z \in D$, also $\chi(z) = 0$, d. h. $\chi(u) = \chi(u')$. Daß diese somit in W eindeutig erklärte Funktion ein Charakter ist und in allen Elementen

von V den Wert 0 hat, ist klar. Ist nun $x \in G$, m eine ganze Zahl und $mx \in W$, so ist $mx = w = u + v$, $u = mx - v$, also nach Voraussetzung $\chi(u)$ durch m teilbar, also, da $\chi(w) = \chi(u)$ ist, $\chi(w)$ durch m teilbar. Daher läßt sich nach c) χ auf die ganze Gruppe G erweitern.

5. Beweis des Satzes II

Wir machen zunächst die spezielle Annahme, daß die in $\bar{P}$ gegebene Abbildung F simplizial sei. Dabei ist eine feste Simplexzerlegung von P^{n+1} , und damit auch von $\bar{P}$, zugrunde gelegt. Für diese Zerlegung sei G die Gruppe der n -dimensionalen Komplexe in P^{n+1} , d. h. der Linearformen mit ganzen Koeffizienten in den orientierten n -dimensionalen Simplexen, die wir mit x_i^n bezeichnen; U sei die Gruppe der zu $\bar{P}$ gehörigen n -dimensionalen Komplexe, V die Gruppe der n -dimensionalen Ränder in P^{n+1} , d. h. derjenigen Zyklen, welche $(n+1)$ -dimensionale Komplexe beranden; U und V sind Untergruppen von G . τ^n sei ein festes n -dimensionales Simplex der bei der simplizialen Abbildung F benutzten Zerlegung von S^n . Jedes x_i^n aus $\bar{P}$, das durch F auf τ^n abgebildet wird, hat dabei den Grad $+1$ oder -1 ; wir nennen ihn $\chi(x_i^n)$; für diejenigen x_i^n aus $\bar{P}$, die nicht auf dieses τ^n abgebildet werden, setzen wir $\chi(x_i^n) = 0$. Für einen beliebigen Komplex $x^n = \sum a_i x_i^n$ aus $\bar{P}$ hat dann F in dem Simplex τ^n den Grad $\chi(x^n) = \sum a_i \chi(x_i^n)$. χ ist ein Charakter in der Gruppe U . Ich behaupte, daß er in bezug auf die Gruppen G , U und V die Voraussetzungen des Hilfssatzes d) aus Nr. 4 erfüllt. In der Tat: ist, in der Bezeichnung von d), $mx + v = u$, so bedeutet das jetzt: der in $\bar{P}$ gelegene, n -dimensionale Komplex u ist mod. m einem Rande v in P^{n+1} kongruent, er ist also ein Zyklus mod. m , der ~ 0 mod. m in P^{n+1} ist; dann ist nach der Voraussetzung des Satzes II der Grad mod. m seines Bildes $F(u)$ Null; das gilt insbesondere in dem Simplex τ^n von S^n , und das bedeutet in unserer neuen Ausdrucksweise: $\chi(u) \equiv 0$ mod. m . Da somit die Voraussetzung von d) erfüllt ist, gilt auch die Behauptung, und wir können daher χ auf die Gruppe G aller n -dimensionalen Komplexe von P^{n+1} so erweitern, daß dieser Charakter für jeden Rand v den Wert 0 hat.

Nachdem damit die algebraischen Vorbereitungen erledigt sind, wird die gewünschte Ausdehnung von $F(\bar{P})$ auf das ganze Polyeder P^{n+1} in zwei Schritten vorgenommen werden: 1) Q sei das Polyeder, das aus allen nicht zu $\bar{P}$ gehörigen n -dimensionalen Simplexen von P^{n+1} besteht; dann wird F derart auf $\bar{P} + Q$ ausgedehnt, daß für jedes Simplex x_i^n

von $P^{n+1} \chi(x_i^n)$ der Grad des Bildes $F(x_i^n)$ in dem Simplex T^n ist; (dabei wird $F(Q)$ im allgemeinen nicht mehr simplizial sein); 2) die Abbildung $F(\bar{P} + Q)$ wird zu einer Abbildung $F(P^{n+1})$ erweitert.

Wir zeigen zunächst, wie man den zweiten Schritt vornimmt, wenn der erste bereits ausgeführt ist: da $\chi(x_i^n)$ der Grad in τ^n für jedes Simplex x_i^n ist, ist $\sum a_i \chi(x_i^n) = \chi(x)$ der Grad in T^n bei der Abbildung des Komplexes $x = \sum a_i x_i^n$; das gilt insbesondere, wenn $x = v$ ein Rand ist; für einen solchen ist der Grad daher $\chi(v) = 0$, und zwar ist dies, da v ein Zyklus ist, nicht nur der Grad in τ^n , sondern der Grad der Abbildung $F(v)$ schlechthin. Ist nun y^{n+1} ein (nicht zu $\bar{P}$ gehöriges) $(n+1)$ -dimensionales Simplex von P^{n+1} , so läßt sich, da sein Rand v mit dem Grade 0 abgebildet wird, diese Abbildung F auf Grund des Satzes II' auf y^{n+1} ausdehnen. Tun wir dies für jedes y^{n+1} , so erhalten wir die gewünschte Abbildung von P^{n+1} .

Die Ausführung des ersten Schrittes, die nun noch nachzuholen ist, ist ganz elementar und unabhängig von dem Satz II' und den algebraischen Betrachtungen. Im Inneren jedes n -dimensionalen Simplexes x_i^n von Q wählen wir ein System von zueinander fremden n -dimensionalen Teilsimplexen in der Anzahl $|\chi(x_i^n)|$; jedes von ihnen bilden wir affin auf τ^n ab und zwar mit dem Grade $+1$ oder -1 , je nachdem $\chi(x_i^n)$ positiv oder negativ ist. Wenn wir nun die Abbildung F , die jetzt außer in $\bar{P}$ auch in diesen Teilsimplexen erklärt ist, so auf den Rest von Q ausdehnen, daß die noch hinzukommenden Bildpunkte nicht im Innern von τ^n liegen, so sind wir fertig; denn dann hat jedes x_i^n in τ^n den Grad $\chi(x_i^n)$. Q' sei der Teil von Q , der entsteht, wenn man die Innengebiete aller der eben betrachteten n -dimensionalen Teilsimplexe aus Q entfernt. Die Ränder dieser Teilsimplexe und der Durchschnitt $Q \cdot \bar{P}$ bilden die Teilmenge $\bar{Q}$ von Q' , auf der F schon erklärt ist; sie ist ein $(n-1)$ -dimensionales Polyeder, und F ist auf ihm simplizial; daher liegt keiner der zugehörigen Bildpunkte im Inneren von τ^n . a sei ein innerer Punkt von τ^n ; wir fassen jetzt für einen Augenblick S^n als euklidischen Raum R^n mit a als unendlich fernem Punkt auf. Dann liegt die Bildmenge $F(\bar{Q})$ in R^n ; die euklidischen Koordinaten der Bildpunkte sind stetige Funktionen auf $\bar{Q}$; nach dem allgemeinen Erweiterungssatz für stetige Funktionen¹⁷⁾ können wir diese Funktionen auf ganz Q' ausdehnen; dadurch wird $F(\bar{Q})$ zu einer Abbildung $F_1(Q')$ in den R^n er-

¹⁷⁾ Hausdorff, Mengenlehre (2. Aufl. 1927), S. 248; von Kerékjártó, Vorlesungen über Topologie (1923), S. 75.

weitert; kehren wir zu der früheren Auffassung der S^n zurück, so kann bei dieser Abbildung die Bildmenge zwar ins Innere von τ^n eintreten; jedoch bleibt der Punkt a unbedeckt. Wenn wir daher jeden im Inneren von τ^n liegenden Bildpunkt durch den Punkt des Randes von τ^n ersetzen, in den er von a aus projiziert wird, so erhalten wir eine stetige Abbildung $F(Q')$, die alle Anforderungen erfüllt.

Wir haben uns jetzt noch von der am Anfang des Beweises gemachten Annahme zu befreien, daß F auf $\bar{P}$ simplizial sei. F sei also eine beliebige stetige Abbildung von $\bar{P}$, die die Voraussetzungen des Satzes II erfüllt; dann sei F' eine so gute simpliziale Approximation von F , daß sie auch noch diese Voraussetzungen erfüllt und daß für jeden Punkt $\bar{p} \subset \bar{P}$ die Entfernung $\rho(F'(\bar{p}), F(\bar{p})) < 1$ ist; dabei fassen wir S^n als Kugel vom Radius 1 im euklidischen R^{n+1} auf. $F'(p)$ dürfen wir für alle $p \subset P^{n+1}$ als definiert betrachten. Unter $v(\bar{p})$ verstehen wir den Vektor mit dem Anfangspunkt $F'(\bar{p})$ und dem Endpunkt $F(\bar{p})$. Die Komponenten dieser Vektoren sind stetige Funktionen auf $\bar{P}$; wir können sie nach dem allgemeinen Erweiterungssatz¹⁷⁾ zu stetigen Funktionen auf P^{n+1} erweitern; damit ist jedem Punkt $p \subset P^{n+1}$ ein Vektor zugeordnet; dabei können wir die Erweiterung so ausführen, daß nicht nur die Vektoren $v(\bar{p})$, sondern alle Vektoren $v(p)$ kürzer als 1 sind. $F''(p)$ sei der Endpunkt des im Punkte $F'(p)$ angebrachten Vektors $v(p)$; dann ist $F''(\bar{p}) = F(\bar{p})$ für alle $\bar{p} \subset \bar{P}$, und der Mittelpunkt m der Kugel fällt mit keinem $F''(p)$ zusammen. Ist nun $F(p)$ der Schnittpunkt des Halbstrahls $mF''(p)$ mit S^n , so erfüllt die damit erklärte Abbildung $F(P^{n+1})$ alle Anforderungen.

6. Aufzählung der Abbildungsklassen

Da auf Grund des Satzes I die Aufzählung der Klassen der Abbildungen von P^n auf S^n mit der Aufzählung der algebraischen Abbildungstypen zusammenfällt, handelt es sich hier im wesentlichen um eine algebraische Aufgabe. Wir beginnen mit einigen rein algebraischen Betrachtungen, die an diejenigen aus Nr. 4 anknüpfen.

Neben den Charakteren, die homomorphe Abbildungen einer Gruppe in die additive Gruppe der ganzen Zahlen sind und die wir jetzt als „ganze“ Charaktere bezeichnen wollen, werden noch „rationale“ Charaktere betrachtet, die homomorphe Abbildungen in die additive Gruppe der rationalen Zahlen sind. Ferner werden jetzt außer denjenigen Abelschen Gruppen mit endlichen Erzeugendensystemen, die nur Elemente unend-

licher Ordnung enthalten und die wir jetzt „freie“ Abelsche Gruppe nennen werden, auch endliche Abelsche Gruppen vorkommen und in ihnen „zyklische“ Charaktere, d. h. homomorphe Abbildungen in die additive Gruppe der Restklassen mod. 1.

Für rationale Charaktere gilt der folgende einfache Erweiterungssatz:

e) Wenn U eine Untergruppe von endlichem Index in G und wenn in U ein rationaler Charakter χ gegeben ist, so läßt sich dieser auf eine und nur eine Weise auf die ganze Gruppe G erweitern.

Beweis: Infolge der Endlichkeit des Index gibt es zu jedem $x \in G$ eine von Null verschiedene ganze Zahl m so, daß $mx = u \in U$ ist. Wenn χ für alle x erklärt ist, so ist $m \cdot \chi(x) = \chi(u)$, also $\chi(x) = \frac{1}{m} \chi(u)$; mithin ist die Erweiterung auf höchstens eine Weise möglich. Daß umgekehrt durch $\chi(x) = \frac{1}{m} \chi(u)$ ein Charakter in G erklärt wird, erkennt man wie in Nr. 4, c.

e') Falls der eben betrachtete Charakter χ für alle Elemente von U ganzzahlig ist, ist $\chi(x) \equiv \chi(y) \pmod{1}$ für je zwei Elemente x, y , die einer der Restklassen angehören, in welche G nach U zerfällt. Daher ist in der endlichen Restklassengruppe R ein zyklischer Charakter ξ durch die Bestimmung definiert, daß $\xi(X) \equiv \chi(x) \pmod{1}$ ist, falls X die das Element x enthaltende Restklasse ist. Infolge von e) ist ξ bereits durch den Charakter $\chi(U)$, und nicht erst durch $\chi(G)$, vollständig bestimmt. Wir sagen daher, daß der zyklische Charakter ξ in R durch den ganzen Charakter χ in U „induziert“ wird.

f) U sei eine Untergruppe von endlichem Index in der freien Gruppe G , R die zugehörige endliche Restklassengruppe; ξ sei ein gegebener zyklischer Charakter von R . Dann gibt es (unendlich viele) ganze Charaktere von U , die ξ induzieren.

Beweis: Es sei $G = Z_1 + \dots + Z_r$, wobei die Z_i unendliche Zyklen sind; x_i sei erzeugendes Element von Z_i , X_i sei die x_i enthaltende Restklasse. Wir setzen $\chi(x_i) = \xi_0(Z_i)$, wobei wir unter $\xi_0(Z_i)$ irgend eine bestimmte Zahl aus der Restklasse mod. 1 $\xi(Z_i)$ verstehen. Dadurch wird in G ein rationaler Charakter χ erklärt, für den $\chi(x) \equiv \xi(X) \pmod{1}$ ist, wenn x irgend ein Element von G , X die x enthaltende Restklasse ist. Ist insbesondere $x \in U$, so ist daher $\chi(x) \equiv 0 \pmod{1}$; χ ist daher in U ganz. Daß ξ durch χ induziert wird, folgt unmittelbar aus der Definition.

Wir betrachten jetzt das Polyeder P^n in einer festen Simplexzzerlegung. Unter L^n verstehen wir die Gruppe der n -dimensionalen Komplexe von P^n (in dieser Zerlegung), unter Z^n die Gruppe der n -dimensionalen Zyklen. L^n ist eine freie Gruppe, Z^n eine abgeschlossene Untergruppe von L^n . Daher gibt es nach Nr. 4, a) eine Untergruppe V^n von L^n so, daß $L^n = Z^n + V^n$ ist; die Gruppe V^n ist durch den Satz aus Nr. 4, a) nicht eindeutig bestimmt, wir wählen sie aber ein für alle Mal fest. Ferner sei R^{n-1} die Gruppe der $(n-1)$ -dimensionalen Ränder, $\bar{R}^{n-1}$ die Gruppe der „Randteiler“, d. h. derjenigen $(n-1)$ -dimensionalen Zyklen, von denen gewisse Vielfache Ränder sind; R^{n-1} ist Untergruppe von $\bar{R}^{n-1}$ mit endlichem Index, die zugehörige Restklassengruppe T^{n-1} ist die $(n-1)$ -dimensionale Torsionsgruppe. Verstehen wir für jedes Element $v^n \in V^n$ unter $\dot{v}^n$ seinen Rand, so wird, indem man jedem $v^n \in V^n$ den zugehörigen $\dot{v}^n \in R^{n-1}$ zuordnet, V^n homomorph auf R^{n-1} abgebildet; dies ist aber sogar ein Isomorphismus; denn ist $v_1^n = v_2^n$ so ist $\dot{v}_1^n - \dot{v}_2^n = (v_1^n - v_2^n)^\cdot = 0$, d. h. $v_1^n - v_2^n$ ist Zyklus, also $v_1^n - v_2^n \in Z^n$ und $v_1^n - v_2^n \in V^n$, mithin $v_1^n - v_2^n = 0$.

Es sei nun f eine Abbildung von P^n auf S^n . Verstehen wir für jeden Zyklus $z^n \in Z^n$ unter $\chi(z^n)$ den Grad der Abbildung $f(z^n)$, so ist χ ein ganzer Charakter von Z^n . Wir nehmen nun weiter an, daß f simplizial sei, und daß dabei die ursprüngliche Zerlegung von P^n oder eine ihrer Unterteilungen zugrundeliegt. Ist dann τ^n ein festes n -dimensionales Simplex der in S^n zugrundeliegenden Zerlegung, und verstehen wir für jeden Komplex $x^n \in L^n$ unter $\chi(x^n)$ den Grad der Abbildung $f(x^n)$ in τ^n , so stimmt diese Definition in Z^n mit der eben gegebenen überein, und χ ist ein ganzer Charakter von L^n . Infolge der Isomorphie zwischen V^n und R^{n-1} wird durch die Bestimmung $\dot{\chi}(\dot{v}^n) = \chi(v^n)$ auch in R^{n-1} ein ganzer Charakter $\dot{\chi}$ definiert. Durch ihn wird — gemäß e') — in T^{n-1} ein zyklischer Charakter ζ induziert, wobei ζ nach folgender Vorschrift gebildet ist: X^{n-1} sei ein Element von T^{n-1} , also eine $(n-1)$ -dimensionale Homologiekategorie, welche Randteiler enthält (Restklasse von $\bar{R}^{n-1}$ nach R^{n-1}); x^{n-1} sei einer dieser Randteiler, und es sei $mx^{n-1} = \dot{v}^n$; dann ist $\zeta(X^{n-1}) \equiv \frac{1}{m} \chi(v^n) \pmod{1}$, oder: $m \cdot \zeta(X^{n-1}) \equiv \chi(v^n) \pmod{m}$; infolge von $\dot{v}^n = mx^{n-1}$ ist v^n ein Zyklus mod. m ; die Restklasse mod. m von $\chi(v^n)$ ist der Grad mod. m der Abbildung $f(v^n)$, da $\chi(v^n)$ der Grad in dem Simplex τ^n ist. Mithin ist der zyklische Charakter ζ von T^{n-1} durch die Grade mod. m der Abbildungen der n -dimensionalen Zyklen mod. m für $m > 1$ vollständig bestimmt, und umgekehrt bestimmt ζ diese Grade eindeutig¹⁸⁾.

Demnach ist klar: sind f und g zwei simpliziale Abbildungen aus derselben Klasse, so bewirken sie sowohl denselben ganzen Charakter χ von Z^n als auch denselben zyklischen Charakter ζ von T^{n-1} ; da jede Abbildungsklasse simpliziale Abbildungen enthält, gehören daher zu jeder Klasse ein bestimmter Charakter $\chi(Z^n)$ und ein bestimmter Charakter $\zeta(T^{n-1})$. Gehören dagegen f und g verschiedenen Klassen an, so besitzen sie nach Satz I verschiedene algebraische Typen, d. h. es gibt einen n -dimensionalen Zyklus oder Zyklus mod. m , der durch sie mit verschiedenen Graden bzw. Graden mod. m abgebildet wird; folglich bewirken sie nicht sowohl denselben $\chi(Z^n)$ als auch denselben $\zeta(T^{n-1})$. Mithin entsprechen den Klassen eineindeutig Paare χ, ζ von Charakteren; umgekehrt gibt es, wenn χ und ζ willkürlich gegeben sind, Abbildungen, die diese Charaktere bewirken. Denn zunächst gibt es nach f) einen ganzen Charakter $\dot{\chi}$ der Gruppe R^{n-1} , der ζ induziert; erklären wir dann durch $\chi(v^n) = \dot{\chi}(\dot{v}^n)$ einen Charakter χ in V^n , so ist in Verbindung mit dem in Z^n gegebenen Charakter jetzt in $L^n = Z^n + V^n$ ein ganzer Charakter χ definiert. Wir konstruieren nun eine stetige Abbildung h , so daß für jeden Komplex $x^n \subset L^n$ die Zahl $\chi(x^n)$ der Grad der Abbildung $h(x^n)$ in dem festen Simplex τ^n von S^n ist: a sei ein nicht zu τ^n gehöriger Punkt von S^n ; in jedem n -dimensionalen Simplex x_i^n von P^n wählen wir $|\chi(x_i^n)|$ zueinander fremde n -dimensionale Simplexe und bilden jedes von ihnen so auf S^n ab, daß der Rand von x_i^n auf a abgebildet wird, die Abbildung im Inneren von x_i^n eineindeutig ist und den Grad $+1$ oder -1 hat, je nachdem $\chi(x_i^n)$ positiv oder negativ ist; alle übrigen Punkte von P^n bilden wir ebenfalls auf a ab. Dann ist $\chi(x_i^n)$ der Grad von $h(x_i^n)$ in τ^n für jedes Simplex x_i^n , und mithin $\chi(x^n) = \sum a_i \chi(x_i^n)$ der Grad von $h(x^n)$ für jeden Komplex $x^n = \sum a_i x_i^n$. h , sowie jede simpliziale Abbildung f aus derselben Klasse, bewirkt dann die gegebenen Charaktere χ und ζ von Z^n und T^{n-1} .

Damit ist folgendes bewiesen:

Satz III. Jede Klasse von Abbildungen des Polyeders P^n auf die Sphäre S^n bewirkt einen ganzen Charakter der n -dimensionalen Zyklengruppe Z^n von P^n und einen zyklischen Charakter der $(n-1)$ -dimensionalen Torsionsgruppe T^{n-1} von P^n durch die folgenden Festsetzungen: $\chi(Z^n)$ ist der Grad, mit dem der Zyklus $z^n \subset Z^n$ abgebildet wird;

¹⁸⁾ Man vergesse aber nicht, daß in der Wahl der Gruppe V^n eine Willkür liegt. Ohne die Auszeichnung von V^n ist, wenn x^{n-1} gegeben ist, v^n durch $mx^{n-1} = \dot{v}^n$ nicht eindeutig bestimmt, da auch $mx = (v^n + z^n)$ mit irgend einem (gewöhnlichen) Zyklus z^n ist; und im allgemeinen ist $\chi(v^n) \neq \chi(v^n + z^n)$.

ferner sei die Gruppe L^n der n -dimensionalen Komplexe als direkte Summe $L^n = Z^n + V^n$ dargestellt; ist dann X^{n-1} eine Homologiekategorie aus T^{n-1} , so gibt es in V^n einen Zyklus mod. m v_m^n , dessen Rand $\dot{v}_m^n \subset X^{n-1}$ ist; der Grad mod. m , mit dem v_m^n abgebildet wird, ist $\equiv m \cdot \zeta(X^{n-1})$ mod. m . Dies ist eine eindeutige Zuordnung zwischen den Abbildungsklassen und der Gesamtheit aller Charakterenpaare $\chi(Z^n)$, $\zeta(T^{n-1})$.

Hiernach kann man leicht die Anzahl der Klassen bestimmen:

Satz III'. Ist die n -te Bettische Zahl von P^n positiv, so gibt es unendlich viele Klassen; ist sie 0, so ist die Anzahl der Klassen endlich, und zwar gleich der Ordnung der $(n-1)$ -dimensionalen Torsionsgruppe.

Beweis: Daß die n -te Bettische Zahl positiv ist, bedeutet, daß Z^n nicht nur aus dem Nullelement besteht; sie besitzt als freie Gruppe dann unendlich viele ganze Charaktere χ ; denn man kann, wenn $Z^n = X_1 + \dots + X_r$ ist und die X_i unendliche Zyklen sind, die Werte von χ für die erzeugenden Elemente der X_i willkürlich vorschreiben. Ist die n -te Bettische Zahl 0, so besteht Z^n nur aus der 0, und $\chi = 0$ ist der einzige Charakter von Z^n . Man hat also zu zeigen, daß die Anzahl der zyklischen Charaktere einer endlichen Gruppe T^{n-1} gleich der Ordnung von T^{n-1} ist. Nun ist T^{n-1} direkte Summe endlicher zyklischer Gruppen: $T^{n-1} = X_1 + \dots + X_r$; x_i seien erzeugende Elemente der X_i , ihre Ordnungen seien e_i . Da $e_i \cdot x_i = 0$ ist, muß $e_i \cdot \zeta(x_i) \equiv 0$ mod. 1, also $\zeta(x_i) \equiv \frac{k_i}{e_i}$ mod. 1 sein, wobei k_i eine der Zahlen 0, 1, ..., $e_i - 1$ ist.

Wählt man umgekehrt die k_i willkürlich und setzt $\zeta(x_i) \equiv \frac{k_i}{e_i}$ für $i = 1, \dots, r$, so entsteht ein zyklischer Charakter von T^{n-1} . Daraus folgt, daß die Anzahl dieser Charaktere gleich $\prod e_i$, also gleich der Ordnung von T^{n-1} ist.

Als Spezialfall des Satzes III' sei noch hervorgehoben:

Satz III". Die Abbildungen von P^n auf die S^n bilden dann und nur dann eine einzige Klasse, wenn für P^n die n -te Bettische Zahl 0 und keine $(n-1)$ -dimensionale Torsion vorhanden ist¹⁹⁾.

(Eingegangen den 12. März 1932)

¹⁹⁾ Satz I der unter 6) zitierten Arbeit.

Asymptotische Abschätzung des absoluten Betrages einer Funktion, die die Werte 0 und 1 nicht annimmt

Von ALEXANDER OSTROWSKI, Basel

Es sei $p(z)$ eine Funktion, die für $|z| < 1$ regulär und von 0 und 1 verschieden ist. Dann gilt bekanntlich nach *Schottky* die Ungleichung

$$(a) \quad \left| p(z) \right| < \Omega(p_0, r), \quad p_0 = p(0), \\ |z| \leq r < 1$$

wo Ω nur von r und p_0 , sogar nur von r und einer Schranke für $|p_0|$ abhängt. Daraus folgen insbesondere die Abschätzungen

$$(b) \quad |p^{(v)}(0)| \leq L_v(p_0), \quad v = 1, 2, \dots,$$

wo auch L_v nur von einer Schranke für $|p_0|$ abhängen.

Ueber Ω ist durch Hrn. *Landau*¹⁾ bekannt, daß es sich durch

$$(a) \quad \Omega(p_0, r) \leq (A|p_0| + B)^{\frac{C}{1-r}}$$

abschätzen läßt, mit absoluten positiven Konstanten A, B, C ; ferner durch Hrn. *Valiron*, daß sogar

$$(b) \quad \Omega(p_0, r) \leq (A(|p_0| + 1))^{\frac{1+r}{1-r}}$$

gilt für eine gewisse absolute Konstante A , deren Wert nicht abgeschätzt wurde²⁾. Andererseits kann man mit einer elementaren (d. h. in diesem Falle nur algebraische Hilfsfunktionen benutzenden) Methode beweisen: Für $|p_0| \leq e^d$, $d \geq 0$ gilt

$$(c) \quad \Omega(p_0, r) \leq 1 + e^{-16} e^{12,1 \cdot \frac{1}{1-r} + 2,7 \left(\lg \left(\frac{e}{1-r} \right) \right) \frac{1}{1-r} \frac{4,5}{e^{1-r}} d},$$

¹⁾ Vgl. den von Hrn. *Landau* herrührenden § 6 in der Abhandlung von *Bohr* und *Landau*, Gött. Nachr. 1910, pp. 303—330, übrigens auch die Formel (54) in § 12 der Abhandlung von Hrn. *Landau*: Ueber den Picard'schen Satz. Vierteljahresschr. d. Zürch. Naturf. Ges. 1906.

²⁾ Vgl. *G. Valiron*, Bull. d. Sc. Math., Bd. 51 (1927).

wo also die Größenordnung in $1 - r$ schlechter ist, wohl aber die numerischen Schranken verhältnismäßig kleine Werte haben, was beim Gebrauche für kleinere r , z. B. $r = \frac{1}{2}$ von Wichtigkeit ist³⁾.

Wir wollen nun mit $S^*(\alpha, r)$ die obere Grenze der absoluten Beträge von $|p(z)|$ für $|z| \leq r < 1$ für alle Funktionen $p(z)$ mit $|p(0)| = \alpha$ bezeichnen, die für $|z| < 1$ regulär und $\neq 0$, $\neq 1$ sind; ferner mit $S(\alpha, r)$ die obere Grenze von $|p(z)|$ für $|z| \leq r < 1$ für alle solche Funktionen mit $p(0) = \alpha$; endlich mit $\Phi(\alpha, r)$ die obere Grenze von

$$|\arg p(z) - \arg p(0)|$$

für $|z| \leq r < 1$ für eine solche Funktion $p(z)$ mit $p(0) = \alpha$ (natürlich bei stetiger Fortsetzung des Arguments). Dann kann man die obigen Abschätzungen (α), (β), (γ) als Abschätzungen von $S^*(\alpha, r)$ und $S(\alpha, r)$ nach oben auffassen. Für $\Phi(\alpha, r)$ ist durch Hrn. P. Lévy⁴⁾ bekannt

$$(\delta) \quad \Phi(\alpha, r) \leq \frac{c(\alpha)}{1-r}, \quad c(\alpha) > 0.$$

An Abschätzungen nach unten ist bisher nur bekannt (durch Hrn. P. Lévy, l. c.)

$$(\epsilon) \quad |\lg S(\alpha, r) + i \Phi(\alpha, r)| \geq \frac{c_1(\alpha)}{1-r}, \quad c_1(\alpha) > 0.$$

Nun sind für L_v asymptotisch genaue Schranken bekannt durch den (übrigens elementar, d. h. ohne Benutzung der Modulfunktion beweisbaren) Satz:

Ist für $p_0 = \alpha$, $\varphi_v(\alpha)$ die obere Grenze aller $|p^{(v)}(0)|$, so gilt für $\alpha \rightarrow \infty$ und ebenso für $\alpha \rightarrow 0$

$$(b^*) \quad \frac{\varphi_v(\alpha)}{|\alpha| |\lg |\alpha||^v} \rightarrow \frac{2^v}{v!} {}^5).$$

Es sollen nun im folgenden in demselben Sinne asymptotisch genaue Schranken für S^* und S ermittelt werden. Hier handelt es sich aller-

³⁾ Vgl. A. Ostrowski, Studien über den Schottky'schen Satz, Rektoratsprogramm der Univ. Basel für 1931 (als selbständige Schrift erschienen im Verlag von B. Wepf & Cie., Basel), pp. 96—102.

⁴⁾ Vgl. P. Lévy, Bull. Soc. Math. d. France, 1912. Einen anderen Beweis gibt J. E. Littlewood, Proc. Lond. Math. Soc. (2), 23 (1924), pp. 490, 509—510.

⁵⁾ Vgl. A. Ostrowski, Berliner Sitzungsberichte, 1925, (math.-physik. Klasse) pp. 483—484. Für $v=1$ rührt diese Formel von Hrn. Landau her (Vierteljahrsschrift Zürich, Nat. Ges. 51, 1906), dessen Resultat später von Gronwall, Paris C. R. 155 (1912), pp. 764—766 und Bernays, Zürich, Vierteljahrsschr. 58 (1913), pp. 203—238 auf anderem Wege hergeleitet wurde

dings um Funktionen von zwei Variablen (a, r) bzw. (α, r) . Dementsprechend sind verschiedene Auffassungen des asymptotischen Verhaltens möglich.

Wir werden im ersten Teil dieser Mitteilung (§§ 1—3, Formelnummern 1—20) r fest ($0 < r < 1$) annehmen und $|p_0| \rightarrow \infty$ gehen lassen. Es ergibt sich eine asymptotische Relation für $S^*(a, r)$:

$$(A) \quad 16 S^*(|p_0|, r) \sim (16 |p_0|)^{\frac{1+r}{1-r}}, |p_0| \rightarrow \infty,$$

und zwar *gleichmäßig* in r . Hieraus folgt insbesondere, daß der *Valiron'sche* Exponent $\frac{1+r}{1-r}$ in (β) sogar in dem Sinne der genaue ist, daß er sich für kein einziges r mit $0 < r < 1$ verkleinern läßt, selbst auf Kosten der Vergrößerung der Zahlenkoeffizienten.

Der Beweis beruht wesentlich auf der Benutzung der elliptischen Modulfunktion. Zur Untersuchung von $p(z)$ ist die Heranziehung von $\lambda(z)$, des *Legendre'schen* Modulquadrates, also einer Modulfunktion 2^{ter} Stufe notwendig. Andererseits hat die absolute Invariante $J(z)$ den Vorteil, daß ihre Entwicklung nach Potenzen von $q = e^{\pi iz}$ *positive* Koeffizienten hat. Wir kombinieren daher die Betrachtung beider Funktionen. Dies bringt mit sich, daß wir zugleich — und sogar zuerst — die analogen Abschätzungen für Funktionen $m(z)$ herleiten, die für $|z| < 1$ regulär, den Wert 0 nur in dreifacher Mehrfachheit, den Wert 1 nur in doppelter Mehrfachheit annehmen. Es ergeben sich ganz analoge Resultate, nur muß die Konstante 16 durch 1728 ersetzt werden⁶⁾.

Es ist übrigens von prinzipiellem Interesse, daß die ganze Untersuchung sogar unter Benutzung der absoluten Invariante $J(z)$ allein durchgeführt werden kann, wenn man vom folgenden Satz ausgehen will, der ja aus der Theorie der Modulfunktionen leicht herzuleiten ist, aber auch elementar nicht schwer zu beweisen ist: Es gilt die Relation

$$(c) \quad m(z) = \frac{4}{27} \frac{(p^2(z) - p(z) + 1)^3}{p^2(z) (1 - p(z))^2},$$

in dem Sinne, daß wenn rechts ein $p(z)$ eingesetzt wird, links ein $m(z)$ herauskommt und umgekehrt, wenn man die Gleichung bei gegebenem

⁶⁾ Bei *Valiron* und in einer früheren Untersuchung von *Landau* (l. c.) wird zur Abschätzung von $p(z)$ nur die J -Funktion benutzt, sodaß die betreffenden Abschätzungen sich zugleich auf alle $m(z)$ -Funktionen beziehen. Zum Nachweis aber, daß die so gefundenen Abschätzungen für $p(z)$ die „besten“ sind, ist prinzipiell die Heranziehung von $\lambda(z)$ nötig.

$m(z)$ nach $p(z)$ auflöst, ergeben sich 6 verschiedene Funktionen $p(z)$, die auseinander mit Hilfe der bekannten Ausdrücke $1 - p$, $\frac{1}{p}$, $1 - \frac{1}{p}$, $\frac{p}{p-1}$, $\frac{1}{1-p}$ hervorgehen.

Mit Hilfe dieses Satzes kann man erstens das Analogon des *Schottky'schen* Satzes für Funktionen $m(z)$ direkt und elementar aus dem elementar beweisbaren *Schottky'schen* Satz für $p(z)$ folgern, während der direkte elementare Beweis für $m(z)$ mit keiner der bisherigen Methoden geht, wobei man ja $\lg m(z)$ oder $\sqrt[n]{m(z)}$ betrachten müßte. Sodann aber lassen sich auf analogem Wege auch die Fälle quantitativ genau untersuchen, die verschiedenen *Schwarz'schen* Dreiecksfunktionen entsprechen, und insbesondere auch für Scharen meromorpher Funktionen, die Hr. *Montel* daraus gewonnen hat, quantitative Abschätzungen finden. Darauf hoffe ich in späteren Mitteilungen eingehen zu können.

Die Benutzung der (c) entsprechenden Relation zwischen $J(z)$ und $\lambda(z)$ stellt den eigentlichen Gedanken dar, auf dem die Betrachtungen des ersten Teiles dieser Mitteilung beruhen. Im Uebrigen lassen sich diese Ueberlegungen durchführen, ohne daß man auf Einzelheiten der zu $J(z)$ und $\lambda(z)$ gehörenden Modulfiguren eingeht. Ueberdies reicht die Formel (A) für viele Zwecke aus.

Ferner legt (A) die Frage nahe, ob die gleiche Relation auch für $S(\alpha, r)$ mit $\alpha \rightarrow \infty$ gültig bleibt. Endlich kann man fragen, ob für feste α der Grenzwert $(1-r) \log S(\alpha, r)$ für $r \rightarrow 1$ existiert und bestimmt werden kann. Beide Fragen werden in bejahendem Sinne im zweiten Teil dieser Mitteilung (§§ 4–8, Formelnummern 21–52) beantwortet. Zugleich wird auch die Frage nach dem Verhalten von $S(\alpha, r)$ für $\alpha \rightarrow 0$ entschieden. In dieser Beziehung lautet das einzige in der Literatur zu findende Resultat, das (1927 l. c.) von *G. Valiron* gefunden wurde:

$$S(p_0, r) < B^{\frac{1}{1-r}} |p_0|^{\frac{1-r}{1+r}}.$$

Andererseits ließe sich aus der auch hierbei ihre Geschmeidigkeit neu erweisenden *Montel'schen* Theorie der normalen Funktionenfamilien leicht direkt folgern, daß $S(\alpha, r) \rightarrow 0$ mit $\alpha \rightarrow 0$ gelten muß. Wir finden nun, daß

$$(B) \quad \frac{S(p_0, r)}{16} \sim \left| \frac{p_0}{16} \right|^{\frac{1-r}{1+r}}$$

gilt für $|p_0|^{1-r} \rightarrow 0$ gleichmäßig in p_0 und r , wobei auch das Fehlerglied sich weiter abschätzen läßt. Man kann also in der *Valironschen* Formel $B^{\frac{1}{1-r}}$ durch eine *absolute Konstante* ersetzen. Was nun das Verhalten von $S(\alpha, r)$ für $\alpha \rightarrow \infty$ oder $r \rightarrow 1$ anbetrifft, so bezeichnen wir mit $\omega(z)$ denjenigen Zweig der Umkehrfunktion von $\lambda(z)$, dessen Werte im Fundamentalbereiche von $\lambda(z)$ liegen, der zwischen den Halbstrahlen $\Im z > 0$, $\Re z = \pm 1$ und den Halbkreisen $|z \pm \frac{1}{2}| = \frac{1}{2}$, $\Im z > 0$ liegt. Dann lautet unser wesentlichstes Ergebnis, das als Verallgemeinerung von (A) aufzufassen ist:

$$(I) \quad S(p_0, r) \sim \frac{1}{16} e^{\pi \Im \omega(p_0) \frac{1+r}{1-r}}, \quad \Im \omega(p_0) \frac{1+r}{1-r} \rightarrow \infty,$$

und sogar noch schärfer:

$$(I^*) \quad S(p_0, r) - \frac{1}{16} e^{\pi \Im \omega(p_0) \frac{1+r}{1-r}} = o(1), \quad \Im \omega(p_0) \frac{1+r}{1-r} \rightarrow \infty.$$

Daraus folgt insbesondere für feste p_0

$$(A) \quad (1-r) \lg S(p_0, r) \rightarrow 2\pi \Im \omega(p_0), \quad r \uparrow 1.$$

Und für feste r und $p_0 \rightarrow \infty$ ergibt sich aus (I)

$$(A^*) \quad 16 S(p_0, r) \sim (16 |p_0|)^{\frac{1+r}{1-r}}, \quad p_0 \rightarrow \infty,$$

gleichmäßig in r und $\arg p_0$. Für die genaue Formulierung der zum Teil noch schärferen Ergebnisse vgl. man in § 8 die Formeln I—VI**.

Methodisch ist aber zum zweiten Teil dieser Mitteilung folgendes zu bemerken: Es ist seit der ersten Arbeit von Hrn. *Carathéodory* über den *Picard'schen* Satz bekannt, daß $S(p_0, r)$ gleich dem $\max |\lambda(z)|$ für alle Punkte einer gewissen Kreislinie K ist, die ganz in der oberen Halbebene verläuft und den Punkt $\omega(p_0)$ zu ihrem „nicht euklidischen Mittelpunkt“ hat. Und die Höhe μ des höchsten Punktes von K über der reellen Axe ist gleich $\Im \omega(p_0) \frac{1+r}{1-r}$. Man würde also ohne weiteres eine Abschätzung von $S(p_0, r)$ erhalten durch das Maximum von $|\lambda(z)|$ auf der Strecke $\Im z = \mu$, $|\Re z| \leq 1$, wenn nicht auch Punkte der unteren

Kreishälfte in der Nähe gewisser Punkte der reellen Axe liegen würden, wo $\lambda(z)$ auch ∞ wird. Man kann nun dem so begegnen, daß man solche Punkte von K , die außerhalb des oben angegebenen Fundamentalbereiches liegen, durch eine Substitution der zu $\lambda(z)$ gehörenden Gruppe in den Fundamentalbereich bringt. Dann muß man aber die Höhe der „reduzierten“ Punkte über der x -Axe kennen und insbesondere mit μ vergleichen können. Die Frage läßt sich so wenden, daß man sämtliche Kreise betrachtet, in die K durch die hier zulässigen Modulsubstitutionen übergeführt wird, und nach dem Maximum der zugehörigen Höhen μ' fragt. Und die Wendung, die damit der Frage gegeben wird, findet nun ihre Rechtfertigung im Resultat, daß *alle sich so ergebenden Höhen die Höhe des höchsten Punktes von K nicht überschreiten können*, sofern der nichteuklidische Mittelpunkt von K im obigen Fundamentalbereich liegt. Mit dieser Tatsache (Hilfssatz B in § 4), deren Beweis nachträglich sehr leicht zu führen ist, ist die Hauptschwierigkeit überwunden, die wohl bisher einer genauen Durchrechnung der asymptotischen Werte von S im Wege stand⁷⁾. Eine ähnliche Tatsache (Hilfssatz A § 4) gilt übrigens auch für die gesamte Modulgruppe, die zur Funktion $J(z)$ gehört. Es dürften sich damit auch analoge Abschätzungen für die Funktion $m(z)$ herleiten lassen. Obgleich man, was die Form dieser Abschätzung anbetrifft, einige neue und interessante Momente erwarten dürfte, bin ich dieser Frage nicht mehr nachgegangen. Ich hoffe, daß darüber demnächst eine Mitteilung von anderer Seite erfolgen wird.

Daß aber die Betrachtung der Strecke $\Im z = \mu$ auf den asymptotisch genauen Wert von $S(p_0, r)$ führt, ist dem glücklichen Umstand zu verdanken, daß das asymptotische Verhalten von $\lambda(z)$ für $\Im z \uparrow \infty$ in der hier in betracht kommenden Näherung unabhängig vom Realteil von z ist.

Analoge sehr genaue Abschätzungen lassen sich auch für $\Phi(p_0, r)$ — die Argumentschranke von $p(z)$ — aufstellen. Man findet für $r \uparrow 1$

$$(A_1) \quad (1 - r) \Phi(p_0, r) \rightarrow \pi \max \left(\Im \omega(p_0), \Im \frac{-1}{\omega(p_0)} \right),$$

⁷⁾ Ein Versuch einer ähnlichen Reduktion des Problems findet sich am Schlusse der bekannten Arbeit von *G. Pick* (Ueber eine Eigenschaft der konformen Abbildung kreisförmiger Bereiche, Math. Ann. Bd. 77 (1916), pp. 1—6), in der zum ersten Mal die „nichteuklidische“ Auffassung des Schwarzschen Lemmas herausgearbeitet wurde. Doch sind die bezüglichen Angaben von *Hrn. Pick* nicht stichhaltig, da die Bemerkung auf p. 5 unten: „weil $|\lambda(z)|$ auf Parallelen zur Axe der imaginären Zahlen nach oben zunimmt“, sicher unzutreffend ist, wenn das betreffende Stück jener Parallelen aus dem Fundamentalbereich austritt und in die Nähe von Unendlichkeitsstellen von $\lambda(z)$ auf der reellen Axe kommt.

und noch schärfer

$$(A_2) \quad \left| \Phi(p_0, r) - \frac{2\pi r}{1-r^2} \max \left(\Im \omega(p_0), \Im \frac{-1}{\omega(p_0)} \right) \right| \leq 11.$$

Zum Beweis dieser Abschätzungen haben wir indessen sehr ausführlich auf die geometrische Struktur der Modulfigur eingehen müssen. Und da eine Darstellung dieser Entwicklungen im Rahmen der vorliegenden Mitteilung keine Abkürzung der gesamten Darstellung bedeuten würde, soll der Beweis von (A_1) und (A_2) an einer anderen Stelle (in der Math. Zeitschrift, unter dem Titel: *Asymptotische Abschätzung der Argumentvariation einer Funktion, die die Werte 0 und 1 nicht annimmt*) erscheinen.

§ 1. Funktionen $J(z)$ und $\lambda(z)$

Wir gehen von der elliptischen Modulfunktion $J(z)$ aus, von der wir folgende bekannte und leicht herleitbare Eigenschaften voraussetzen:

Die Funktion $J(z)$ ist in der oberen Halbebene $\Im z > 0$ definiert, durchweg regulär und genügt insbesondere der Relation

$$(1) \quad J(z) = J\left(-\frac{1}{z}\right).$$

Setzt man $e^{\pi iz} = q$, so gilt für $J(z)$ die bekannte Darstellung

$$(2) \quad J(z) = \frac{1}{q^2} \left(\frac{1}{12} + 20 \sum_{n=1}^{\infty} n^2 \frac{q^{2n}}{1-q^{2n}} \right)^3 \prod_{n=1}^{\infty} (1-q^{2n})^{-24},$$

deren rechte Seite mit q^2 multipliziert eine für $|q| < 1$ konvergente Potenzreihe mit *positiven* Koeffizienten und dem Anfangsgliede $\frac{1}{12^3}$ darstellt, also insbesondere monoton in q^2 ist. Es folgt daher für $|q| \downarrow 0$, $\Im z \uparrow \infty$, $P \uparrow \infty$:

$$(3) \quad 12^3 q^2 J(z) \rightarrow 1, \quad 12^3 e^{-2\pi P} J(iP) \downarrow 1.$$

Für rein imaginäre ω ist $J(\omega)$ positiv. Es gilt wegen der Positivität der Koeffizienten von (2)

$$(4) \quad |J(z)| \leq J(i \Im z).$$

Endlich sei noch erwähnt, daß $J(z)$ alle endlichen Werte im „Fundamentalbereich“

$$(5) \quad |z| > 1, -\frac{1}{2} < \Re z \leq \frac{1}{2} \text{ und } |z| = 1, 0 \leq \Re z \leq \frac{1}{2}$$

annimmt, und zwar jeden von 0 und 1 verschiedenen Wert genau einmal, den Wert 0 im „Eckpunkte“ $\frac{1}{2} + i\frac{\sqrt{3}}{2}$, der an drei verschiedene Äquivalenzbereiche anstößt, als eine dreifache Nullstelle, und den Wert 1 im „Eckpunkte“ i , der zwei verschiedenen Äquivalenzbereichen angehört, als eine Doppelwurzel von $J(z) - 1$. Insbesondere konvergiert die einzige im Bereich (5) liegende Wurzel der Gleichung

$$(5a) \quad J(z) = w, \quad w \neq 0, \quad w \neq 1,$$

mit ins Unendliche wachsendem $|w|$ gegen ∞ (und dann gilt natürlich auch $\Im z \rightarrow \infty$) und umgekehrt. Beachtet man, daß $J(z)$ auf der imaginären Achse von i bis $i \cdot \infty$ jeden Wert ≥ 1 nur einmal annimmt, und auf der Strecke von i bis 0 dieselben Werte von neuem annimmt, so folgt, daß wenn man längs der imaginären Achse von $i \cdot \infty$ nach dem Nullpunkt geht, die reellen Werte von $J(z)$ zuerst monoton bis 1 abnehmen und sodann wieder bis ∞ anwachsen. Daraus folgt:

Ist $a > b \geq 1$ bzw. $0 < a < b \leq 1$, so gilt

$$(5b) \quad J(ib) < J(ia).$$

Neben der Modulfunktion $J(z)$ müssen wir noch eine sogenannte „Modulfunktion 2-ter Stufe“ $\lambda(z)$ betrachten, die man am einfachsten erhält, wenn man das Innere des Gebietes ($\mathfrak{H}$ in der Fig. des § 4):

$$(6) \quad 0 < \Re z < 1, \quad |z - \frac{1}{2}| > \frac{1}{2}$$

auf die obere λ -Halbebene $\Im \lambda > 0$ so konform abbildet, daß die Randpunkte 0, 1, ∞ sich selbst entsprechen, und sodann die Abbildungsfunktion durch fortgesetzte Spiegelung analytisch fortsetzt. Diese Funktion hängt mit $J(z)$ durch die folgende Relation zusammen:

$$(7) \quad J(z) = \frac{4}{27} \frac{(\lambda^2 - \lambda + 1)^3}{\lambda^2 (1 - \lambda)^2}.$$

Aus (7) folgt für ins Unendliche wachsende Werte von $\lambda(z)$ bzw. (für $|\lambda| \geq 2$) für ins Unendliche wachsende Werte von $J(z)$:

$$(8) \quad J(z) \sim \frac{4}{27} \lambda^2(z), \quad \lambda(z) \sim \frac{3\sqrt{3}}{2} \sqrt{J(z)},$$

wo natürlich eine passende Bestimmung der Wurzel zu nehmen ist⁸⁾.

Es sei endlich an die Formel erinnert:

$$(8a) \quad \lambda\left(\frac{-1}{z}\right) = \frac{1}{\lambda(z)}.$$

§2. Abschätzung von $m(z)$

Für die Anwendungen der Funktionen $J(z)$ und $\lambda(z)$ auf Probleme aus dem *Picardschen* Ideenkreis sind nun die beiden Tatsachen fundamental, die wir so formulieren können:

1. *Dafür, daß $p(z)$ eine für $|z| < 1$ reguläre und dort die Werte 0 und 1 nicht annehmende Funktion ist, ist notwendig und hinreichend, daß die Darstellung gilt:*

$$(9) \quad p(z) = \lambda(C(z)),$$

wo $C(z)$ eine für $|z| < 1$ reguläre Funktion ist, deren Imaginärteil für $|z| < 1$ stets positiv ist.

Hat $p(0) = p_0$ nicht negativen Imaginärteil, so kann und soll $C(z)$ so gewählt werden, daß $C(0) = c_0$ im Bereiche (6) liegt.

2. *Dafür, daß $m(z)$ für $|z| < 1$ regulär ist und dort in jedem Punkte 1 nur in gerader, 0 nur in durch 3 teilbarer Mehrfachheit annimmt, ist notwendig und hinreichend, daß die Darstellung gilt:*

$$(10) \quad m(z) = J(C(z)),$$

wo $C(z)$ eine für $|z| < 1$ reguläre Funktion mit $\Im C(z) > 0$ ist. Insbesondere kann und soll $C(z)$ so gewählt werden, daß $C(0) = c_0$ im Bereich (5) liegt.

Für die Herleitung des *Schottkyschen* Satzes ist weiter die folgende Tatsache wichtig, die zur Untersuchung der „*Picardschen* Probleme“ zum ersten Male von *Borel* und sodann von *Carathéodory* herangezogen und vom letzteren auf einem besonders einfachen Wege — mit Hilfe des *Schwarzschen* Lemmas — bewiesen worden ist:

⁸⁾ In der Literatur wird die λ -Funktion oft so definiert, daß sie für $\omega = 0, 1, \infty$ die Werte $1, \infty, 0$ annimmt. Diese λ -Funktion ergibt sich aus der obigen durch die Transformation

$\lambda^* = \frac{1}{1 - \lambda}$ und genügt gleichfalls der Relation (7).

Ist für $|z| < 1$ durchweg $\Im C(z) > 0$, so gilt, $\omega = C(0)$ gesetzt,

$$(11) \quad \Re \omega \frac{1 - |z|}{1 + |z|} \leq \Re C(z) \leq \Re \omega \frac{1 + |z|}{1 - |z|}.$$

Hier gilt für die Funktion

$$(12) \quad C^*(z) = iP \frac{1 + z}{1 - z}, \quad P > 0,$$

offenbar das Gleichheitszeichen für reelle $z > 0$.

Genauer liegen die Werte von $C(z)$ innerhalb des Kreises $K_{\omega, |z|}$

$$(11a) \quad |C(z) - \Re \omega - \Im \omega \frac{1 + |z|^2}{1 - |z|^2}| \leq \Im \omega \frac{2|z|}{1 - |z|^2}$$

mit dem Mittelpunkt $\Re \omega + \Im \omega \frac{1 + r^2}{1 - r^2}$ und dem Radius $\Im \omega \frac{2r}{1 - r^2}$, $|z| = r$ gesetzt. Und für die Funktion

$$(12a) \quad C^{**}(z) = x + iy \frac{1 + z}{1 - z}, \quad y > 0, \quad C^{**}(0) = \omega = x + iy$$

erfüllen die Werte von $C^{**}(z)$ für $|z| \leq r$ den Kreis $K_{\omega, r}$ vollständig.

Nun folgt aus (10) und (4)

$$|m(z)| \leq J(i \Im C(z)).$$

Ist hier $\Im C(z) \geq 1$, so folgt aus (5b) und (11) wegen des monotonen Anwachsens von $J(iu)$ für u zwischen 1 und ∞ :

$$|m(z)| \leq J\left(i \Im c_0 \frac{1 + |z|}{1 - |z|}\right).$$

Ist aber $\Im C(z) < 1$, so folgt nach (5b) und (11) wegen des monotonen Abnehmens von $J(iu)$ für u zwischen 0 und 1 unter Benutzung von (1):

$$|m(z)| \leq J\left(i \Im c_0 \frac{1 - |z|}{1 + |z|}\right) = J\left(i \frac{1}{\Im c_0} \frac{1 + |z|}{1 - |z|}\right).$$

Da aber jetzt wegen (11) $\Im c_0 \frac{1-|z|}{1+|z|} < 1$ ist, muß $\frac{1}{\Im c_0} \frac{1+|z|}{1-|z|} > 1$ sein, so daß für $\Im c_0 \geq 1$, wegen des Monotoniecharakters von $J(iz)$, $J\left(i \frac{1}{\Im c_0} \frac{1+|z|}{1-|z|}\right)$ nicht verkleinert wird, wenn $\frac{1}{\Im c_0}$ durch $\Im c_0$ ersetzt wird. Daher folgt, daß jedenfalls für $\Im c_0 \geq 1$

$$(13) \quad |m(z)| \leq J\left(i \Im c_0 \frac{1+|z|}{1-|z|}\right), \quad \Im c_0 \geq 1,$$

gilt. Wird aber insbesondere nach (12)

$$(13a) \quad m^*(z) = J(C^*(z)) = J\left(iP \frac{1+z}{1-z}\right)$$

gesetzt, so gilt für $z = r > 0$:

$$(13b) \quad m^*(r) = J\left(iP \frac{1+r}{1-r}\right), \quad 0 < r < 1.$$

Nach (3) folgt nun aus (13):

$$(14) \quad |m(z)| \leq \frac{1+\varepsilon_1}{12^3} e^{2\pi \Im c_0 \frac{1+|z|}{1-|z|}}, \quad \Im c_0 \geq 1,$$

wo $\varepsilon_1 \downarrow 0$ für $\Im c_0 \uparrow \infty$ gilt. Für $m^*(r)$ folgt aber aus der zweiten Relation (3)

$$(14a) \quad m^*(r) > \frac{1}{12^3} e^{2\pi P \frac{1+r}{1-r}}, \quad 0 < r < 1.$$

Wie hängt nun $\Im c_0$ mit $m_0 = m(0)$ zusammen? Aus (10) folgt insbesondere

$$m_0 = J(c_0).$$

Da hier c_0 im Bereiche (5) liegt, folgt aus dem oben über die Gleichung (5a) Gesagten, daß für $|m_0| \rightarrow \infty$ auch $\Im c_0 \rightarrow \infty$ gilt. Daher liefert dann (3):

$$12^3 e^{2\pi i c_0} m_0 \rightarrow 1,$$

$$12^3 e^{-2\pi \Im c_0} |m_0| \geq \frac{1}{1 + \varepsilon_2},$$

wo $\varepsilon_2 = \varepsilon_2(|m_0|) \downarrow 0$ für $|m_0| \uparrow \infty$ ist. Daraus folgt

$$e^{2\pi \Im c_0} \leq (1 + \varepsilon_2) 12^3 |m_0|,$$

und daher nach (14):

$$(15) \quad |12^3 m(z)| \leq ((1 + \varepsilon_3) 12^3 |m_0|)^{\frac{1+r}{1-r}}, \quad \varepsilon_3 \downarrow 0 \text{ für } |m_0| \uparrow \infty.$$

Man kann (15) auch so schreiben:

$$(16) \quad |m(z)| \leq 12^6 \frac{r}{1-r} ((1 + \varepsilon) |m_0|)^{\frac{1+r}{1-r}}.$$

Andererseits ist aber nach dem Obigen

$$m_0^* = m^*(0) = J(C^*(0)) = J(iP)$$

eine monotone Funktion von $P \geq 1$, die für $1 \leq P < \infty$ alle positiven Werte von $J(i)$ = 1 bis ∞ durchläuft. Daher gilt für $m_0^* \uparrow \infty$:

$$12^3 m_0^* e^{-2\pi P} \downarrow 1,$$

$$e^{2\pi P} = (1 - \varepsilon) 12^3 m_0^*, \quad \varepsilon \downarrow 0 \text{ für } m_0^* \uparrow \infty.$$

Dann liefert (14a):

$$(16a) \quad 12^3 m^*(r) > ((1 - \varepsilon) 12^3 m_0^*)^{\frac{1+r}{1-r}}, \quad \varepsilon \downarrow 0 \text{ für } m_0^* \uparrow \infty.$$

§ 3. Abschätzung von $p(z)$

Wir betrachten nun $p(z)$ für absolut große $p(0) = p_0$, und zwar zuerst mit $\Im p_0 \geq 0$, und machen Gebrauch von der Relation (9), betrachten aber zugleich für das $C(z)$ aus (9) die Funktion $m(z) = J(C(z))$. Es gilt dann nach (15):

$$|12^8 m(z)| \leq |(1 + \varepsilon) 12^8 m(0)|^{\frac{1+r}{1-r}}, \quad \varepsilon = \varepsilon(|m(0)|) \downarrow 0 \text{ mit } |m(0)| \uparrow \infty.$$

$$|z| \leq r < 1$$

Für $p_0 = \lambda(c_0) \rightarrow \infty$ folgt nun aus (8) für $m(0) = m_0$:

$$m_0 \sim \frac{4}{27} p_0^2, \quad p_0 \rightarrow \infty,$$

so daß die obige Abschätzung in

$$(17) \quad |12^8 m(z)| \leq |(1 + \varepsilon) 2^8 p_0^2|^{\frac{1+r}{1-r}}, \quad \varepsilon \downarrow 0 \text{ mit } |p_0| \uparrow \infty$$

$$|z| \leq r < 1$$

übergeht.

Aus $m(z) = J(C(z))$, $p(z) = \lambda(C(z))$ folgt weiter nach (8), sobald $|p(z)| \geq 2$ ist:

$$(18) \quad |2^8 p^2(z)| \leq |(1 + \varepsilon_1) 2^8 p_0^2|^{\frac{1+r}{1-r}}.$$

$$|z| \leq r < 1$$

Sobald aber $|p_0| > 2$ ist, gilt diese Relation auch für $|p(z)| \leq 2$, wenn man $\varepsilon_1(p_0)$ langsam genug gegen 0 abnehmen läßt, so daß wir nun ganz allgemein schließen können:

$$(19) \quad |16 p(z)| \leq |(1 + \varepsilon) 16 p_0|^{\frac{1+r}{1-r}}, \quad \varepsilon \downarrow 0 \text{ mit } |p_0| \uparrow \infty.$$

$$|z| \leq r < 1$$

Andererseits betrachten wir für die Funktion $C^*(z)$ aus (12)

$$p^*(z) = \lambda(C^*(z)) = \lambda\left(iP \frac{1+z}{1-z}\right).$$

Für positive r liegt nun der Wert von $p^*(r) = \lambda\left(iP \frac{1+r}{1-r}\right)$, wegen

der Symmetrie der konformen Abbildung, auf der Geraden $\Re z = \frac{1}{2}$ und wandert auf dieser Geraden mit P monoton ins Unendliche. Daher folgt aus (16a) wegen (8)

$$|2^8 p^{**}(r)| \geq |(1-\varepsilon) 2^8 p^{**}(0)|^{\frac{1+r}{1-r}}, \varepsilon \downarrow 0 \text{ mit } P \uparrow \infty,$$

oder

$$(20) \quad |16 p^*(r)| \geq |(1-\varepsilon) 16 p^*(0)|^{\frac{1+r}{1-r}}, \varepsilon \downarrow 0 \text{ mit } P \uparrow \infty.$$

Ist aber $\Im p_0 < 0$, so läßt sich die ganze Deduktion auf die Potenzreihe um den Nullpunkt mit konjugiert komplexen Koeffizienten anwenden, so daß dann (19) für diese Potenzreihe gilt und daher auch für $p(z)$.

Die Relation (16a) zeigt nun, daß in einer Abschätzung vom Typus

$$|m(z)| \leq |(1+\varepsilon) A_r m(0)|^{\varphi(r)}, \varepsilon = \varepsilon(|m(0)|) \downarrow 0 \text{ mit } |m(0)| \uparrow \infty.$$

für kein einziges r , $0 < r < 1$, für $\varphi(r)$ eine kleinere Zahl als $\frac{1+r}{1-r}$ gesetzt werden kann, noch für $\varphi(r) = \frac{1+r}{1-r}$ ein kleineres A_r als 12^3 .

Betrachtet man allgemein eine Abschätzung vom Typus

$$|A_r m(z)| \leq |(1+\varepsilon) B_r m(0)|^{\varphi(r)}, \varepsilon \downarrow 0, |m(0)| \uparrow \infty,$$

so kann in ihr $\varphi(r)$ für kein einziges r durch eine kleinere Zahl als $\frac{1+r}{1-r}$ ersetzt werden, wie A_r und B_r auch gewählt werden mögen.

Soll aber diese Abschätzung für $\varphi(r) = \frac{1+r}{1-r}$ für unendlich viele $r, \uparrow 1$ gelten und hängen A_r, B_r von r nicht ab, so muß $B_r \geq 12^3$ sein.

In demselben Sinne ist unsere Abschätzung (19) für $p(z)$ die beste, wenn in den obigen Formulierungen 12^3 durch 16 ersetzt wird.

§ 4. Ueber die Aequivalenz von Kreisen in der Modulfur

Wir betrachten für $\omega = x + iy$, $y > 0$ und $0 < r < 1$ den Kreis $\mathcal{K}_{\omega, r}$ um den Mittelpunkt e_0 mit dem Radius R , wo

$$(21) \quad e_0 = x + iy \frac{1+r^2}{1-r^2}, \quad R = y \frac{2r}{1-r^2}$$

ist. Dann gilt für den Abstand σ von $\mathfrak{K}_{\omega, r}$ von der x -Axe und die Höhe μ des höchsten Punktes von $\mathfrak{K}_{\omega, r}$ über der x -Axe

$$(22) \quad \sigma = y \frac{1-r}{1+r}, \quad \mu = y \frac{1+r}{1-r}, \quad \sigma \mu = (\Im e_0)^2 - R^2 = y^2.$$

Den Punkt ω nennen wir den *Pseudomittelpunkt*, die Zahl r den *Pseudoradius* von $\mathfrak{K}_{\omega, r}$). Nun gilt für jedes reelle δ

$$(23) \quad |e_0 - \delta|^2 - R^2 = |\omega - \delta|^2 = (x - \delta)^2 + y^2, \quad \frac{1-r}{1+r} = \frac{\sigma}{y} = \frac{\Im e_0 - R}{y},$$

so daß durch e_0 und R auch ω und r *eindeutig* bestimmt sind. Ist uns umgekehrt ein ganz in der oberen Halbebene verlaufender Kreis $\mathfrak{K}$ mit dem Mittelpunkt e_0 und dem Radius R gegeben, so kann man aus (21) und (23) solche Zahlen $x = \Re e_0$, $y > 0$ und r ($0 < r < 1$) bestimmen, daß $\mathfrak{K} \equiv \mathfrak{K}_{\omega, r}$ wird.

Was geschieht nun mit $\mathfrak{K}_{\omega, r}$, wenn auf die z -Ebene eine Substitution

$$(24) \quad w = \frac{az + b}{cz + d}, \quad a, b, c, d \text{ reell, } ad - bc = \Delta > 0$$

ausgeübt wird?

Man kann (24) so schreiben:

$$(25) \quad w = \frac{a}{c} - \frac{ad - bc}{c^2 \left(z + \frac{d}{c} \right)} = \alpha - \frac{\Delta}{c^2 (z - \delta)}, \quad \alpha = \frac{a}{c}, \quad \delta = -\frac{d}{c}.$$

Ist $e_0 - \delta = \varepsilon |e_0 - \delta|$, $|\varepsilon| = 1$, so schneidet der von δ nach e_0 gehende, Halbstrahl die Kreislinie $\mathfrak{K}$ in den Punkten $\delta + e_{1,2} = \delta + \varepsilon(|e_0 - \delta| \pm R)$ denen vermöge (25) die beiden Punkte (vgl. (23))

⁹⁾ In der im Inneren der oberen Halbebene geltenden nichteuklidischen Maßbestimmung ist ω der nichteuklidische Mittelpunkt von $\mathfrak{K}$, der nichteuklidische Radius von $\mathfrak{K}$ ist aber $\lg \frac{1+r}{1-r}$. Der Kreis $\mathfrak{K}_{\omega, r}$ geht aus dem Kreise um den Ursprung mit dem Radius r hervor, wenn man das Innere des Einheitskreises so auf die obere Halbebene konform abbildet, daß der Ursprung in den Punkt ω übergeht.

$$(26) \quad e_{1,2}^* = \alpha - \frac{D}{c^2} \frac{|e_0 - \delta| \mp R}{|e_0 - \delta|^2 - R^2} = \alpha - \frac{D}{c^2} \frac{|e_0 - \delta| \mp R}{|\omega - \delta|^2}$$

entsprechen. In den beiden Punkten (26) wird aber die aus $\mathcal{K}$ durch (25) hervorgehende Kreislinie $\mathcal{K}^*$ von der Geraden geschnitten, in die der Strahl von δ nach e_0 übergeht (da δ in den unendlich fernen Punkt transformiert wird), und die als Orthogonalsehne von $\mathcal{K}^*$ durch den Mittelpunkt e_0^* von $\mathcal{K}^*$ hindurch geht. Daher ergibt sich für den Mittelpunkt e_0^* und den Radius R^* von $\mathcal{K}^*$

$$\begin{aligned} e_0^* &= \frac{e_1^* + e_2^*}{2} = \alpha - \frac{D}{c^2} \frac{\overline{e_0 - \delta}}{|\omega - \delta|^2}, \quad R^* = \frac{|e_1^* - e_0^*|}{2} = \frac{|D|}{c^2} \frac{R}{|\omega - \delta|^2}, \\ \Re e_0^* &= \alpha - \frac{D}{c^2} \frac{x - \delta}{|\omega - \delta|^2} = \alpha - \frac{D}{c^2} \Re \frac{1}{\omega - \delta} = \Re \left(\alpha - \frac{D}{c^2} \frac{1}{\omega - \delta} \right), \\ \Im e_0^* &= \frac{D}{c^2} \frac{\Im e_0}{|\omega - \delta|^2} = \frac{D}{c^2} \frac{1 + r^2}{1 - r^2} \frac{y}{|\omega - \delta|^2} = \frac{1 + r^2}{1 - r^2} \Im \left(\alpha - \frac{D}{c^2} \frac{1}{\omega - \delta} \right), \\ R^* &= \frac{2r}{1 - r^2} \frac{D}{c^2} \frac{y}{|\omega - \delta|^2} = \frac{2r}{1 - r^2} \Im \left(\alpha - \frac{D}{c^2} \frac{1}{\omega - \delta} \right). \end{aligned}$$

Setzt man also

$$\omega^* = \frac{a\omega + b}{c\omega + d} = \alpha - \frac{D}{c^2(\omega - \delta)},$$

so ist

$$\Re e_0^* = \Re \omega^*, \quad \Im e_0^* = \frac{1 + r^2}{1 - r^2} \Im \omega^*, \quad R^* = \frac{2r}{1 - r^2} \Im \omega^*,$$

so daß $\mathcal{K}^*$ den gleichen Pseudoradius hat wie $\mathcal{K}$ und den Pseudomittelpunkt ω^* , der aus ω durch (24) hervorgeht¹⁰⁾.

Es sei nun G eine Gruppe von linearen Substitutionen $w = \frac{az + b}{cz + d}$ mit $ad - bc > 0$ und reellen a, b, c, d . $\mathcal{K}$ sei ein ganz in der oberen Halbebene $\Im z > 0$ liegender Kreis mit dem Pseudomittelpunkt ω und der Ordinate des höchsten Punktes μ . Man betrachte nun alle aus $\mathcal{K}$

¹⁰⁾ Die damit bewiesene Tatsache folgt offensichtlich unmittelbar daraus, daß (24) eine Bewegung in unserer „nichteuclidischen Ebene“ ist, durch die der nichteuclidische Radius invariant bleibt, der Mittelpunkt aber kovariant transformiert wird.

vermöge der Substitutionen von G hervorgehenden Kreise und insbesondere die Ordinaten ihrer höchsten Punkte. Ist dann die obere Grenze M dieser Ordinaten zu bestimmen, so beachte man, daß, da alle diese Kreise den gleichen Pseudoradius haben, jene Ordinate wegen des Ausdrucks (22) für μ um so größer wird, je größer die Ordinate des Pseudomittelpunktes ist. Ist also die obere Grenze von $\Im \frac{a\omega + b}{c\omega + d}$ für alle Substitutionen von G gleich η , so gilt

$$(26a) \quad M = \mu \frac{\eta}{\Im \omega}.$$

Ist nun z. B. *erstens* die Gruppe G die Gesamtheit aller ganzzahligen Substitutionen mit der Determinante 1, so ist bekannt, daß es unter den Transformaten einer Zahl ω mit $\Im \omega > 0$ stets eine und nur eine gibt, die im Bereich (5) liegt, und zwar hat dieses „reduzierte“ ω nach Hurwitz die *maximale* Ordinate ¹¹⁾. Daraus folgt:

Hilfssatz A. Ist $\mathbb{K}$ ein Kreis, der ganz innerhalb der obern Halbebene liegt, und liegt sein Pseudomittelpunkt im Bereich (5), ist ferner die Ordinate des höchsten Punktes von $\mathbb{K}$ gleich μ , so sind die Ordinaten aller Transformaten sämtlicher Punkte von $\mathbb{K}$ vermöge der Substitutionen von G höchstens gleich μ .

Es sei *zweitens* Γ die sogenannte „Hauptkongruenzgruppe“ zweiter Stufe, d. h. die Gesamtheit aller Substitutionen

$$(27) \quad w = \frac{az + b}{cz + d}$$

mit ganzen a, b, c, d und $ad - bc = 1$,

$$a \equiv d \equiv 1 \pmod{2}, \quad b \equiv c \equiv 0 \pmod{2}.$$

Zu dieser Gruppe gehört die oben benutzte $\lambda(z)$ -Funktion als Invariante (d. h. es gilt $\lambda\left(\frac{az + b}{cz + d}\right) = \lambda(z)$ für alle Substitutionen von Γ), sowie die in der Figur a. p. 72 ausschnittsweise dargestellte Modulfigur. Diese Figur entsteht aus dem *Hauptdreieck* $\mathbb{H}$, $0 < \Re z < 1$, $|z - \frac{1}{2}| > \frac{1}{2}$ durch fortgesetzte Spiegelung an den drei Seiten und besteht aus lauter nullwinkligen Dreiecken, die die obere Halbebene einfach und lückenlos überdecken. $\lambda(z)$ bildet $\mathbb{H}$ auf die obere λ -Halbebene $\Im \lambda > 0$ ab, jedes

¹¹⁾ Hurwitz, Dissertation, Math. Werke, Bd. I, p. 6.

der an $\mathcal{H}$ angrenzenden Dreiecke wird auf die untere λ -Halbebene $\Im \lambda < 0$ abgebildet und so fort. Insbesondere wird das in der Figur mit $\mathcal{H}'$ bezeichnete Dreieck auf die untere λ -Halbebene abgebildet.

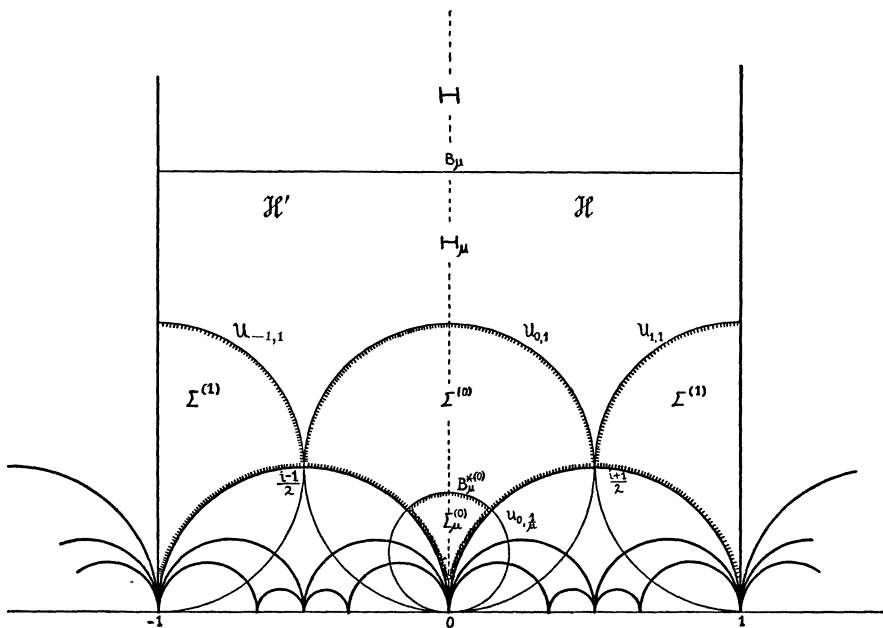


Fig. 1.

Durch jede Substitution von Γ wird nun unsere ganze Modulfigur in sich übergeführt, jedes Dreieck D der Figur geht durch eine eindeutig bestimmte Substitution von Γ entweder in $\mathcal{H}$ oder in $\mathcal{H}'$ über. Ist (27) diese Substitution, so wird dabei offenbar $z = -\frac{d}{c}$ in den unendlich fernen Punkt von $\mathcal{H}$ (oder $\mathcal{H}'$) übergeführt. — Fügt man die Kreisbogendreiecke $\mathcal{H}$ und $\mathcal{H}'$ aneinander, läßt aber dann die Strecke $\Re z = -1$, sowie den Halbkreis $|z + \frac{1}{2}| = \frac{1}{2}$ bis auf $z = 0$ fort, so ergibt sich ein sogenannter Fundamentalbereich H der Gruppe Γ , so daß jeder Punkt z mit $\Im z > 0$ einem und nur einem Punkt aus H vermöge Γ äquivalent ist. Ist nun ω ein Punkt aus H und geht daraus vermöge einer Substitution (27) von Γ ein Punkt ω^* hervor, so gilt, wie eine kurze Rechnung zeigt,

$$(28) \quad \Im \omega^* = \frac{\Im \omega}{|c\omega + d|^2}.$$

Nun ist aber für jeden Punkt ω aus H $|c\omega + d| \geq 1$ für $c \equiv 0, d \equiv 1 \pmod{2}$. Denn es genügt beim Beweis $\Re \omega \geq 0$ anzunehmen, da man sonst ω, c durch $-\bar{\omega}, -c$ ersetzen kann. Für $\Re \omega \geq 0$ ist nun

$$\begin{aligned} |c\omega + d|^2 &= (c(\omega - \tfrac{1}{2}) + \tfrac{c}{2} + d)(c(\overline{\omega - \tfrac{1}{2}}) + \tfrac{c}{2} + d) = \\ &= c^2 |\omega - \tfrac{1}{2}|^2 + 2c \left(\tfrac{c}{2} + d \right) (\Re \omega - \tfrac{1}{2}) + \left(\tfrac{c}{2} + d \right)^2, \end{aligned}$$

wegen $|\omega - \tfrac{1}{2}| \geq \tfrac{1}{2}, -\tfrac{1}{2} \leq \Re \omega - \tfrac{1}{2} \leq \tfrac{1}{2}$ ist dies aber

$$\geq \frac{c^2}{4} \pm c \left(\frac{c}{2} + d \right) + \left(\frac{c}{2} + d \right)^2 = \left(d + \frac{c}{2} \pm \frac{c}{2} \right)^2 \geq 1,$$

da c gerade, d ungerade ist. Daher folgt aus (28) $\Im \omega^* \leq \Im \omega$, so daß auch für die Gruppe Γ (und den Fundamentalbereich H) ein Analogon der Hurwitzschen Eigenschaft gilt. Nun können wir aus dem Obigen schließen:

Hilfssatz B. Ist $\mathcal{K}$ ein ganz in $\Im z > 0$ verlaufender Kreis mit dem Pseudomittelpunkt aus H und der Ordinate des höchsten Punktes μ , so sind die Ordinaten aller Transformaten der Punkte von $\mathcal{K}$ vermöge der Substitutionen von Γ höchstens gleich μ .

§ 5. Die parabolischen Umgebungen von rationalen Punkten in der Modulfigur

Als eine parabolische s -Umgebung ($s > 0$) des unendlich fernen Punktes bezeichnen wir die Halbebene $\Im z > s$. Diese Umgebung ist invariant gegenüber den linearen Substitutionen der ganzen Modulgruppe, die den unendlich fernen Punkt in sich überführen und also die Gestalt haben $w = z + b, b$ ganz.

Führt eine Substitution von G :

$$(29) \quad w = \frac{az + b}{cz + d} = \alpha - \frac{1}{c^2(z - \delta)},$$

a, b, c, d ganz, $ad - bc = 1, \alpha = \frac{a}{c}, \delta = -\frac{a}{c}, s = \infty$ in $w = \alpha = \frac{a}{c}$ über, so geht die Halbebene $\Im z > s$ in das Innere einer Kreislinie $U_{\alpha, s}$ über, die die reelle Axe im Punkte α berührt. Man erhält den

Durchmesser dieses Kreises, wenn man für die Punkte $z = x + is$,
 $-\infty < x < \infty$, die entsprechenden Punkte $w = \alpha - \frac{1}{c^2(x + is - \delta)}$ von
 $U_{\alpha, s}$ bildet und die größte Distanz eines solchen Punktes von α :

$$(30) \quad \max |w - \alpha| = \max \frac{1}{c^2 |x + is - \delta|}$$

bestimmt. (30) ist offenbar gleich

$$\frac{1}{sc^2}$$

und hängt nur von c ab, dagegen nicht von der speziellen Substitution
 (29), die $z = \infty$ in $w = \alpha$ überführt. Wir bezeichnen das Innere des in
 α berührenden Kreises $U_{\alpha, s}$ mit dem Durchmesser $\frac{1}{sc^2}$ als die parabolische
 s -Umgebung von α (größeren Werten von s entspricht also eine „kleinere“
 Umgebung)¹²⁾. Führt allgemein eine Substitution (S) $w = \frac{az + b}{cz + d}$ einen
 reellen rationalen Punkt α in einen andern α' über, so wird durch sie

¹²⁾ An den Begriff dieser parabolischen Umgebungen knüpft das folgende Analogon des
 Hilfssatzes B in § 4 an:

Ist $U = U_{0, a}$ ein die reelle Axe im o-Punkt berührender Kreis mit dem Durchmesser
 $\frac{1}{a}$ und berührt U einen in $\Im z > 0$ verlaufenden Kreis K mit dem Pseudomittelpunkt
 ω aus H von außen, so dringt keiner der Kreise, die aus K durch Substitutionen von
 Γ hervorgehen, in das Innere von U ein. Beweis: Wir bezeichnen allgemein den Kreis,
 der aus einem Kreis K durch $w = \frac{pz + q}{sz + r}$ entsteht, durch $\frac{pK + q}{sK + r}$. Nach Voraussetzung
 liegt der Kreis $P = \frac{-1}{K}$ im Bereich $\Im z \leq a$. Ist $w = \frac{\alpha z + \beta}{\gamma z + \delta}$ eine Substitution von Γ ,
 so ist zu beweisen, daß $\frac{\alpha K + \beta}{\gamma K + \delta}$ nicht in U eindringt, d. h., daß $\frac{-\gamma K - \delta}{\alpha K + \beta}$ im Bereich
 $\Im z \leq a$ liegt.

Nun ist aber

$$\frac{-\gamma K - \delta}{\alpha K + \beta} = \frac{\frac{\gamma}{P} - \delta}{\frac{-\alpha}{P} + \beta} = \frac{\delta P - \gamma}{-\beta P + \alpha},$$

und da $w = \frac{\delta z - \gamma}{-\beta z + \alpha}$ zu Γ gehört und $\frac{-1}{\omega}$ mit ω in H liegt, folgt dies aus dem
 Hilfssatz B.

auch $U_{\alpha,s}$ in $U_{\alpha',s}$ übergeführt, und zwar für jedes $s > 0$. Denn man kann S „auf Umwege über den unendlich fernen Punkt“ ausführen. Die Gesamtheit der parabolischen s -Umgebungen aller reellen rationalen Punkte geht also durch jede ganzzahlige lineare Substitution von der Determinante 1 in sich über. Wir zeigen nun ferner, daß für $s = 1$ die s -Umgebungen von zwei verschiedenen rationalen Punkten $\alpha = \frac{a}{c}, \beta = \frac{b}{d}$, $(a, c) = (b, d) = 1$, keine Punkte gemeinsam haben können¹³⁾. Dann können erst recht für $s > 1$ die parabolischen s -Umgebungen von verschiedenen rationalen Punkten keine Punkte gemeinsam haben.

Es genügt zu zeigen, daß die Mittelpunktsdistanz von $U_{\alpha,1}$ und $U_{\beta,1}$ nicht kleiner ist als die Summe der Radien, d. h.:

$$(\alpha - \beta)^2 + \left(\frac{1}{2c^2} - \frac{1}{2d^2} \right)^2 \geq \left(\frac{1}{2c^2} + \frac{1}{2d^2} \right)^2,$$

$$(\alpha - \beta)^2 = \left(\frac{ad - bc}{cd} \right)^2 \geq \frac{1}{c^2 d^2},$$

was wegen $ad - bc \neq 0$ ($\alpha \neq \beta$) sicher richtig ist.

Wir betrachten nun den Durchschnitt $\Sigma^{(0)}$ von H und $U_{0,1}$. Ferner sei die Vereinigungsmenge der Durchschnitte von $U_{\pm 1,1}$ mit H (vgl. Fig.) mit $\Sigma^{(1)}$ bezeichnet.

Es möge nun ein Kreis $K = \mathbb{K}_{\omega,r}$ den Pseudomittelpunkt ω in $\Sigma^{(0)}$ haben, und die Höhe des höchsten Punktes von K sei gleich μ . Wir wollen nun das größte $s = \frac{1}{\mu^*}$, d. h., die kleinste parabolische s -Umgebung $U_{0,s}$ von 0 bestimmen, für die das Innere von K ganz im Inneren von $U_{0,s}$ liegt. Dann muß die Radiendifferenz $\geq$ der Mittelpunktsdistanz der beiden Kreise sein, d. h., da der Radius von $\mathbb{K}_{\omega,r}$ nach (21) gleich $R = \gamma \frac{2r}{1 - r^2}$ ist,

¹³⁾ Die Konfiguration der parabolischen s -Umgebungen dürfte zuerst für $s = \frac{\sqrt{3}}{2}$ bei

Humbert (J. d. m. p. e. a., (7) 2 (1916), pp. 84 ff.) auftreten. Für beliebige s wurde sie anscheinend zum ersten Mal von L. R. Ford (Proc. Ed. M. S., Bd. 35 (1916–17), pp. 40 ff.) angegeben, von beiden Autoren im Zusammenhang mit der Theorie der Kettenbrüche. Anschließend an eine kurze Mitteilung von A. Speiser (Verh. d. Schweiz. Nat. Ges. (1923), pp. 113–114) wurde sie ferner in der Zürcher Dissertation von J. Züllig (Ueber eine geometrische Deutung unendlicher Kettenbrüche und ihre Approximation durch rationale Zahlen, Zürich, Orell Füssli, 1928) behandelt. Die Invarianz dieser Konfiguration für $s = 1$ ist implizite in den Betrachtungen von R. Picard, (Untersuchung einer Untergruppe der unimodularen Picard'schen Gruppe, Basel, Dissertation, 1927) sowie in den älteren Untersuchungen von v. Dick (Leipz. Ber. (1883), pp. 61 ff.) und R. Fricke (Fricke-Klein, Automorphe Funktionen, Bd. 1 (1897), p. 431) enthalten. Verallgemeinerungen auf höhere Räume finden sich in den Arbeiten von L. R. Ford (Trans. A. M. S., Bd. 27 (1925), pp. 146 ff.) und A. Speiser (Journal f. r. u. a. M., Bd. 167 (1932), pp. 88 ff.).

$$\left(\frac{1}{2s} - y \frac{2r}{1-r^2}\right)^2 \geq x^2 + \left(\frac{1}{2s} - y \frac{1+r^2}{1-r^2}\right)^2,$$

$$x^2 \leq \left(\frac{1}{2s} - y \frac{2r}{1-r^2} + \frac{1}{2s} - y \frac{1+r^2}{1-r^2}\right) \left(y \frac{1+r^2}{1-r^2} - y \frac{2r}{1-r^2}\right)$$

$$x^2 \leq \left(\frac{1}{s} - y \frac{1+r}{1-r}\right) y \frac{1-r}{1+r}, \quad x^2 \leq \left(\frac{1}{s} - \mu\right) \sigma,$$

$$\frac{x^2}{\sigma} = \mu \frac{x^2}{y^2} \leq \frac{1}{s} - \mu, \quad \frac{1}{s} \geq \mu \left(\frac{x^2 + y^2}{y^2}\right) = \mu \frac{|\omega|^2}{(\Im \omega)^2} = \frac{\mu}{\Im \omega \Im \frac{-1}{\omega}}.$$

Daher ist

$$(31) \quad \mu^* = \frac{\mu}{\Im \omega \Im \frac{-1}{\omega}} = \mu \frac{x^2 + y^2}{y^2} \leq 2\mu, \quad \frac{1}{\mu^*} = \Im \frac{-1}{\omega} \frac{1-r}{1+r},$$

da ω in $\Sigma^{(0)}$ liegt und daher $\frac{x^2}{y^2} \leq 1$, $\frac{x^2 + y^2}{y^2} \leq 2$ ist. So ist μ^* für $\mu \leq \frac{1}{2}$ sicher nicht größer als 1, so daß insbesondere für $\mu \leq \frac{1}{2}$ das Innere von K ganz in der parabolischen 1-Umgebung $U_{0,1}$ von $z = 0$ liegt, d. h. im Kreis $\left|z - \frac{i}{2}\right| < \frac{1}{2}$.

Daraus folgt aber nun, daß wenn ein beliebiger Punkt des Kreises K durch eine Substitution S von Γ in H gebracht wird, er im Bereich $\Sigma^{(0)}$ liegen muß. (Die Alternative wäre, da ja seine Ordinate jedenfalls nicht größer als $\frac{1}{2}$ ist, daß er in die Nähe der Spitzen ± 1 (d. h. in den Bereich $\Sigma^{(1)}$) transformiert würde.)

Es genügt zu zeigen, daß S für $\omega \notin \Sigma^{(0)}$ den Nullpunkt und daher auch $U_{0,1}$ invariant läßt, da ja dann K in $U_{0,1}$ bleibt. Transformiert nun S den Nullpunkt in einen Punkt $\alpha \neq 0$, so kann α auch nicht ± 1 oder ∞ sein, da ja diese Punkte vermöge Γ nicht mit 0 äquivalent sind. Dann geht $U_{0,1}$ in $U_{\alpha,1}$ über, und $U_{\alpha,1}$ hat keinen Punkt mit $U_{0,1}$ noch mit $U_{\pm 1,1}$ gemeinsam, also auch nicht mit dem ganzen Gebiet H . Da aber K ganz in $U_{0,1}$ liegt, muß das Bild von K ganz in $U_{\alpha,1}$ liegen, hat also auch keine Punkte mit H gemeinsam, entgegen der Voraussetzung. Daraus folgt insbesondere, daß ein solcher Kreis K , ebenso wie $U_{0,1}$

nur solche Dreiecke der Modulfigur durchsetzen kann, die eine Spitze im o-Punkt haben.¹⁴⁾

Die gleiche Ueberlegung zeigt, daß wenn ein Kreis $K_{\omega, r}$ mit $\omega \prec \Sigma^{(0)}$ ganz in einer parabolischen s -Umgebung $U_{0, s}$ ($s \geq 1$) des Nullpunktes verläuft, jeder Transformierte eines Punktes von K vermöge einer Substitution von I' , der in H liegt, dann innerhalb $U_{0, s}$, also innerhalb des Durchschnitts von $U_{0, s}$ und H liegt.

§ 6. Abschätzung von $S(p_0, r)$ durch Ω_μ und ω_μ

Ist nun wieder $p(z)$ für $|z| < 1$ regulär und $\neq 0$, $\neq 1$, so daß (9) gilt, so ist für $|z| \leq r < 1$

$$(32) \quad |p(z)| \leq \max |\lambda(z)| \text{ für } z \prec K_{\omega, r},$$

wobei $K_{\omega, r}$ der Kreis mit dem Pseudomittelpunkt ω und dem Pseudoradius r ist. Dabei ist ω eine solche Zahl aus H , daß $p_0 = \lambda(\omega)$ ist. Bezeichnet man also den Zweig der Umkehrung von $\lambda(z)$, dessen Werte in H liegen, mit $\omega(\lambda)$, so ist $\omega = \omega(p_0)$. — (32) folgt daraus, daß wegen (IIa) die Werte von $C(z)$ für $|z| \leq r$ stets im $K_{\omega, r}$ liegen.

Wir bezeichnen nun mit B_μ für $\mu \geq \frac{1}{2}$ die in H enthaltene Strecke der Geraden $\Im z = \mu$. Ferner bezeichnen wir mit $B_\mu^{*(0)}$ für $\mu < 1$ den in $\Sigma^{(0)}$ enthaltenen Teilbogen des in o berührenden Kreises mit dem Durchmesser μ :

$$\left| z - \frac{i\mu}{2} \right| = \frac{\mu}{2},$$

d. h. des Begrenzungskreises von $U_{0, \frac{1}{\mu}}$. Diese Strecken und Bögen sind in der Figur eingezeichnet.

Wir bestimmen nun die Maxima und Minima von $|\lambda(z)|$ auf B_μ ($\mu \geq \frac{1}{2}$) und $B_\mu^{*(0)}$ ($\mu < 1$) und bezeichnen sie resp. mit

$$(33) \quad \Omega_\mu, \omega_\mu; \Omega_\mu^{*(0)}, \omega_\mu^{*(0)}.$$

¹⁴⁾ Genauer gilt folgendes: Der Kreis um den Punkt i mit dem Radius 1, also der Randkreis von $U_{0, \frac{1}{2}}$, durchsetzt sämtliche Dreiecke mit der Spitze im Nullpunkt und berührt jedesmal ihre dritten, nicht an o anstoßenden Halbkreise, so daß er der größte im Nullpunkt berührende Kreis ist, der nur in die Dreiecke mit der Spitze im Nullpunkt eindringt. Aus dieser Eigenschaft folgt übrigens unmittelbar, daß dieser Kreis bei jeder Modulsstitution, die den Nullpunkt in sich überführt, unverändert bleibt. Dies würde für unsere Zwecke bereits ausreichen, so daß bei Benutzung dieses Kreises anstatt $U_{0, 1}$ unsere Betrachtungen über parabolische Umgebungen unnötig wären. Doch liefert dieser Kreis eine relativ unsymmetrische Abgrenzung der Spitze o gegen die Spitzen ± 1 .

Wir bezeichnen ferner den Teil von H mit $\Im z \leq \mu$ ($\mu \geq \frac{1}{2}$) mit H_μ und den Durchschnitt von $\Sigma^{(0)}$ mit $U_0, \frac{1}{\mu}$ mit $\Sigma_\mu^{(0)}$ (vgl. Fig.).

Ist nun μ die größte Ordinate der Punkte von $K_{\omega, r}$, also nach § 4

$$\mu = \Im \omega \frac{1+r}{1-r} = \Im \omega(p_0) \frac{1+r}{1-r},$$

so gilt sicher:

$$(34) \quad |\lambda(z)| \leq \Omega_\mu \quad \text{für } z \prec H_\mu, \mu \geq \frac{1}{2};$$

$$(35) \quad |\lambda(z)| \leq \Omega_\mu^{*(0)} \quad \text{für } z \prec \Sigma_\mu^{(0)}, \mu < 1.$$

Denn da die Randteile von H_μ , die zum Rand von H gehören, durch $\lambda(z)$ auf ein Stück der positiven reellen Axe (doppelt überdeckt) abgebildet werden, wird auf ihnen der größte Wert von $|\lambda(z)|$ in einem der beiden Schnittpunkte mit der Geraden $\Im z = \mu$ erreicht. Und das Analoge gilt für $\Sigma_\mu^{(0)}$.

Liegt nun ferner ein Punkt von $K_{\omega, r}$ nicht in H , so wird er durch eine Substitution von Γ , die den Wert von $\lambda(z)$ unverändert läßt, in H übergeführt und liegt alsdann nach §§ 4 und 5 in H_μ für $\mu \geq \frac{1}{2}$. Falls ferner $\mu^* < 1$ ist, liegt ein solcher Punkt in $\Sigma_{\mu^*}^{*(0)}$, wenn ω in $\Sigma^{(0)}$ liegt, wo μ^* nach (31) gleich $\frac{\mu}{\Im \omega \Im \frac{-1}{\omega}} \leq 2\mu$ ist. Daher ist $|\lambda(z)|$ für alle

Punkte von $K_{\omega, r}$ resp. höchstens gleich $\left(\mu^* = \frac{\mu}{\Im \omega \Im \frac{-1}{\omega}} \right)$:

$$\Omega_\mu, \Omega_{\mu^*}^{*(0)}.$$

Andererseits hat $K_{\omega, r}$ sicher wenigstens einen Punkt mit B_μ bzw. $B_{\mu^*}^{*(0)}$ gemeinsam, und da nach § 2 die Werte der entsprechend gebildeten Funktion (12 a)

$$(36) \quad C^{**}(z) = \Re \omega + i \Im \omega \frac{1+z}{1-z}, \quad \omega = \omega(p_0)$$

für $|z| \leq r$ den ganzen Kreis $K_{\omega, r}$ erfüllen, ist $S(p_0, r)$ sicher wenigstens gleich dem entsprechenden Minimum ω_μ , bzw. $\omega_{\mu^*}^{*(0)}$. Daher ergibt sich

$$(37) \quad \omega_{\mu} \leq S(p_0, r) \leq \Omega_{\mu} \quad \text{für } \mu = \Im \omega(p_0) \frac{1+r}{1-r} \geq \frac{1}{2},$$

$$(38) \quad \omega_{\mu^*}^{*(0)} \leq S(p_0, r) \leq \Omega_{\mu^*}^{*(0)} \quad \text{für } \mu = \Im \omega(p_0) \frac{1+r}{1-r} < \frac{1}{2},$$

$$\mu^* = \frac{\mu}{\Im \omega(p_0) \Im \frac{1}{\omega(p_0)}}.$$

Um hieraus weitere Folgerungen zu ziehen, leiten wir zunächst asymptotische Abschätzungen von $|\lambda(z)|$ in der Nähe der Spitzen von H her.

§ 7. Asymptotische Abschätzungen von $J(z)$ und $\lambda(z)$

Wir setzen $q = e^{\pi i z}$ und gehen von der Formel (2) für $\omega = z$ aus, deren erste Glieder für $q \rightarrow 0$, d. h. $\Im z \uparrow \infty$ liefern:

$$(39) \quad 12^3 q^2 J(z) = (1 + 720 q^2 + \dots)(1 + 24 q^2 + \dots) = 1 + 744 q^2 + O(q^4).$$

Hieraus folgt

$$(40) \quad 1 - 744 |q^2| + O(|q|^4) \leq |12^3 q^2 J(z)| \leq 1 + 744 |q|^2 + O(|q|^4),$$

$$(41) \quad \left| |J(z)| - \frac{1}{12^3 |q|^2} \right| \leq \frac{31}{72} + O(q^2), \quad q \rightarrow 0,$$

$$(42) \quad \left| |J(z)| - \frac{e^{2\pi \Im z}}{12^3} \right| \leq \frac{31}{72} + O(e^{-2\pi \Im z}), \quad \Im z \uparrow \infty.$$

Andererseits setzen wir in der Formel (7)

$$2\lambda - 1 = \mu, \quad \mu^2 - 1 = 4\lambda(\lambda - 1) = \nu,$$

und erhalten

$$27 J(z) = \frac{(\nu + 4)^3}{\nu^2} = \nu + 12 + \frac{48}{\nu} + \frac{64}{\nu^2} = \nu + O(1),$$

da $\lambda \rightarrow \infty$ für $\Im z \uparrow \infty$ gilt. Daraus und aus (41) folgt weiter

$$\mu^2 = 27 J(z) + O(1) = \frac{1}{64 q^2} + O(1) = \left(\frac{1}{8 q} \right)^2 (1 + O(q^2)),$$

daher

$$\mu = \pm \frac{1}{8q} (1 + O(q^2)).$$

Da aber q für rein imaginäre z positiv ist und λ , daher auch μ Punkte der imaginären Axe auf Punkte der *negativen* reellen Axe abbildet, gilt hier das Minuszeichen und wir erhalten

$$2\lambda - 1 = \frac{-1}{8q} (1 + O(q^2)), \quad 2\lambda = \frac{-1}{8q} (1 - 8q + O(q^2)),$$

$$|16q\lambda| = |1 - 8q + O(q^2)|,$$

$$1 - 8|q| + O(q^2) \leq |16q\lambda| \leq 1 + 8|q| + O(q^2),$$

$$\left| |\lambda| - \frac{1}{16|q|} \right| \leq \frac{1}{2} + O(q),$$

$$(43) \quad \left| |\lambda(z)| - \frac{1}{16} e^{\pi \Im z} \right| \leq \frac{1}{2} + O(e^{-\pi \Im z}), \quad \Im z \uparrow \infty.$$

Lassen wir aber z gegen 0 innerhalb H konvergieren, so benutze man die Formel

$$\lambda\left(-\frac{1}{z}\right) = \frac{1}{\lambda(z)}.$$

Dann gilt nach (43)

$$(44) \quad \left| \frac{1}{|\lambda(z)|} - \frac{1}{16} e^{\pi \Im \frac{-1}{z}} \right| \leq \frac{1}{2} + O(e^{-\pi \Im \frac{-1}{z}}), \quad z \rightarrow 0.$$

Beachten wir aber, daß in H für $z = x + iy$ ($x \pm \frac{1}{2}$)² + $y^2 \geq \frac{1}{4}$ ist, d. h. $|x| \leq x^2 + y^2$ und für $z \rightarrow 0$ daher $x = O(y^2)$, so gilt

$$\Im \frac{-1}{z} = \frac{y}{x^2 + y^2} = \frac{1}{y} - \frac{x^2}{y^3} + \frac{x^4}{y^5} + \dots = \frac{1}{y} + O(y),$$

und daher $e^{\pi \Im \frac{-1}{z}} = e^{\frac{\pi}{y}} (1 + o(\Im z))$,

$$(45) \quad \frac{1}{|\lambda|} \sim \frac{e^{\frac{\pi}{\Im z}}}{16}, \quad z \rightarrow 0, \quad z \not\sim H,$$

$$-\lg |\lambda| = \frac{\pi}{\Im z} - \lg 16 + o(1), \quad \frac{\pi}{\Im z} = \lg 16 + \lg \frac{1}{|\lambda|} + o(1),$$

(45*)

$$\Im z = \frac{-\pi}{\lg |\lambda|} \left(1 + \frac{\lg 16}{\lg |\lambda|} + o\left(\frac{1}{\lg |\lambda|}\right) \right) = \frac{-\pi}{\lg |\lambda|} - \frac{\pi \lg 16}{(\lg |\lambda|)^2} + o\left(\frac{1}{\lg |\lambda|^2}\right).$$

§8. Abschätzungen von $S(p_0, r)$ durch $\omega(p_0)$ und p_0

Das für uns Wesentliche an den in § 7 erhaltenen Abschätzungen ist, daß sie in der dort angegebenen Näherung nur von $\Im z$ bzw. $\Im \frac{-1}{z}$ abhängen, sodaß wir $\mathcal{Q}$ und ω im Wesentlichen in gleicher Weise abschätzen, woraus, wegen (37), (38), die gleiche Abschätzung auch für $S(p_0, r)$ folgt.

Es sei zuerst $\mu \geq \frac{1}{2}$. Dann liefert (43) eine Abschätzung für alle Werte von $|\lambda(z)|$ auf B_μ , die also auch auf $S(p_0, r)$ anwendbar ist und ergibt

$$(I) \quad S(p_0, r) = \frac{1}{16} e^{\pi\mu} + \frac{\theta}{2} + O(e^{-\pi\mu}), \quad |\theta| \leq 1, \quad \mu = \Im \omega(p_0) \frac{1+r}{1-r} \geq \frac{1}{2},$$

wo O sich auf den Grenzübergang $\mu \uparrow \infty$ bezieht, und (I) im Übrigen gleichmäßig in p_0 und r gilt. Hieraus folgt weiter

$$(46) \quad S(p_0, r) = \frac{1}{16} e^{\pi\mu} (1 + 8\theta e^{-\pi\mu} + O(e^{-2\pi\mu})),$$

$$(II) \quad \lg(16 S(p_0, r)) = \pi\mu + 8\theta e^{-\pi\mu} + O(e^{-2\pi\mu}),$$

sodaß nicht nur $\lg(16 S(p_0, r)) - \pi\mu \rightarrow 0$, sondern sogar $e^{\pi\mu} (\lg(16 S(p_0, r)) - \pi\mu)$ beschränkt für $\frac{1}{2} \leq \mu \rightarrow \infty$ bleibt. Für die in diesen Abschätzungen vorkommende GröÙe $e^{\pi\mu} = (e^{\pi \Im \omega})^{\frac{1+r}{1-r}}$ erhalten wir für $p_0 \rightarrow \infty$ aus (43)

$$|p_0| = |\lambda(\omega(p_0))| = \frac{1}{16} e^{\pi \Im \omega} + \frac{\theta}{2} + O(e^{-\pi \Im \omega}),$$

$$\begin{aligned}
 (47) \quad e^{\pi \Im \omega} &= 16 |p_0| + 8\theta + O\left(\frac{1}{p_0}\right), \\
 (48) \quad e^{\pi \Im \omega} &= 16 |p_0| \left(1 + \frac{\theta}{2} \frac{1}{|p_0|} + O\left(\frac{1}{p_0^2}\right)\right), \\
 (49) \quad \pi \Im \omega &= \lg(16 |p_0|) + \frac{\theta}{2} \frac{1}{|p_0|} + O\left(\frac{1}{p_0^2}\right),
 \end{aligned}
 \left. \vphantom{\begin{aligned} (47) \\ (48) \\ (49) \end{aligned}} \right\} p_0 \rightarrow \infty, |\theta| \leq 1.$$

Konvergiert also insbesondere p_0 gegen ∞ , so ergibt sich aus (46) und (48)

$$S(p_0, r) = \frac{1}{16} \left| 16 p_0 \right|^{\frac{1+r}{1-r}} \left\{ 1 + \frac{\theta}{2 |p_0|} + O\left(\frac{1}{p_0^2}\right) \right\}^{\frac{1+r}{1-r}} \left(1 + O\left(p_0^{-\frac{1+r}{1-r}}\right) \right).$$

Hieraus folgt aber für *alle* p_0

$$(I^*) \quad 16 S(p_0, r) = (16 |p_0| + \theta c)^{\frac{1+r}{1-r}}, \quad |\theta| \leq 1, \quad c = \text{const.},$$

und insbesondere für $r \geq \frac{1}{3}$, $\frac{1+r}{1-r} \geq 2$ und $p_0 \rightarrow \infty$

$$(I^{**}) \quad 16 S(p_0, r) = \left(16 |p_0| + 8\theta + O\left(\frac{1}{p_0}\right) \right)^{\frac{1+r}{1-r}}, \quad |\theta| \leq 1, \quad 1 > r \geq \frac{1}{3}.$$

Durch Logarithmieren von (I*) und (I**) ergibt sich

$$(II^*) \quad \lg(16 S(p_0, r)) = \frac{1+r}{1-r} \left(\lg |16 p_0| + O\left(\frac{1}{p_0}\right) \right), \quad p_0 \rightarrow \infty,$$

$$\begin{aligned}
 (II^{**}) \quad \lg(16 S(p_0, r)) &= \frac{1+r}{1-r} \left(\lg |16 p_0| + \frac{\theta}{2 |p_0|} + O\left(\frac{1}{p_0^2}\right) \right), \\
 & \quad r \geq \frac{1}{3}, \quad |\theta| \leq 1, \quad p_0 \rightarrow \infty.
 \end{aligned}$$

Es sei nun $\mu < \frac{1}{2}$ oder allgemeiner $\mu^* < 1$. Konvergiert μ gegen 0, so muß dann offenbar $\Im \omega \rightarrow 0$ gelten, sodaß $\omega \rightarrow 0$ oder $\omega \rightarrow 1$ gilt. Es möge nun $\omega \rightarrow 0$ und daher $p_0 \rightarrow 0$ gelten. Dann liefert (38) recht scharfe Ergebnisse, wenngleich in ihnen auch der Realteil von $\omega(p_0)$

eine Rolle spielt. Denn die Werte von $|\lambda(z)|$ auf dem Kreisbogen $B_{\mu^*}^{*(0)}$ ($\mu^* < 1$) werden von $\frac{1}{|\lambda(z)|} = \left| \lambda\left(-\frac{1}{z}\right) \right|$ auf der Strecke $B_{\frac{1}{\mu^*}}$ angenommen. Dort gilt aber wegen (43)

$$|\lambda(z)| = \frac{1}{16} e^{\frac{\pi}{\mu^*}} + \frac{\theta}{2} + O\left(e^{-\frac{\pi}{\mu^*}}\right), \quad |\theta| \leq 1, \mu^* \rightarrow 0,$$

und daher

$$(50) \quad S(p_0, r) = 16 e^{-\frac{\pi}{\mu^*}} \left(1 + 8\theta e^{-\frac{\pi}{\mu^*}} + O\left(e^{-\frac{2\pi}{\mu^*}}\right) \right),$$

$$|\theta| \leq 1, \mu^* = \frac{1}{\Im \frac{1+r}{-1-r}} \rightarrow 0, p_0 \rightarrow 0.$$

$$(III) \quad S(p_0, r) = 16 e^{-\frac{\pi}{\mu^*}} + 128\theta e^{-\frac{2\pi}{\mu^*}} + O\left(e^{-\frac{3\pi}{\mu^*}}\right),$$

$$|\theta| \leq 1, \mu^* = \frac{1}{\Im \frac{1+r}{-1-r}} \rightarrow 0, p_0 \rightarrow 0.$$

Man überzeugt sich leicht, daß die Bedingung $\mu^* \rightarrow 0$ mit $\frac{\Im \omega}{1-r} \rightarrow 0$ äquivalent ist; dies ist aber, wegen (45*), äquivalent mit $|p_0|^{1-r} \rightarrow 0$.

Logarithmiert man nun (50), so ergibt sich

$$(IV) \quad \lg \frac{S(p_0, r)}{16} = -\frac{\pi}{\mu^*} + 8\theta e^{-\frac{\pi}{\mu^*}} + O\left(e^{-\frac{2\pi}{\mu^*}}\right),$$

$$|\theta| \leq 1, \mu^* = \frac{1}{\Im \frac{1+r}{-1-r}} \rightarrow 0, p_0 \rightarrow 0.$$

Der Zusammenhang mit p_0 ergibt sich jetzt aus (31):

$$e^{-\frac{\pi}{\mu^*}} = \left(e^{-\pi \Im \frac{-1}{\omega} \frac{1}{1+r}} \right), \quad p_0 = \lambda(\omega(p_0)) = \frac{1}{\lambda\left(\frac{-1}{\omega(p_0)}\right)},$$

wenn man beachtet, daß wegen (43) für $p_0 \rightarrow 0$

$$\begin{aligned} \frac{1}{p_0} &= \left| \lambda \left(\frac{-1}{\omega(p_0)} \right) \right| = \frac{1}{16} e^{\pi \Im \frac{-1}{\omega(p_0)}} + \frac{\theta}{2} + O \left(e^{-\pi \Im \frac{-1}{\omega(p_0)}} \right) = \\ &= \frac{1}{16} e^{\pi \Im \frac{-1}{\omega(p_0)}} \left(1 + 8 \theta e^{-\pi \Im \frac{-1}{\omega(p_0)}} + O \left(e^{-2\pi \Im \frac{-1}{\omega(p_0)}} \right) \right), \\ e^{-\pi \Im \frac{-1}{\omega(p_0)}} &= \frac{|p_0|}{16} \left(1 + \frac{\theta}{2} |p_0| + O(p_0^2) \right) \end{aligned}$$

gilt. Denn dann folgt aus (50)

$$\begin{aligned} S(p_0, r) &= 16 \left| \frac{p_0}{16} \right|^{\frac{1-r}{1+r}} \left\{ 1 + \frac{\theta}{2} |p_0| + O(p_0^2) \right\}^{\frac{1-r}{1+r}} \\ &\quad \left\{ 1 + 8 \theta_1 \left| \frac{p_0}{16} \right|^{\frac{1-r}{1+r}} + O \left((1-r) p_0 |p_0|^{\frac{1-r}{1+r}} \right) + O \left(|p_0|^{2 \frac{1-r}{1+r}} \right) \right\}, \\ (51) \quad \frac{S(p_0, r)}{16} &= \left| \frac{p_0}{16} \right|^{\frac{1-r}{1+r}} \left(1 + \frac{\theta}{2} \frac{1-r}{1+r} |p_0| + 8 \theta_1 \left| \frac{p_0}{16} \right|^{\frac{1-r}{1+r}} + O \left(|p_0|^{2 \frac{1-r}{1+r}} \right) \right). \end{aligned}$$

Hieraus folgt dann erstens für alle r

$$(III^*) \quad \frac{S(p_0, r)}{16} = \left| \frac{p_0}{16} \right|^{\frac{1-r}{1+r}} + O \left(\left| \frac{p_0}{16} \right|^{2 \frac{1-r}{1+r}} \right), \quad |p_0|^{1-r} \rightarrow 0.$$

Für $r \geq \frac{1}{3}$, $\frac{1-r}{1+r} \leq \frac{1}{2}$ aber gilt zweitens

$$(III^{**}) \quad \frac{S(p_0, r)}{16} = \left| \frac{p_0}{16} \right|^{\frac{1-r}{1+r}} + 8 \theta \left| \frac{p_0}{16} \right|^{2 \frac{1-r}{1+r}} + O \left(\left| \frac{p_0}{16} \right|^{3 \frac{1-r}{1+r}} \right),$$

wo $|p_0|^{1-r} \rightarrow 0$, $|\theta| \leq 1$ ist.

Durch Logarithmieren ergibt sich aus diesen Formeln aber

$$(IV^*) \quad \lg \frac{S(p_0, r)}{16} = \frac{1-r}{1+r} \lg \left| \frac{p_0}{16} \right| + O \left(|p_0|^{\frac{1-r}{1+r}} \right), \quad |p_0|^{1-r} \rightarrow 0,$$

$$(IV^{**}) \quad \lg \frac{S(p_0, r)}{16} = \frac{1-r}{1+r} \lg \left| \frac{p_0}{16} \right| + 8\theta \left| \frac{p_0}{16} \right|^{\frac{1-r}{1+r}} + O \left(|p_0|^{2\frac{1-r}{1+r}} \right),$$

$$r \geq \frac{1}{3}, \quad |p_0|^{1-r} \rightarrow 0, \quad |\theta| \leq 1.$$

Konvergiert für $\mu^* < 1$ aber p_0 gegen 1, so hat man nur die Formeln III — IV** auf $1 - p(z)$ anzuwenden und erhält Abschätzungen für $|1 - p(z)|$.

Ist im Falle der Formel (I) ($\mu \rightarrow \infty$) $p_0 \rightarrow \infty$ für festes r , oder allgemeiner für $|p_0|^{1-r} \rightarrow \infty$, so läßt sich diese Formel mit Hilfe von (III) durch die Angabe einer (asymptotisch genauen) *unteren Grenze* von $|p(z)|$ für $|z| \leq r$ ergänzen. Denn für $|p_0|^{1-r} \rightarrow \infty$ lassen sich auf $q(z) = \frac{1}{p(z)}$ wegen $|q(0)|^{1-r} \rightarrow 0$ (III), sowie (III*), (III**) anwenden und es ergibt sich für $|z| \leq r$ wegen $\mu^* = \frac{1}{\Im \omega(q(0))} \cdot \frac{1+r}{1-r} = \frac{1}{\Im \omega(p(0))} \cdot \frac{1+r}{1-r} = \frac{1}{\sigma}$,

$$\left| \frac{1}{p(z)} \right| \leq 16 e^{-\pi\sigma} + 128 \theta e^{-2\pi\sigma} + O(e^{-3\pi\sigma}), \quad |\theta| \leq 1,$$

$$\left| \frac{1}{16 p(z)} \right| \leq \left| \frac{1}{16 p_0} \right|^{\frac{1-r}{1+r}} + O \left(\left| \frac{1}{p_0} \right|^{2\frac{1-r}{1+r}} \right),$$

$$\left| \frac{1}{16 p(z)} \right| \leq \left| \frac{1}{16 p_0} \right|^{\frac{1-r}{1+r}} + 8\theta \left| \frac{1}{16 p_0} \right|^{2\frac{1-r}{1+r}} + O \left(\left| \frac{1}{p_0} \right|^{3\frac{1-r}{1+r}} \right), \quad r \geq \frac{1}{3},$$

und daraus folgt in Verbindung mit (I), (I*), (I**) mit $\mu = \Im \omega(p_0) \frac{1+r}{1-r} > \sigma = \Im \omega(p_0) \frac{1-r}{1+r} \rightarrow \infty$.

$$(V) \quad e^{\pi\sigma} + 8\theta + O(e^{-\pi\sigma}) \leq |16 p(z)| \leq e^{\pi\mu} + 8\theta_1 + O(e^{-\pi\mu}), \quad |\theta| \leq 1, \quad |\theta_1| \leq 1,$$

$$(V^*) \quad (16 |p_0| - c)^{\frac{1-r}{1+r}} \leq |16 p(z)| \leq (16 |p_0| + c)^{\frac{1+r}{1-r}}, \quad 0 \leq c < \infty, \quad |z| \leq r$$

$$(V^{**}) |16p_0|^{\frac{1-r}{1+r}} + 8\theta + O\left(\left|\frac{1}{p_0}\right|^{\frac{1-r}{1+r}}\right) \leq |16p(z)| \leq \left(16|p_0| + 8\theta_1 + O\left(\frac{1}{p_0}\right)\right)^{\frac{1+r}{1-r}},$$

$$|\theta| \leq 1, |\theta_1| \leq 1, r \geq \frac{1}{3}.$$

Andererseits folgen aus den linken Hälften dieser Formeln umgekehrt (III) (III*) (III**). Es ist noch von Interesse zu bemerken, daß die linke Hälfte der Formel (V*) (und damit auch (III*)) in etwas schwächerer Fassung aus (I*) direkt mit Hilfe eines Gedankens von Hrn. Valiron hergeleitet werden kann. Valiron bemerkt nämlich¹⁵⁾, daß eine lineare Abbildung des Einheitskreises in sich, die ein z mit $|z| = r$ in den Nullpunkt bringt, den Nullpunkt in einen Punkt z_1 überführt mit $|z_1| = r$. In einer Relation zwischen $f(0)$ und $f(z)$, in der ein z mit $|z| = r$ vorkommt, darf man daher z mit 0 vertauschen, sofern die Bedingungen für die Gültigkeit der Relation gegenüber einer linearen Transformation des Einheitskreises in sich invariant sind. Danach folgt aus (I*)

$$|16p(0)| \leq (16|p(z)| + c)^{\frac{1+r}{1-r}}, |c| < \infty,$$

$$(52) \quad 16|p(z)| \geq (16|p_0|)^{\frac{1-r}{1+r}} + O(1).$$

Dies ist aber nur insofern schwächer als die linke Hälfte von (V*), als die linke Seite von (V*) gleich $16|p_0|^{\frac{1-r}{1+r}} + O\left(\frac{1}{|p_0|^{\frac{2r}{1-r}}}\right)$ ist.

Ähnliches gilt auch für (I**). Denn aus $|p_0| \xrightarrow{1-r} \infty$ folgt ja, da mit (I**) auch (I*) gilt, nach (52) $p(z) \rightarrow \infty$, so daß in (I**) $p(0)$ mit $p(z)$ vertauscht werden kann und man eine Abschätzung von $p(z)$ nach unten erhält, die sogar etwas schärfer als die linke Hälfte von (V**) ist.

Durch ganz analoge Betrachtungen oder durch Benutzung von (V), (V*), (V**) für $\frac{-1}{p(z)}$ ergänzt man auch die Formeln (III), (III*), (III**) (wenn $\sigma^* = \frac{1}{\mathfrak{J} \frac{-1}{\omega(p_0)} \frac{1-r}{1+r}}$ gesetzt wird) zu

¹⁵⁾ Vgl. die a. p. 55 zitierte Abhandlung.

$$\text{VI} \quad e^{-\frac{\pi}{\sigma^*}} + 8\theta e^{-\frac{2\pi}{\sigma^*}} + O\left(e^{-\frac{3\pi}{\sigma^*}}\right) \leq \left| \frac{p(z)}{16} \right|_{|z|=r} \leq e^{-\frac{\pi}{\mu^*}} + 8\theta_1 e^{-\frac{2\pi}{\mu^*}} + O\left(e^{-\frac{3\pi}{\mu^*}}\right),$$

$$\text{VI}^* \quad \left(\frac{|p_0|}{16} + O(p_0^3) \right)^{\frac{1+r}{1-r}} \leq \left| \frac{p(z)}{16} \right| \leq \left(\frac{|p_0|}{16} + O(p_0^3) \right)^{\frac{1-r}{1+r}},$$

$$\begin{aligned} \text{VI}^{**} \quad \left(\frac{|p_0|}{16} + \frac{\theta_1}{32} |p|^2 + O(p_0^3) \right)^{\frac{1+r}{1-r}} &\leq \left| \frac{p(z)}{16} \right| \leq \left| \frac{p_0}{16} \right|^{\frac{1-r}{1+r}} + \\ &+ 8\theta \left| \frac{p_0}{16} \right|^{2\frac{1-r}{1+r}} + O(|p_0|^{3\frac{1-r}{1+r}}), \quad r \geq \frac{1}{3}. \end{aligned}$$

Diese Formeln gelten unter der Voraussetzung $\sigma^* \rightarrow 0$, d. h., wegen (45*), $|p_0|^{1-r} \rightarrow 0$. Dabei ist $|\theta| \leq 1$, $|\theta_1| \leq 1$.

(Eingegangen den 14. März 1932)

Die Existenz der Zahlenreihe und des Kontinuums

Von P. FINSLER, Zürich

1. Vorbemerkungen

Es soll die widerspruchsfreie Existenz der natürlichen Zahlenreihe und des Kontinuums in absolutem Sinne dargetan werden.

Der Beweis stützt sich auf eine frühere Arbeit¹⁾, in der ein widerspruchsfreies Axiomensystem für die Mengenlehre aufgestellt wird und einige Sätze daraus abgeleitet werden. Diese Arbeit hat gelegentlich Ablehnung erfahren²⁾, jedoch sind mir stichhaltige Einwendungen wesentlicher Art nicht bekannt geworden, und ich halte solche auch für ausgeschlossen. Um meinen Standpunkt klarzulegen, mögen hier folgende kurze Bemerkungen genügen:

1. Wer die Paradoxien und Antinomien der Logik und Mengenlehre nicht lösen kann (oder sie gar für unlösbar hält), der kann auch keine Kritik üben, denn mit einer Antinomie kann man alles beweisen, also auch alles widerlegen.

2. Wer die Antinomien in richtiger Weise lösen kann, der weiß, daß die reine (absolute) Logik einen sicheren Grund darstellt, auf dem man aufbauen kann. Ein System von Formeln als „schärfer“ zu betrachten, ist ein Irrtum; Formeln allein genügen nicht, um die Antinomien auszuschließen³⁾, dies kann nur der Gedanke tun, der darüber steht und der sich auf die Logik stützt.

3. Die endliche, aber nicht beschränkte Induktion als gegeben anzunehmen, wäre eine *petitio principii*; nur Finites zuzulassen, eine Einschränkung. Die Mathematik ist mehr als ein Handwerk oder ein Schachspiel. Auch transfinite Widersprüche müssen ausgeschlossen werden.

¹⁾ P. Finsler. Ueber die Grundlegung der Mengenlehre. Erster Teil. Die Mengen und ihre Axiome. Math. Zeitschrift, Bd. 25, 1926, S. 683 („Grundlegung“).

²⁾ So z. B. bei R. Baer, Ueber ein Vollständigkeitsaxiom in der Mengenlehre. Math. Zeitschrift, Bd. 27, 1928, S. 536.

³⁾ Vgl. z. B. G. Frege, Grundgesetze der Arithmetik. I und II, Jena 1893 und 1903, insbes. Nachwort.

2. Die Notwendigkeit eines Beweises

Die natürlichen Zahlen werden gerne als etwas unmittelbar Gegebenes betrachtet. Wenn es sich dabei nur um einen vorläufigen Standpunkt handelt, ist dies sicher berechtigt. Wenn man aber in kritischer Weise nichts Unbewiesenes zulassen will, so gilt es vielleicht doch noch für sehr kleine Zahlen, soweit sie sich vollständig im einzelnen überblicken lassen. Es gilt jedoch nicht mehr für beliebig große Zahlen und insbesondere nicht für die Zahlenreihe als Ganzes.

Dies könnte man schon aus der Tatsache entnehmen, daß es viele Mathematiker gab (oder noch gibt), welche die Existenz der Zahlenreihe als eines fertigen Systems mit unendlich vielen Elementen durchaus ablehnen. Man wird aber noch die Gründe dafür untersuchen müssen, denn man wird auch nichts ohne Beweis ablehnen wollen. Auch wird man sich, wenn man einen Beweis führen will, darüber klar sein müssen, was eigentlich bewiesen werden muß, wo also die Schwierigkeiten liegen.

Die Worte „natürliche Zahl“ oder „Zahlenreihe“ sind ohne Inhalt, wenn nicht gesagt wird, was darunter zu verstehen ist. Wenn wir über irgendwelche Dinge eine Aussage machen oder etwas beweisen wollen, so müssen wir genau sagen oder definieren, welche Dinge gemeint sind, und alle Eigenschaften, die wir ableiten, müssen sich aus der Definition ergeben. Für die natürlichen Zahlen kommen hier insbesondere zwei Definitionen in Betracht, die genetische und die axiomatische, die wir beide betrachten wollen.

a) Bei der genetischen Definition geht man von einer bestimmt gegebenen Anfangszahl aus, etwa von der Zahl 1. (Ob mit 0 oder 1 begonnen wird, ist an sich gleichgültig.) Dann nimmt man noch eine Operation $+ 1$ hinzu, die zu jeder schon gefundenen Zahl eine neue, von allen bisherigen verschiedene Zahl liefert. So erhält man die Zahlen $1 + 1$, $1 + 1 + 1$, $1 + 1 + 1 + 1$ usw., die auch mit 2, 3, 4 usw. bezeichnet werden.

Kann man nun von der Gesamtheit aller dieser Zahlen reden?

Wenn man annimmt, es liege schon ein fest gegebenes System von Dingen vor, aus dem die einzelnen Zahlen der Reihe nach entnommen werden können, dann könnte man allerdings die Gesamtheit aller dieser Zahlen definieren: es sind alle die und nur die Dinge des gegebenen Systems, die, von der Zahl (dem Ding) 1 ausgehend, durch beliebige, aber alleinige Anwendung der Operation $+ 1$ erreicht werden können.

Wenn man aber ein solches System von Dingen nicht als gegeben voraussetzen kann, dann ist dieser Schluß nicht zulässig. Die Reihe der Zahlen ist dann eine „werdende“, denn jede Zahl wird erst auf Grund aller vorhergehenden Zahlen sozusagen neu geschaffen, und man sieht zunächst nicht ein, ob dieser Prozeß schließlich zu einem Ende führt oder nicht.

Man könnte vielleicht denken: man sieht ein, daß dieser Prozeß zu keinem Ende führt. Dann übersieht man aber einen wichtigen Punkt: es ist nämlich noch gar nicht gesagt, daß sich der Prozeß in jedem einzelnen Fall wirklich durchführen, die Reihe also immer weiterführen läßt. Es ist noch gar nicht bewiesen, daß wirklich zu jeder Zahl eine folgende existiert.

Die Existenz einer nächstfolgenden Zahl scheint durch die Definition der Zahlen selbst gefordert zu werden. Aber eine Definition kann noch nicht die Existenz eines Dinges sicherstellen. Und darf man die Existenz eines Dinges fordern, das vielleicht einen logischen Widerspruch in sich trägt und deshalb gar nicht existieren kann? Das ist genau der Weg zu den Antinomien. Etwas logisch Widerspruchsvolles, also nicht Existierendes, können wir auch nicht durch einen Willkürakt erschaffen.

Welche Gründe veranlassen uns zu dem Glauben, daß es zu jeder Zahl eine nächstfolgende geben müsse?

Man könnte sich auf die Erfahrung berufen: zu jeder wirklich gegebenen Zahl können wir eine größere angeben. Aber die wirklich gegebenen Zahlen bilden einen so verschwindend winzigen Teil der Zahlen überhaupt, daß dies nicht als Beweis gelten kann.

Man könnte weiter sagen: Es ist kein Grund vorhanden, der gegen die Existenz einer nächstfolgenden Zahl spricht, also kann die Annahme dieser Existenz nicht zu einem Widerspruch führen.

Darauf ist zu erwidern, daß eben doch ein solcher Grund vorhanden ist. Er besteht darin, daß die genetische Definition der natürlichen Zahlen von zirkelhafter Natur ist, und zirkelhafte Definitionen können unerfüllbar sein.

Wenn man nämlich die Definition auf eine schärfere Form zu bringen sucht, so tritt der Zirkel hervor. Die genetische Definition der Zahlenreihe ist im wesentlichen eine Konstruktionsvorschrift, der man etwa die folgende Form geben kann:

Konstruktion $\mathfrak{U}$: Man beginne mit der Zahl 1 und setze hinter jede Zahl, die sich aus dieser Konstruktion $\mathfrak{U}$ ergibt, eine neue Zahl.

Bei der zuerst angegebenen Definition der einzelnen Zahlen hat man in ganz entsprechender Weise hinter jede durch ebendieselbe Vorschrift gefundene Zahl ein neues Zeichen $+ 1$ zu setzen.

Diese Vorschrift bezw. Konstruktion bezieht sich also ganz ausdrücklich auf sich selbst, und das ist ein offener Zirkel. Daß aber dieser Zirkel nicht ungefährlich ist, erkennt man aus der anderen, aber in ganz analoger Weise gebildeten Konstruktionsvorschrift:

Konstruktion $\mathfrak{B}$: Man beginne mit der Zahl 1 und setze hinter jede Zahlenreihe, die sich aus dieser Konstruktion $\mathfrak{B}$ ergibt, eine neue Zahl.

Diese Konstruktion $\mathfrak{B}$ ist sicher nicht ⁴⁾ in jedem Falle erfüllbar, man käme sonst zu der Antinomie von *Burali-Forti*. Es ist daher auch nicht selbstverständlich, daß die Konstruktion $\mathfrak{A}$ in jedem Falle erfüllbar ist. *Der Satz, daß es zu jeder Zahl eine folgende gibt* ⁵⁾, bedarf eines Beweises.

b) Bei der axiomatischen Definition bilden die natürlichen Zahlen ein System von Dingen, die einem gegebenen Axiomensystem genügen. Diesem Axiomensystem kann man (nach *Peano*) etwa die folgende Gestalt geben:

- I. 1 ist eine Zahl.
- II. Wenn n eine Zahl ist, so ist auch $n + 1$ eine Zahl.
- III. Sind m und n Zahlen und ist $m + 1 = n + 1$, so ist $m = n$.
- VI. Für jede Zahl n ist $n + 1 \neq 1$.
- V. Eine Aussage, die für die Zahl 1 gilt, und die für die Zahl $n + 1$ gilt, sofern sie für n gilt, gilt für jede Zahl n .

Ein solches Axiomensystem hat den Vorteil, daß es die Forderungen, die man an den Begriff der natürlichen Zahl stellt und die zu seiner Definition notwendig sind, einzeln und vollständig angibt, so daß man ein Fundament besitzt, auf dem sich die Lehre von den natürlichen Zahlen aufbauen läßt. Es bleibt aber die Frage, ob dieses Fundament selbst gesichert ist, d. h. ob das Axiomensystem keinen Widerspruch in sich enthält.

Um diese Frage zu beantworten, könnte man zu zeigen versuchen, daß es nicht möglich ist, von dem Axiomensystem ausgehend „in endlich vielen

⁴⁾ Dabei wird allerdings vorausgesetzt, daß man jede Reihe von Zahlen (z. B. die Reihe aller natürlichen Zahlen) auch als Ganzes auffassen kann. Dies wird für die natürlichen Zahlen unter $\mathfrak{B}$ gezeigt und läßt sich auf Grund des in der „Grundlegung“ definierten Systems aller Mengen auch allgemein zeigen.

⁵⁾ Der Satz dürfte hinfällig werden, wenn man verlangt, daß sich jede Zahl durch ein materielles Zeichen darstellen läßt. Sobald alle Atome der Welt für die Zeichen verbraucht sind, gibt es kein neues mehr.

Schritten“ einen Widerspruch herzuleiten. Wenn ein solcher Nachweis gelingt, so ist das wohl von Wichtigkeit, aber noch nicht genügend.

Wenn man nämlich annimmt, es gäbe nicht zu jeder Zahl eine folgende, dann gibt es auch nicht zu jedem Schritt einen folgenden, und die beschränkt vielen Schritte brauchen keinen Widerspruch zu ergeben. Es ist also wieder nicht gezeigt, daß es zu jeder Zahl eine folgende gibt, obschon das Gegenteil dieses Satzes mit dem Axiomensystem in offenkundigem Widerspruch steht. Dieser Widerspruch ließe sich aber nicht formalisieren und wäre deshalb für eine formale Methode nicht angreifbar.

Die Widerspruchsfreiheit des Axiomensystems läßt sich aber in einem absoluten Sinne beweisen, wenn es gelingt, ein widerspruchsfrei existierendes System von Dingen anzugeben, für welches sämtliche Axiome erfüllt sind. Daß man ein solches System nicht aus der genetischen Definition der Zahlenreihe allein entnehmen kann, wurde oben gezeigt. Es gelingt aber, ein solches System aus der Mengenlehre zu entnehmen.

Ebenso, wie sich die Widerspruchsfreiheit der Geometrie aus der Arithmetik beweisen läßt, läßt sich auch die Widerspruchsfreiheit der Arithmetik aus der Mengenlehre beweisen. Für die Mengenlehre ist ein anderer Weg notwendig; hier kann man sich direkt auf die Logik stützen, wie in der „Grundlegung“ gezeigt wurde.

Die widerspruchsfreie Existenz des Kontinuums folgt aus der der Zahlenreihe nur dann, wenn man weiß, daß die Operationen der Mengenlehre, insbesondere die beliebige Teilmengenbildung, auf die Reihe der natürlichen Zahlen angewendet werden dürfen.

3. Der Beweis

Als „Mengen“ bezeichnen wir ideelle Dinge, die durch eine bestimmte Beziehung („als Element enthalten“) miteinander verknüpft sind. Zusammenfassungen von Mengen heißen Systeme. Nicht jedem System braucht eine Menge zu entsprechen. Das durch 3 Axiome festgelegte System aller Mengen werde mit Σ bezeichnet. Wegen des Begriffs „zirkelfrei“ muß auf die „Grundlegung“ selbst verwiesen werden; ebenso beziehen sich die angeführten Sätze auf diese Arbeit.

Wir betrachten Systeme S , welche, wie z. B. das System Σ , die Nullmenge und mit jeder Menge M stets auch die Menge $\{M\}$ enthalten, sofern diese Menge, die M als einziges Element enthalten soll, existiert. Der Durchschnitt aller dieser Systeme, der also alle und nur die Mengen enthält, die in jedem System S vorkommen, werde mit $\mathfrak{J}$ bezeichnet. $\mathfrak{J}$ ent-

hält ebenfalls die Nullmenge und mit jeder Menge M auch die Menge $\{M\}$, sofern diese Menge existiert. Außerdem gilt aber für das System $\mathfrak{J}$ das Induktionsprinzip:

Eine Aussage $\mathfrak{A}$, die für die Nullmenge gilt und die für die Menge $\{M\}$ gilt, sofern sie für M gilt und $\{M\}$ existiert, gilt für jede Menge von $\mathfrak{J}$.

Zum Beweis betrachten wir das System aller der Mengen, für die die Aussage $\mathfrak{A}$ gilt. Dieses System enthält die Nullmenge und mit jeder Menge M auch die Menge $\{M\}$, sofern $\{M\}$ existiert, und gehört daher zu den Systemen $\mathfrak{S}$. Nach der Definition von $\mathfrak{J}$ muß also jede Menge von $\mathfrak{J}$ diesem System angehören, d. h. es gilt für sie die Aussage $\mathfrak{A}$.

Es folgt nun weiter, daß die Mengen von $\mathfrak{J}$ sämtlich zirkelfrei sind. Nach Satz 14 der „Grundlegung“ gilt dies nämlich für die Nullmenge, und nach Satz 15 existiert zu jeder zirkelfreien Menge M stets auch die Menge $\{M\}$ und sie ist ebenfalls zirkelfrei. Nach dem Induktionsprinzip ergibt sich also noch, daß $\mathfrak{J}$ tatsächlich mit jeder Menge M stets auch eine Menge $\{M\}$ enthält.

Wenn die Menge $\{M\}$ mit der Menge $\{N\}$ identisch ist, so ist auch M mit N identisch nach dem Axiom der Beziehung. Ferner ist jede Menge $\{M\}$ von der Nullmenge verschieden, da diese im Gegensatz zu $\{M\}$ kein Element enthält.

Damit sind aber für das System $\mathfrak{J}$ die Axiome *Peanos* sämtlich als erfüllt nachgewiesen, wenn man für die Zahl 1 die Nullmenge und für die Zahl $n + 1$ die Menge $\{M\}$ einsetzt, sofern n die Menge M bedeutet.

Bei der Definition von $\mathfrak{J}$ wurde der Begriff zirkelfrei nicht benutzt, das System $\mathfrak{J}$ ist also von diesem Begriff unabhängig. Nach Satz 12 existiert daher eine zirkelfreie Menge Z , welche gerade die Mengen von $\mathfrak{J}$ als Elemente enthält ⁶⁾.

Nach Satz 17 gibt es eine zirkelfreie Menge, welche die sämtlichen Teilmengen von Z und nur diese enthält. Diese Menge ist aber bekanntlich dem Kontinuum äquivalent.

4. Schlussbemerkung

Man könnte fragen, ob sich die Existenz der Zahlenreihe und des Kontinuums nicht auch nachweisen läßt, ohne daß man den Begriff zirkelfrei zu Hilfe nimmt.

⁶⁾ Damit ist auch das 7. Axiom *E. Zermelos* (Math. Ann. Bd. 65, 1908, S. 261), das in der „Grundlegung“ noch fehlte, für die zirkelfreien Mengen als erfüllt nachgewiesen und somit die Widerspruchsfreiheit dieses Axiomensystems der Mengenlehre dargetan.

Für die Zahlenreihe könnte dies möglich sein, da man, allerdings nicht in einfacher Weise (und vielleicht auch nicht ohne eine Zirkeldefinition), **einsehen kann**, daß es zu jeder Menge M eine Menge $\{M\}$ gibt, also auch dann, wenn M nicht zirkelfrei ist.

Für den Nachweis jedoch, daß das System aller Mengen eine dem Kontinuum äquivalente Menge enthält, dürfte der Begriff zirkelfrei oder ein äquivalenter Begriff jedenfalls nicht zu vermeiden sein.

Man könnte aber, wenn die Zahlenreihe gegeben ist, für das Kontinuum die Gesamtheit aller *Teilsysteme* der Zahlenreihe nehmen, dies wäre jedoch ein System höherer Stufe. Für die Bildung von Funktionen usw. kämen dann Systeme noch höherer Stufe in Frage. Dies wäre unbequem und könnte auch zu Schwierigkeiten führen. Die allgemeine Mengenlehre könnte man auf diese Weise jedenfalls nicht erhalten.

Der Vorteil der angewendeten Methode ist gerade der, daß die erste Stufe (die Mengen) schon die ganze Mengenlehre umfaßt; die zweite Stufe (Systeme von Mengen) und die dritte (alle Systeme einer Eigenschaft) braucht man nur für die Begründung, so daß man sich also z. B. in der Analysis, nachdem sie einmal begründet ist, um verschiedene Stufen nicht mehr zu kümmern braucht.

(Eingegangen den 1. April 1932)

Ein Satz über die konforme Abbildung Riemannscher Flächen

Von ROLF NEVANLINNA, Helsingfors

§ 1. Einleitung

1. Herr *Speiser*¹⁾ hat in zwei interessanten Abhandlungen die konforme Abbildung einer speziellen Klasse von Riemannschen Flächen untersucht. Auf folgenden Seiten soll ein Beitrag zu diesem Speiserschen Problem gegeben werden. Ich benutze hierbei, außer dem *Koebeschen* Verzerrungssatz, eine Methode, welche von Herrn *Ahlfors*²⁾ herrührt und von ihm mit großem Erfolg zur Untersuchung verschiedener Probleme aus der Theorie der konformen Abbildung angewendet worden ist.

2. Die von Herrn *Speiser* betrachteten Flächen können als diejenigen unendlich vielblättrigen Ueberlagerungsflächen F der in drei gegebenen Punkten w_1, w_2, w_3 punktierten w -Ebene definiert werden, welche an sämtlichen Stellen $w \neq w_1, w_2, w_3$ unverzweigt sind und außerdem folgenden speziellen Bedingungen genügen:

a) F ist einfach zusammenhängend.

b) Ueber den drei kritischen Stellen w_1, w_2, w_3 liegt, außer eventuellen schlichten Blättern, eine endliche oder unendliche Anzahl von Windungspunkten *unendlich hoher Ordnung*.

Um die Struktur einer solchen Fläche F zu beschreiben, denke man sie auf einen schlichten Kreis $|z| < R$ topologisch abgebildet. Es besteht dann eine stetige und umkehrbar eindeutige Zuordnung zwischen den Punkten z des Kreisinneren und den inneren, d. h. den von den Windungspunkten verschiedenen Stellen der Riemannschen Fläche F . Die Windungspunkte sind Randpunkte der Fläche F ; wenn der Punkt w gegen einen solchen Punkt rückt, so strebt der Bildpunkt $z = z(w)$ gegen den Rand $|z| = R$.

1) *A. Speiser*: Probleme aus dem Gebiet der ganzen transzendenten Funktionen und Ueber Riemannsche Flächen (Diese Zeitschrift, Bd. 1 und 2, 1929, 1930). Ich bin Herrn *Speiser* auch für viele anregende briefliche Mitteilungen, welche meine Arbeit auf dem Gebiete der Theorie der Riemannschen Flächen wesentlich befördert haben, zu großem Dank verpflichtet.

2) *L. Ahlfors*: Untersuchungen zur Theorie der konformen Abbildung und der ganzen Funktionen (Acta Soc. Scient. Fennicae, Nova Series, T. 1, No. 9, 1930), Zur Bestimmung des Typus einer Riemannschen Fläche (Diese Zeitschrift, Bd. 3, 1931), sowie eine Arbeit über die konforme Abbildung Riemannscher Flächen, welche in den Acta Mathematica erscheinen wird.

Die drei Grundpunkte können ohne Einschränkung der Allgemeinheit in $w_\nu = e^{\frac{2\pi\nu i}{3}}$ ($\nu = 1, 2, 3$) gebracht werden. Wir ziehen durch diese Punkte den Einheitskreis und betrachten die unendlich vielen schlichten Kreisscheiben $K_0: |w| < 1$ und $K_\infty: |w| > 1$, welche hierdurch aus der Fläche F ausgestanzt werden. Da jedes Blatt $K_0 + K_\infty$ mittels mindestens zweier Windungspunkte mit der Fläche zusammenhängen muß, so sind entweder alle drei oder gewisse zwei von den drei Punkten w_ν , welche auf der Peripherie einer gegebenen der Kreisscheiben K_0 oder K_∞ belegen sind, Windungspunkte (und somit Randpunkte) der Fläche F . In jenem Fall entspricht vermöge der Abbildung $z = z(w)$ der betreffenden Kreisscheibe ein Teilgebiet k_0 bzw. k_∞ des Kreises $|z| < R$, welches von drei Querschnitten des Kreises, den Bildkurven l_{12} , l_{23} , l_{31} der Kreisbogen $w_1 w_2, w_2 w_3, w_3 w_1$, begrenzt wird; ein solches Teilgebiet heie ein *Dreieck*. Wenn dagegen nur zwei von den Peripheriepunkten der betrachteten Kreisscheibe, z. B. w_1 und w_2 , Windungspunkte der Fläche darstellen, so entspricht dieser Kreisscheibe ein *Zweieck* k_0 bzw. k_∞ , welches von zwei Querschnitten l_{12} und l_{231} des Kreises $|z| < R$ begrenzt wird. Auf letzterem Bogen liegt der Bildpunkt des Peripheriepunktes w_3 , der jetzt ein innerer Punkt der Fläche ist.

Die unendlich vielen Zweiecke und Dreiecke k füllen den Kreis $|z| < R$ aus und häufen sich gegen den Rand $|z| = R$. Wenn G_0 ein beliebiges der Gebiete k , z. B. ein Gebiet k_0 ist, so lät sich die Fläche F von diesem Grundgebiet ausgehend durch kranzförmige Erweiterung ausschöpfen. Zu G_0 (die nullte Generation) werden die unmittelbar angrenzenden (zwei oder drei) Gebiete k_∞ (die erste Generation G_1) hinzugefügt, zum Gebiet $G_0 + G_1$ die von den angrenzenden Gebieten k_0 gebildete zweite Generation G_2 , u. s. w.

Der Aufbau der Fläche F lät sich nach Herrn *Speiser* noch einfacher durch einen „topologischen Baum“ darstellen. Zu dieser Darstellung gelangt man, wenn man in der oben erklärten Figur der z -Ebene in jedem Gebiet k_0 bzw. k_∞ denjenigen Punkt z_0 bzw. z_∞ einzeichnet, welcher dem Punkt $w = 0$ bzw. $w = \infty$ entspricht, und den Punkt z_0 des Grundgebietes G_0 mit den (zwei oder drei) Punkten z_∞ der ersten Generation G_1 verbindet durch Kurvenbogen, die unter den Bildern der Halbstrahlen $\arg w = \frac{\nu\pi}{3}$ ($\nu = 1, \dots, 6$) gewählt werden können. Setzt man diese Kurvenbogen in entsprechender Weise, ausgehend von den Endpunkten z_∞ , bis zu den Punkten $z_0, z_\infty, \dots$ der zweiten, dritten, ... Generation fort, so entsteht ein baumartiges Liniensystem B , welches

die unendlich vielen „Knotenpunkte“ z_0, z_∞ verbindet. Auf jedem, von zwei aufeinander folgenden Verzweigungsknoten begrenzten Astsegment liegt stets eine gerade Anzahl (≥ 0) von Knotenpunkten z_0, z_∞ . Umgekehrt entspricht jeder derartigen Baumfigur eine wohlbestimmte Riemannsche Fläche F der hier betrachteten Art³⁾.

3. Dem Hauptsatz der Theorie der konformen Abbildung gemäß läßt sich die Fläche F nicht nur topologisch, sondern sogar konform auf einen Kreis $|z| < R$ abbilden, dessen Radius R entweder endlich (hyperbolischer Fall) oder unendlich (parabolischer Fall) ist. Beim *Speiserschen* Problem handelt es sich nun um die Frage, welcher von diesen beiden möglichen Fällen für eine gegebene, durch das obige Verfahren dargestellte Fläche F vorliegt. Es seien hier zwei Fälle erwähnt, wo das Resultat bekannt ist.

a) Die Fläche sei so schwach verzweigt, daß die Figur B nur eine endliche Anzahl p von Verzweigungsstellen enthält. Dann besitzt die Fläche genau $p + 2$ Windungspunkte und ebenso viele „logarithmische Enden“, denen die $p + 2$ offenen Äste der Figur B zugeordnet sind. Eine derartige Fläche F gehört zum parabolischen Typus und läßt sich mithin auf die punktierte Ebene $z \neq \infty$ konform abbilden. Die Umkehrfunktion $w(z)$ der Abbildungsfunktion ist eine meromorphe Funktion der Ordnung $\frac{p}{2} + 1$, welche eine Differentialgleichung

$$\{w, z\} \equiv \frac{w'''}{w'} - \frac{3}{2} \left(\frac{w''}{w'} \right)^2 = P(z)$$

befriedigt, wo $P(z)$ ein gewisses Polynom p -ten Grades ist⁴⁾.

b) Wenn F möglichst stark verzweigt ist, d. h. wenn die Figur B sich an jedem Knotenpunkt z_0, z_∞ verzweigt, so handelt es sich um die universelle Ueberlagerungsfläche der in $w = w_1, w_2, w_3$ punktierten Ebene. Ueber jedem Grundpunkt w_v liegt eine unendliche Anzahl von Windungspunkten. Die Fläche wird durch die Umkehrfunktion der Modulfunktion auf einen endlichen Kreis konform abgebildet, und es liegt mithin der hyperbolische Fall vor.

4. Um die Verzweigtheit einer Riemannschen Fläche der betrachteten Klasse zu charakterisieren, führe man die Anzahl $\sigma(n)$ der den

³⁾ Man vergleiche hierzu meine Untersuchung Ueber Riemannsche Flächen mit endlich vielen Windungspunkten, welche in den Acta Mathematica erscheinen wird.

⁴⁾ Vgl. meine oben zitierte Arbeit.

Generationen $G_0, G_1, \dots, G_{n-1}$ angehörenden Verzweigungspunkte z_0, z_∞ ein. Die Anzahl der Knotenpunkte der n -ten Generation G_n berechnet sich dann auf $2 + \sigma(n)$. Für die unter a) genannten Flächen ist $\sigma(n)$ endlich, für die Modulfunktion wiederum ist $\sigma(n) = 3 \cdot 2^{n-1} - 2$.

Das in Aussicht gestellte Kriterium können wir jetzt folgendermaßen formulieren:

Satz. *Wenn die Reihe*

$$\sum_1^\infty \frac{1}{n \sigma(n)}$$

divergent ist, so gehört die Riemannsche Fläche dem parabolischen Typus an.

§ 2. Beweis des Satzes

5. Wir denken uns die Riemannsche Fläche F mittels der Funktion $z = z(w)$ auf den endlichen oder unendlichen Kreis $|z| < R$ konform abgebildet, und zeichnen in oben dargestellter Weise in diesem Kreis die Querschnitte l sowie das Liniensystem B , welches die Knotenpunkte z_0, z_∞ verbindet, ein. Wir fixieren dann einen beliebigen, außerhalb der Linien B belegenen Punkt z des Kreises und lassen ihn sich stetig bewegen, ohne die erwähnten Linien oder die Kreisperipherie $|z| = R$ zu überschreiten. Er beschreibt dann ein, von einem zur Figur B gehörenden Linienzug Γ sowie von gewissen Punkten des Kreises $|z| = R$ begrenztes einfach zusammenhängendes Gebiet H . Wir numerieren die auf der Kurve Γ liegenden unendlich vielen Knotenpunkte z_0, z_∞ , indem jedem Punkt je die Ordnungszahl der entsprechenden Generation zuerteilt wird. Unter den Knotenpunkten befindet sich eine wohlbestimmte Verzweigungsstelle (z_0 oder z_∞), welche die niedrigste Ordnungszahl n enthält; wir sagen, das Gebiet H gehöre zur Generation G_n und nennen die erwähnte Verzweigungsstelle „Anfangspunkt“ des Gebietes H (im Gebiete G_0 wählen wir den Punkt z_0 , welcher ohne Einschränkung der Allgemeinheit in $z = 0$ verlegt werden kann, als Anfangspunkt der zwei oder drei Gebiete H , welche diesen Punkt als Randpunkt haben). Der Anfangspunkt teilt die Randkurve Γ in zwei Äste, und die auf diesen Ästen liegenden Knotenpunkte z_0, z_∞ erhalten der Reihe nach die Ordnungszahlen $n + 1, n + 2, \dots$.

6. Es sei H ein zur Generation G_n gehörendes Gebiet. Mittels der Umkehrfunktion $w(z)$ der Abbildungsfunktion $z(w)$ wird es auf ein un-

endlich vielblättriges Teilgebiet F_n der Riemannschen Fläche F abgebildet, welches genau einen Windungspunkt der Fläche F enthält. Wenn dieser Windungspunkt z. B. über dem Grundpunkt $w_3 = 1$ liegt, so setzt sich die Berandung von F_n aus unendlich vielen Halbstrahlen $(0, \infty)$ zusammen, welche gewisse der Strahlen $\arg w = \pm \frac{\pi}{3}$, $\arg w = \pm \frac{2\pi}{3}$ als Spurlinien haben. F_n enthält als Teilgebiet die einfach zusammenhängende universelle Ueberlagerungsfläche des zweifach zusammenhängenden Gebietes, welches entsteht, wenn man aus dem Zweieck $|\arg w| < \frac{\pi}{3}$ den Punkt $w_3 = 1$ entfernt. Andererseits liegen sämtliche Punkte des Zweiecks $|\arg w - \pi| < \frac{\pi}{3}$ außerhalb des Gebietes F_n .

Mittels der Funktion

$$\omega(z) = \frac{1}{\pi} \log \frac{1 + w(z)}{1 - w(z)} + i c,$$

wo ein beliebiger Zweig des Logarithmus genommen wird, und wo c eine reelle Konstante ist, wird nun das Gebiet H auf ein schlichtes Gebiet H_ω der ω -Ebene abgebildet, welches die Halbebene $\Re(w) > \frac{\log 3}{2\pi}$ enthält und das selbst ein Teilgebiet der Halbebene $\Re(w) > -\frac{\log 3}{2\pi}$ ausmacht. Die Randkurve I_ω von H_ω verläuft im Streifen $|\Re(w)| < \frac{\log 3}{2\pi}$ und geht, falls man der Konstante c einen geeigneten ganzzahligen Wert gibt, durch die Punkte ik ($k=0, \pm 1, \dots$), so daß der Punkt $\omega = 0$ dem „Anfangspunkt“ des Gebietes H entspricht.

Es sei r eine beliebige Zahl $> n$. Man bestätigt leicht, daß der Kreis $|\omega| = r - n$ die Kurve I_ω in genau zwei Punkten A_r und B_r schneidet; wenn r vom Wert $r = n$ ins Unendliche wächst, so durchlaufen die Punkte A_r und B_r die beiden Bogen, in welche die Randkurve I_ω durch den Punkt $\omega = 0$ geteilt wird. Geht man zu der $\zeta = \log \omega$ -Ebene über, so wird H_ω in ein Streifengebiet H_ζ und der Bogen $A_r B_r$ der Kreislinie $|\omega| = r - n$ in einen zur imaginären Achse parallelen Querschnitt $\Re(\zeta) = \log(r - n)$ übergeführt, dessen Länge kleiner als 2π ist.

Andererseits schneide man den Kreis $|z| < R$ auf längs einem Ast der Figur B , der den Nullpunkt mit der Peripherie $|z| = R$ verbindet, und bilde mittels eines beliebig festgelegten Zweiges der Funktion

$t = \log z$ den aufgeschnittenen Kreisbereich auf einen von zwei Parallelkurven und von der Geraden $\Re(\log z) = \log R$ begrenzten Bereich ab; dem Gebiet H entspricht hierbei ein gewisses Gebiet H_t der t -Ebene und dem Querschnitt A_r, B_r von H_0 ein Querschnitt von H_t , dessen Länge mit $s(r)$ bezeichnet werden soll. Mit Hilfe der Schwarzschen Ungleichung findet man, daß

$$(s(r))^2 = \left(\int_{A_r B_r} \left| \frac{d \log z}{d \log w} \right| |d \log w| \right)^2 \leq \int_{A_r B_r} |d \log w| \cdot \int_{A_r B_r} \left| \frac{d \log z}{d \log w} \right|^2 |d \log w|$$

$$< 2\pi \frac{dA(r)}{d \log(r-n)} \leq 2\pi r \frac{dA(r)}{dr},$$

wo $A(r)$ der Flächeninhalt des von dem betrachteten Querschnitt, von der Bildkurve I_t der Linie I und von der Geraden $\log |z| = \log R_0$ begrenzten Gebietes ist, wobei R_0 eine beliebig festgelegte Zahl des Intervalles $0 < R_0 < R$ ist und der Parameterwert r so groß ($r > r_0$) gewählt werden soll, daß auf dem genannten Querschnitt $|z| > R_0$ gilt. Es wird mithin

$$(1) \quad \Sigma (s(r))^2 < 2\pi r \frac{d\mathcal{F}}{dr},$$

wo die Summation über alle Gebiete H zu erstrecken ist, welche zu den Generationen $G_0, G_1, \dots, G_n$ gehören ($n < r$), und wo $\mathcal{F}(r) = \Sigma A(r)$.

Die Anzahl der Querschnitte $s(r)$ sei gleich $\lambda(r)$, und die Gesamtlänge $\Sigma s(r)$ gleich $\alpha(r)$. Gemäß der Schwarzschen Ungleichung ergibt sich, daß

$$(2) \quad \Sigma (s(r))^2 \geq \frac{(\Sigma s(r))^2}{\lambda(r)} = \frac{(\alpha(r))^2}{\lambda(r)}.$$

Wir bezeichnen mit β die größere der Zahlen 0 und $2\pi - \alpha$:

$$(3) \quad \beta(r) = \max(0, 2\pi - \alpha),$$

so daß also $\alpha \geq 2\pi - \beta$ und somit $\alpha^2 \geq 4\pi^2 - 4\pi\beta$, und finden aus (1) und (2), daß

$$\frac{4\pi^2}{\lambda(r)} - \frac{4\pi\beta(r)}{\lambda(r)} < 2\pi r \frac{d\mathcal{F}}{dr}$$

und also, für $r > r_0$,

$$(4) \quad 2\pi \int_{r_0}^r \frac{dr}{r \lambda(r)} \leq \mathcal{F}(r) - \mathcal{F}(r_0) + 2 \int_{r_0}^r \frac{\beta(r) dr}{r \lambda(r)}.$$

7. Wir benutzen jetzt folgenden Hilfssatz, dessen Beweis im 3. Abschnitt gegeben werden soll:

Hilfssatz. Wenn die Riemannsche Fläche zum hyperbolischen Typus gehört, so ist das Integral

$$(5) \quad \int_{r_0}^{\infty} \frac{\beta(r) dr}{r \lambda(r)}$$

konvergent.

Nimmt man nun an, daß der hyperbolische Fall vorliegt, so strebt der Flächeninhalt $\mathcal{F}(r) - \mathcal{F}(r_0)$ für $r \rightarrow \infty$ gegen einen endlichen Wert $< 2\pi \log \frac{R}{R_0}$, und die rechte Seite der Beziehung (4) liegt somit unter der endlichen Schranke

$$2\pi \log \frac{R}{R_0} + 2 \int_{r_0}^{\infty} \frac{\beta(r) dr}{r \lambda(r)},$$

woraus folgt, daß das links stehende Integral in (4) für $r \rightarrow \infty$ konvergiert.

Die gegebene Riemannsche Fläche gehört folglich sicher dem parabolischen Typus an, sobald das zuletzt genannte Integral divergent ist. Wegen der evidenten Beziehung (vgl. Einleitung): $\lambda(r) = \sigma(v) + 2$ für $v-1 < r \leq v$, sind dasselbe Integral und die Reihe

$$\sum_1^{\infty} \frac{1}{n \sigma(n)}$$

gleichzeitig konvergent oder divergent, woraus sich die Richtigkeit der in unserem Satze ausgesprochenen Behauptung ergibt.

§ 3. Beweis des Hilfssatzes

8. Die Riemannsche Fläche F gehöre dem hyperbolischen Typus an; sie läßt sich also auf einen *endlichen* Kreis $|z| < R$ schlicht abbilden, dessen Radius R ohne Einschränkung der Allgemeinheit größer als 1 angenommen werden kann.

Gibt man dem oben eingeführten Parameter r einen ganzzahligen Wert n , so verbinden die Kurvenstücke a_r, b_r , welche in der z -Ebene als Bilder der Kreisbogen A_r, B_r erscheinen, die Knotenpunkte der n -ten Generation G_n und bilden zusammen eine geschlossene Kurve C_r . Für nichtganzzahlige Werte r ($n - 1 < r < n$) verbinden wir die Endpunkte der zu zwei Nachbargebieten H gehörigen Endpunkte a_r und b_r , welche im allgemeinen nicht zusammenfallen, mittels eines Teilstückes γ_r desjenigen Zweiges Z der Figur B , auf welchem die betreffenden Endpunkte liegen. Der Verbindungsbogen γ_r ist ein Teilbogen desjenigen Segmentes von Z , welches von den zur $(n-1)$ -ten und zur n -ten Generation gehörenden Knotenpunkte s_0, s_∞ (oder s_∞, s_0) begrenzt wird.

In der $t = \log s$ -Ebene entspricht der durch obige Konstruktion hergestellten geschlossenen Linie C_r ein Kurvenbogen, dessen Länge mindestens gleich 2π ist. Wenn $\gamma(r)$ die Länge des Bildbogens eines der hinzugefügten Verbindungsbogen γ_r bezeichnet, so ist die Gesamtlänge des genannten Kurvenbogens andererseits gleich $\sum s(r) + \sum \gamma(r) = \alpha(r) + \sum \gamma(r)$, und es wird mithin $2\pi - \alpha \leq \sum \gamma$, und also, gemäß (3),

$$(6) \quad \beta(r) \leq \sum \gamma(r).$$

9. Um die Größen $\gamma(r)$ abzuschätzen, nehmen wir ein beliebiges der Kurvenstücke γ_r und fixieren das Astsegment $P_1 P_2$, welches γ_r als Teilbogen enthält und von zwei aufeinander folgenden Verzweigungsstellen P_1 und P_2 begrenzt wird. Die Knotenpunkte P_1 und P_2 mögen zu den Generationen G_m und G_{m+2p+1} gehören. Wenn der Parameterwert r das Intervall $m \leq r \leq m + 2p + 1$ durchläuft, so bewegen sich die Endpunkte a_r und b_r des Kurvenstückes γ_r stetig von P_1 bis zu P_2 . Die zwischen diesen Endpunkten liegenden Knotenpunkte, deren Zahl $2p$ beträgt, entsprechen gewissen Zweiecken k_0, k_∞ , welche zusammen mit den beiden Dreiecken, in welchen die Endpunkte P_1 und P_2 liegen, ein Gebiet bilden, welches im folgenden mit K bezeichnet werden soll.

Dem Bogen $P_1 P_2$ entspricht in der w -Ebene eine Linie, welche als Spurlinie eine der Geraden hat, die aus der reellen Achse durch eine

Drehung um $\frac{2\pi\nu}{3}$ ($\nu = 0, 1, 2$) entstehen; im folgenden wird $\nu = 0$ angenommen. Ferner möge dem Endpunkt P_1 der Nullpunkt $w = 0$ entsprechen; der Bildpunkt von P_2 fällt dann in den Punkt $w = \infty$. Vermittels der Funktion

$$y = (-1)^p \log \frac{w - w_1}{w - w_2} - \frac{\pi i}{3},$$

bilden wir die universelle Ueberlagerungsfläche des zweifach zusammenhängenden Gebietes $w \neq w_1, w_2$, auf die schlichte y -Ebene ab. Setzt man hier $w = w(z)$, so läßt sich ein Zweig des Logarithmus bestimmen, so daß das Gebiet K in den schlichten Parallelstreifen $|\Im(y)| < (p+1)\pi$ übergeführt wird. Den Knotenpunkten z_0 des Bogens $P_1 P_2$ entsprechen bei dieser Abbildung die Punkte $y = -(p + \frac{2}{3})\pi i + 2\pi\nu i$, den Knotenpunkten z_∞ die Punkte $y = (p + \frac{2}{3})\pi i - 2\pi\nu i$ ($\nu = 0, 1, \dots, p$), so daß also der Abstand zwischen zwei aufeinander folgenden Bildpunkten entweder gleich $\frac{2\pi}{3}$ oder gleich $\frac{4\pi}{3}$ ist. Dem Bogen $P_1 P_2$ ist die Strecke $|y| \leq (p + \frac{2}{3})\pi$ der imaginären Achse zugeordnet.

Wir nehmen nun einen Wert r des Intervalles $m \leq r \leq m + 2p + 1$. Dem Bogen γ_r entspricht in der y -Ebene ein Segment der imaginären Achse, welche als Teilstrecke eines Knotenintervalles höchstens die Länge $\frac{4\pi}{3}$ hat. Wir wollen jetzt die Länge $\bar{\gamma}_{(r)}$ des Bogens γ_r mit Hilfe der Verzerrungssätze von *Koebe* und *Bieberbach* abschätzen.⁵⁾

Es sei $y = iz_r$ der Punkt, welcher dem Endpunkt $z = a_r$ des betrachteten Bogens γ_r entspricht; nach obigem gilt für $m + p + \nu \leq r \leq m + p + \nu + 1$ ($|\nu| = 0, 1, \dots, p$) die Ungleichung

$$(7) \quad (p+1)\pi - |\gamma_r| \geq \pi \left(\frac{1}{3} + p - |\nu| \right) > 1 + p - |\nu|.$$

⁵⁾ Der hier gemeinte Satz lautet: Wenn $f(z) = a_1 z + \dots$ den Kreis $|z| < \rho$ schlicht und konform abbildet, so liegt der Wert des Quotienten

$$\left| \frac{f'(z_1)}{f'(z_2)} \right|,$$

wo z_1 und z_2 zwei beliebige Punkte des Kreises $|z| \leq \theta\rho$ sind ($0 \leq \theta < 1$), unterhalb einer Schranke $Q(\theta)$, die nur von θ abhängig ist (es ist z. B. $Q(\theta) < \left(\frac{1+\theta}{1-\theta}\right)^4$). — Der *Bieberbachsche* Flächensatz besagt bekanntlich, daß der Flächeninhalt des Bildgebietes wenigstens gleich $\pi |a_1|^2 \rho^2$ ist.

Wir wenden zuerst den *Koebeschen* Verzerrungssatz an in dem Kreis vom Radius 2π , der den Punkt $y = iy_r$ zum Mittelpunkt hat, falls $(p+1)\pi - |y_r| \geq 2\pi$, und den Punkt $y = \pm(p-1)\pi i$, falls $|\pm(p+1)\pi - y_r| < 2\pi$. Man findet derart, daß in jedem Punkt y des Intervalles $|y| \leq (p + \frac{2}{3})\pi$ der imaginären Achse, für welchen der Abstand $|y - iy_r| \leq \frac{4\pi}{3}$ ist, die Ungleichung

$$\left| \frac{dz}{dy} \right| < N_1 \left| \frac{dz}{dy} \right|_{y=iy_r}$$

besteht, wo N_1 eine numerische Konstante ist.

Hierauf schlagen wir um den Punkt $y = iy_r$ einen Kreis mit dem Radius $(p+1)\pi - |y_r| = \varrho(r)$, und erhalten mit Hilfe des Flächensatzes für den Flächeninhalt des in der z -Ebene liegenden Bildgebietes dieses Kreises die untere Schranke $\pi \varrho^2 \left| \frac{dz}{dy} \right|^2$, wo $y = iy_r$ zu nehmen ist. Da dieser Flächeninhalt andererseits höchstens gleich dem Flächeninhalt πq des Gebietes K ist, so wird

$$\left| \frac{dz}{dy} \right|_{y=iy_r} < \frac{\sqrt{q}}{\varrho(r)},$$

und also

$$\left| \frac{dz}{dy} \right| < \frac{N_1 \sqrt{q}}{\varrho(r)}$$

für jeden innerhalb des Intervalles $|y| \leq (p + \frac{2}{3})\pi$ liegenden Punkt der Strecke $|y - iy_r| < \frac{4\pi}{3}$ der imaginären Achse. Diese Ungleichung ergibt für die Länge $\bar{\gamma}(r)$ des Bogens γ_r die obere Schranke

$$\bar{\gamma}(r) < \frac{4\pi N_1}{3} \cdot \frac{\sqrt{q}}{\varrho(r)}.$$

Wir nehmen nun an, daß der Bogen $P_1 P_2$ außerhalb des Einheitskreises $|z| = 1$ gelegen ist. Dann gilt $\frac{d \log z}{dz} \leq 1$, woraus hervorgeht, daß auch die Länge $\gamma(r)$ des Bildbogens von γ_r (in der $\log z$ -Ebene) der Ungleichung

$$\gamma(r) < \frac{N_2 \sqrt{q}}{\rho(r)}$$

genügt, wo N_2 die numerische Konstante $\frac{4\pi N_1}{3}$ bezeichnet. Unter Berücksichtigung der Beziehung (7), wo der rechts stehende Ausdruck den Wert des Radius $\rho(r)$ angibt, wird ferner

$$(8) \quad \int_m^{m+2p+1} \gamma(r) dr < 2N_2 \sqrt{q} \sum_{v=1}^{p+1} \frac{1}{v} < 2N_2 \sqrt{q} \log 3 (p+1).$$

Diese Beziehung haben wir unter der Voraussetzung hergeleitet, daß m und $m+2p+1$ ganze Zahlen sind, welche die Ordnungszahlen der aufeinander folgenden Verzweigungsstellen P_1 und P_2 angeben, die den betrachteten Bogen $P_1 P_2$ begrenzen. Nun sieht man unmittelbar ein, daß das obige Resultat auch dann besteht, wenn die genannten ganzen Zahlen durch beliebige Werte r_1 und r_2 ($0 < r_1 < r_2$) ersetzt werden, wobei für $P_1 P_2$ ein von inneren Verzweigungsstellen freier Bogen zu wählen ist, dessen Endpunkte P_1 und P_2 den Parameterwerten $r=r_1$ bzw. r_2 zugeordnet sind. Es wird somit für einen beliebigen derartigen Bogen $P_1 P_2$

$$(9) \quad \int_{r_1}^{r_2} \gamma(r) dr < 2N_2 \sqrt{q} \log 3 (r_2 - r_1 + 1) < N_3 \sqrt{q} \log 3 (r_2 + 1)$$

wo $N_3 = 2N_2$ eine numerische Konstante und πq den Flächeninhalt des dem Bogen $P_1 P_2$ entsprechenden Gebietes K bezeichnen.

Wir fixieren jetzt einen so großen Wert $n_0 > 1$, daß sämtliche Punkte z der Generation G_n für $n > n_0$ außerhalb des Einheitskreises $|z| = 1$ liegen, was möglich ist, da der Radius R voraussetzungsgemäß größer als 1 ist. Es sei $B(n_0, r)$ derjenige Teil der Baumfigur B , dessen Punkte ($z = a_r$) dem Parameterintervall (n_0, r) entsprechen. Die Zweige von $B(n_0, r)$ zerlegen wir in Segmente $S(P_1, P_2)$, deren Endpunkte P_1, P_2 entweder Verzweigungsstellen oder den äußersten Werten n_0 und r angehörende Punkte (a_{n_0} bzw. a_r) sind, so daß im Innern von S keine Verzweigungsstellen vorhanden sind. Die Anzahl dieser Segmente ist offenbar höchstens gleich der doppelten Anzahl der innerhalb der Kurve C_r belegenen Verzweigungsstellen von B , vermehrt um 1, also höchstens gleich $2\lambda(r)$ (vgl. Nr. 7).

Integriert man die zu einem derartigen Bogen S gehörende Größe $\gamma(r)$ zwischen den Grenzen r_1 und r_2 , welche den Endpunkten P_1 und P_2 zugeordnet sind, so wird, da jedenfalls $r_2 \leq r$ ist,

$$(10) \quad \int_S \gamma(r) dr < N_3 \sqrt{q} \log 3(r+1) < N_4 \sqrt{q} \log r,$$

wo also πq den Flächeninhalt des dem Bogen $P_1 P_2$ entsprechenden Gebietes K ist, und N_4 eine numerische Konstante bezeichnet.

10. Mittels letzter Ungleichung läßt sich nun das Integral

$$b(r) \equiv \int_{n_0}^r \beta(u) du$$

in folgender Weise abschätzen. Zunächst gilt gemäß der Beziehung (6)

$$b(r) \leq \int_{n_0}^r (\Sigma \gamma(u)) du,$$

wo die Summation über sämtliche, dem Parameterwert $r = u$ entsprechenden Bogen S zu erstrecken ist. Das letzte Integral läßt sich aber schreiben

$$\int_{n_0}^r (\Sigma \gamma(u)) du = \Sigma \int_S \gamma(u) du,$$

wo über alle Bogen S der Figur $B(n_0, r)$ zu summieren ist, und es wird somit, nach der Beziehung (10)

$$b(r) < N_4 \log r \cdot \Sigma \sqrt{q}.$$

Beachtet man nun, daß die Anzahl der Bogen S höchstens $2\lambda(r)$ beträgt, so wird nach der Schwarzschen Ungleichung

$$\Sigma \sqrt{q} \leq \sqrt{2\lambda(r)} \cdot \sqrt{\Sigma q}.$$

Hier bedeutet $\pi \Sigma q$ den gesamten Flächeninhalt der dem Bogen S entsprechenden Gebiete K . Unter Beachtung der Definition dieser Gebiete

sieht man unmittelbar ein, daß jeder Punkt des Kreises $|z| < R$, welcher zu einem Zweieck (k_0 oder k_∞) gehört, höchstens in *einem* Gebiet K enthalten ist, während ein Punkt z eines Dreieckes (k_0 oder k_∞) höchstens zu *drei* verschiedenen Gebieten K gehört. Hieraus folgt, daß der Gesamthalt $\pi \Sigma q$ den dreifachen Flächeninhalt des Kreises $|z| < R$ nicht übersteigen kann, so daß also für jedes $r > n_0$

$$\Sigma q \leq 3 R^2.$$

Zusammenfassend wird also schließlich für $r > n_0$

$$b(r) < N \sqrt{\lambda(r)} \cdot \log r,$$

wo N eine von r unabhängige Zahl ist.

Nummehr läßt sich die Endlichkeit des Integrals (5) leicht beweisen. Durch partielle Integration findet man für $r > r_0 > n_0 > 1$, da $\lambda(r) > 1$ ist,

$$\begin{aligned} \int_{r_0}^r \frac{\beta(r)}{r \lambda(r)} dr &= \int_{r_0}^r \frac{db(r)}{r \lambda(r)} \leq \frac{b(r)}{r \lambda(r)} + \int_{r_0}^r \frac{b(r) dr}{r^2 \lambda(r)} + \int_{r_0}^r \frac{b(r) d\lambda(r)}{r \lambda^2} \\ &\leq N \frac{\log r}{r} + N \int_{r_0}^r \frac{\log r dr}{r^2} + N \int_{r_0}^r \frac{\log r}{r} \frac{d\lambda}{\lambda^{3/2}} \\ &< N \left(\frac{\log r_0}{r_0} + \frac{1}{r_0} + \int_{r_0}^r \frac{\log r}{r} \frac{d\lambda}{\lambda^{3/2}} \right). \end{aligned}$$

Da ferner $\frac{\log r}{r} < 1$ für $r \geq r_0 > 1$, so wird der Ausdruck rechts kleiner als

$$N \left(2 + \int_{r_0}^r \frac{d\lambda}{\lambda^{3/2}} \right) = 2N \left(1 + \frac{1}{\sqrt{\lambda(r_0)}} - \frac{1}{\sqrt{\lambda(r)}} \right) < 4N,$$

woraus die behauptete Konvergenz des Integrals (5) hervorgeht.

(Eingegangen den 1. April 1932)

Sur l'axiomatique de la théorie des ensembles et sur la logique des relations

par F. GONSETH, Zurich

1. Remarques sur la méthode axiomatique

Nous allons commencer par soumettre à quelques critiques la méthode dite « axiomatique » par laquelle, suivant l'exemple de M. Zermelo, on a cherché à mettre la théorie des ensembles à l'abri des antinomies bien connues.

Depuis le mémoire de M. Zermelo et les objections présentées par Poincaré, la situation est loin de s'être définitivement éclaircie et l'unanimité ne s'est point faite sur le sens et la portée de la méthode elle-même. Nous allons en reprendre les traits essentiels. La première chose à faire est de renoncer à la définition de Cantor :

Un ensemble est la réunion en un tout de certains objets perçus ou pensés, les éléments de l'ensemble. Cette définition s'est révélée trop large et ouvre la voie aux paradoxes que l'on connaît. Il faut renoncer également à toute autre définition explicite de ce qu'il faut entendre par ensemble, par élément, par la relation d'inclusion d'un élément x dans un ensemble y (la relation $x \varepsilon y$), etc. Ces notions et relations ne doivent prendre — dit-on — que le sens que comportent les axiomes et les définitions qui s'y rapportent, axiomes et définitions qui sont à énoncer explicitement.

On va donc commencer par imaginer certaines choses que l'on appellera *ensembles* et qu'on désignera par les lettres $a, b, c \dots$; on imaginera ensuite entre ces choses trois *relations*

1. la relation $a \varepsilon b$ (a est élément de b)
2. la relation $a < b$ (a est sous-ensemble de b)
- et 3. la relation $a = b$ (a est identique à b).

Ces relations n'ont aucun sens par elles-mêmes. Toute leur signification doit être contenue dans les définitions et axiomes que voici :

Déf. 1. Si tout x qui est à a dans la relation 1 est aussi à b dans la même relation, a et b sont alors dans la relation 2.

Déf. 2. Si a et b sont aussi bien dans la relation $a < b$ que dans la relation $b < a$, alors a et b sont dans la relation 3.

Axiome 1. Si $a = b$ et si a est élément de A , b est aussi élément de A .

Axiome 2. Les éléments de deux ensembles A et B (différents) forment à eux seuls les éléments d'un nouvel ensemble C .

Axiome 3. Si A est un ensemble, les éléments de ses éléments forment un ensemble.

Axiome 4. Tous les sous-ensembles d'un ensemble forment à eux seuls les éléments d'un nouvel ensemble.

Axiome 5. Les éléments d'un ensemble qui possèdent un attribut bien déterminé (défini) forment un sous-ensemble du premier.

Axiome 6. On peut former un ensemble en choisissant un unique élément dans chaque élément de tout ensemble.

Le système de M. Zermelo comprend enfin un septième axiome qui permet d'affirmer l'existence d'un ensemble infini.

Axiome 7. Il existe un ensemble N tel que

- 1. si l'ensemble nul (c.-à-d. qui n'a aucun élément) existe, c'est un élément de N ;*
- 2. si m est un élément de N , celui-ci contient aussi $\{m\}$ (c.-à-d. l'ensemble dont m est l'unique élément).*

La première question que soulève ce système d'axiomes, c'est naturellement de savoir s'il détermine véritablement les objets de pensée (les ensembles) qu'il vise. On se souvient de l'objection de Poincaré¹⁾ : ... « Quelqu'un qui ne sait pas ce que c'est qu'une *Menge* ne le saura pas davantage lorsqu'il aura appris qu'elle est représentée par le symbole ε , puisqu'il ne sait pas ce que c'est que ε » ... Les axiomes peuvent-ils à eux-seuls remplacer la connaissance intuitive et préalable des êtres mathématiques dont ils fixent les lois d'existence ? C'est là le point sur lequel il nous paraît nécessaire de nous arrêter un instant.

Pour justifier la méthode de M. Zermelo, on la met parfois en parallèle avec la méthode axiomatique usuelle en géométrie. Or il y a tout d'abord une différence essentielle entre le système d'axiomes que nous discutons et celui qui est à la base de la géométrie dite élémentaire. Cette différence est justement mise en lumière par la remarque de Poincaré que nous venons de rappeler. Dans la construction de la géométrie élémentaire, les êtres mathématiques soumis aux axiomes sont des *abstractions suggérées par le monde physique*. L'axiomatique fixe simplement les modalités de cette abstraction, fixe dans la sphère du rationnel les règles d'existence de ces objets abstraits. Dire — avec M. Weyl²⁾ — « que la méthode axiomatique consiste simplement à rassembler *tous* les concepts et les faits fondamentaux, à partir desquels tous les autres concepts et

¹⁾ *H. Poincaré. Dernières pensées, p. 124.*

²⁾ *H. Weyl: Philosophie der Mathematik.*

tous les autres faits d'une science peuvent être ou définis ou déduits, » c'est supposer que les concepts fondamentaux ont déjà pris naissance et qu'on sait déjà quelles sortes de relations on peut établir entre eux : c'est supposer que le *travail de la constitution des abstraits est déjà terminé*. Or sur ce point, l'axiomatique de M. Zermelo prend soin de spécifier qu'elle a coupé les ponts qui pourraient la relier avec la notion, intuitivement fondée, de *collection*. Le concept d'ensemble qu'elle vise n'y doit point être envisagé comme un abstrait suggéré par tel ou tel autre concept de la sphère intuitive : au contraire ce sont les axiomes à eux-seuls qui doivent lui créer, de toutes pièces, une signification et l'appeler à l'existence mathématique. Sans vouloir encore en tirer de conclusion quant à la légitimité de la méthode de M. Zermelo, il nous faut constater la présence d'un *hiatus essentiel* qui ne permet pas de la justifier par comparaison ou identification avec l'axiomatique dite élémentaire.

En revanche, il peut sembler à première vue le parallélisme s'établit de lui-même et de façon parfaite avec l'axiomatique au second degré, où toutes les notions sont repoussées dans le domaine de la logique des relations. L'essentiel de cette méthode est déjà dans les toutes premières lignes des « Grundlagen der Geometrie » de Hilbert.

« Nous imaginons trois catégories d'objets : nous nommons *points* les objets de la première catégorie, etc... »

« Nous imaginons qu'entre les points, les droites et les plans, il existe certaines relations que nous désignerons par les termes : « être sur », « entre », « parallèle », « congruent », « continu ». La description de ces relations, description exacte et suffisante pour les buts de la géométrie se fait par le moyen des « axiomes de la géométrie ».

Les objets géométriques doivent ainsi être vidés de leur contenu intuitif — pour employer une expression consacrée — et n'avoir d'autres propriétés que celles que les relations, purement abstraites, à établir entre eux vont leur conférer.

Les axiomes ³⁾ qui fixeront le sens de la locution « être sur » ou si l'on préfère de la « relation d'incidence » seront alors, par exemple, les suivants :

Axiome 1. On peut toujours imaginer pour deux objets de catégories différentes, une relation $J(*,*)$ (la relation d'incidence).

Axiome 2. Pour ces deux objets, ou bien cette relation *est*, ou bien elle *n'est pas*.

Axiome 3. Cette relation est symétrique, etc, etc...

³⁾ M. Geiger. Systematische Axiomatik der euklidischen Geometrie.

En face de cette axiomatisation sur le plan du *logique pur*, on peut prendre deux attitudes. La première est toute naturelle : elle consiste à regarder les « objets » et les « relations » soumis à l'axiomatisation comme de nouveaux abstraits suggérés par les notions et les relations géométriques du degré élémentaire : comme des *abstrais au second degré*. Au moment où nous établissons les axiomes, nous pouvons alors regarder à nouveau le processus de la constitution des notions abstraites comme terminé. De ce point de vue, cette axiomatisation est parfaitement semblable à l'axiomatisation élémentaire, elle répond complètement à la définition de M. Weyl : nous la regardons comme légitime. Il va sans dire que cette première attitude ne peut pas être adoptée dans la méthode de M. Zermelo, le lien avec l'intuitif étant expressément rompu.

La seconde attitude est moins clairement définie. Elle consiste essentiellement à admettre que le système des axiomes fournit en quelque sorte une *détermination implicite* des notions qui y figurent. Nous ne voulons pas discuter à cet endroit si, une fois qu'on a supposé les liens avec l'intuitif complètement dénoués, une détermination de ce genre reste possible. Nous voulons nous borner à remarquer que, dans le cas qui nous occupe, il faut en tous cas admettre que la notion de relation logique soit préalablement en notre possession. Il ne peut être question que cette notion soit elle aussi « définie implicitement », les axiomes caractérisant tout au plus les trois relations fondamentales et non la notion même de relation.

C'est là une grave objection, car la notion de relation, dans sa plénitude intuitive, donne lieu aux mêmes antinomies (*mutatis mutandis*) que la notion d'ensemble selon Cantor. Il faudrait donc que, par une axiomatisation adéquate, cette notion elle-même eût reçu sa justification. Or on sait bien que la Logique théorique n'y parvient que de façon insuffisante.

En résumé, ces critiques cherchent à mettre en lumière les deux faits suivants :

1. La méthode axiomatique de M. Zermelo diffère sur un point essentiel de la méthode axiomatique qui s'est constituée à partir de l'axiomatique élémentaire de la géométrie ; celui de la constitution préalable et par *abstraction*, des notions fondamentales.
2. Dans la mesure où elle prétend se rendre indépendante des notions intuitives, la théorie de M. Zermelo fait appel à la notion de *relation*. Or celle-ci, dans la plénitude de son sens, est antinomique, et les axiomes ne fournissent pas le moyen de la circonscrire.

Les points sur lesquels nous venons d'insister ne sont pas les seuls où les axiomes précédents prêtent à la critique. On a déjà souvent fait observer que le 5^e fait intervenir la notion d'*attribut bien déterminé (definit)* qui aurait elle-même besoin d'être axiomatiquement fondée. Nous ne voulons pas revenir non plus sur la discussion que l'axiome 6 (du choix) a soulevé. En revanche, nous appuyons spécialement sur les critiques que nous venons de formuler, parce qu'elles s'appliquent également aux systèmes axiomatiques dérivés de celui de M. Zermelo (où par exemple on cherche à parer au point faible signalé à l'axiome 5 e).

La position axiomatique de ces systèmes est d'ailleurs telle, qu'elle ne permet pas de porter une clarté suffisante sur la question-même en vue de laquelle ils ont été inventés : celle des antinomies. Il est vrai que les antinomies connues peuvent être évitées, mais la méthode ne permet pas d'apercevoir ce qu'on pourrait appeler la cause ou la racine des paradoxes, et, à cause du « nuage axiomatique » dans lequel la notion d'ensemble reste enveloppée, la possibilité d'autres paradoxes reste ouverte, sans qu'on puisse prévoir si la méthode restera efficace.

Le but de ce travail est d'esquisser une autre axiomatisation des notions de la théorie des ensembles. Cette nouvelle méthode permet — nous semble-t-il — d'éviter les différents écueils que nous venons de signaler. En particulier elle fera voir que la racine commune des antinomies doit être aperçue dans le fait très simple que voici :

Les éléments des ensembles selon Cantor sont susceptibles de posséder par eux-mêmes des attributs intrinsèques, avant même d'être envisagés comme éléments d'un ensemble — attributs qui ne se limitent pas à caractériser l'existence individuelle d'un élément dans l'ensemble.

Des attributs de ce genre sont par exemple « d'être une paire de bas » ou « d'être un ensemble qui ne se contient pas lui-même comme élément » ou même plus simplement « d'être un ensemble ».

2. Les objets et les relations de la logique pure

« *L'attitude axiomatique* » dans laquelle nous allons nous placer est celle que nous avons décrite tout à l'heure en l'opposant à la méthode de M. Zermelo ; l'axiomatisation y comprend les deux phases suivantes :

a) Constitution des notions fondamentales, celles-ci étant à abstraire de certaines autres notions (telles que celles d'objet, de collection, de nombre, etc) qui doivent être considérées comme étant préalablement en notre possession.

b) Énumération des axiomes.

Il faut une analyse assez attentive pour distinguer la première phase dans l'axiomatisation « élémentaire » de la géométrie. Il n'en reste pas moins vrai que c'est elle qui donne sa signification véritable à toute construction axiomatique. Ceci se manifeste spécialement dans le fait que le processus de l'abstraction ne peut être que suggéré, mais jamais « verbalement défini ». C'est ainsi que les droites imparfaitement réalisées dans le monde physique ont suggéré aux premiers géomètres la notion géométrique de droite, la notion de lieu précis, celle de point mathématique. Le passage de la notion intuitive à la notion géométrique est le phénomène mathématique par excellence, mais il reste plus ou moins sous-entendu.

De la même façon nous allons nous efforcer de faire entendre ce que doivent être les objets et les relations de la *logique pure*.

Toute construction mathématique peut être mise sous la forme d'un réseau de relations entre certains êtres géométriques ou arithmétiques. Si l'on abstrait de ces relations et des objets mathématiques qu'elles relient tout ce qui a trait à la grandeur, à la forme, etc, il en reste ce qu'on pourrait appeler le *contenu de pure logique*. Les *objets de la logique pure (éléments logiques)* ne sont alors susceptibles que des propriétés suivantes :

Ils sont ou identiques ou différents (sans qu'il y ait lieu de préciser en quoi ils diffèrent).

Ils n'ont, par ailleurs, pas d'autre rôle que de figurer dans les *relations logiques*, ou mieux encore que de servir de point d'attache aux *liaisons logiques*. (La différence entre les deux expressions précédentes sera précisée tout à l'heure.)

Il est bien entendu que, de tout objet dont on parle, on suppose qu'il est individuellement reconnaissable, et qu'il peut être représenté par un symbole qui lui soit particulier.

Les *liaisons logiques* n'ont à leur tour que les propriétés suivantes :

Deux liaisons logiques sont ou identiques ou différentes (sans qu'il y ait lieu de préciser *en quoi* elles diffèrent).

Par ailleurs, elles n'ont d'autre rôle que de relier les objets logiques dont nous avons parlé.

Une liaison logique s'établit entre deux ou plus de deux éléments. Supposons par exemple qu'elle soit établie entre les deux éléments logiques a et b et désignons-la par le symbole (a, b) . Dans ce symbole, les lettres a et b ne désignent pas deux éléments plus ou moins quelconques (comme ce serait le cas dans la relation arithmétique $a < b$ par exemple).

Il faut au contraire les envisager comme les signes reconnaissables de deux éléments individualisés. La liaison logique (a, b) ne s'établit alors qu'entre ces deux éléments. Si, au contraire, il peut y avoir un sens à dire qu'il existe entre un a et un b , variables au sein de certaines collections, *toujours une même liaison logique*, nous dirons que le symbole (a, b) représente une *relation logique*. Mais le procédé qui permettra de reconnaître quand deux liaisons logiques peuvent être dites identiques ou équivalentes doit encore être expliqué.

Ainsi par exemple, la relation

$F(a, b)$ d'un fils a quelconque à son père b , n'est naturellement pas une relation logique, mais une relation de parenté. De même, ni la relation $G(x, y) \equiv x < y$ entre un nombre x quelconque et tout nombre plus grand y , ni la relation d'incidence $J(a, b)$ d'une droite a quelconque à tout point b de cette droite, ne sont des relations logiques. De même encore ni la liaison « entre Zébédée et son fils », ou bien « entre les nombres 3 et 4 », ou bien « entre l'axe des X et l'origine » ne sont des liaisons logiques. Mais on parvient aux notions de la logique pure à partir des précédentes en faisant abstraction de tout ce qui est parenté, grandeur ou position. Il sera commode de dire que les notions plus ou moins intuitives à partir desquelles une relation logique peut être atteinte par abstraction sont des *réalisations* de la relation logique. Ainsi par exemple, les nombres entiers réalisent certains éléments logiques, et la relation d'un nombre quelconque x à celui qui le suit $x + 1$ réalise une certaine relation logique. De même, on peut imaginer que la succession de 4 à 3 réalise une liaison logique, dont les éléments sont réalisés par les nombres 3 et 4. Nous avons insisté sur ces exemples très simples pour bien opposer le domaine des notions abstraites et les domaines où celles-ci se réalisent. S'il est parfaitement possible de concevoir les objets abstraits et les notions que nous avons en vue, qu'il soit bien clair aussi qu'il est impossible de les réaliser dans leur perfection : toute réalisation *in concreto* est du genre que nous venons de dire. Ceci n'est en aucune façon une faiblesse de notre théorie. C'est au contraire un caractère qui se retrouve dans toute constitution d'abstrait ; un caractère qui est bien visible aussi dans l'axiomatique élémentaire. Le fait qu'il est impossible de réaliser *in concreto* un cercle parfait, ou une droite « absolument rectiligne » n'a jamais été un obstacle à l'érection de la géométrie en science rationnelle. De la manière même dont les notions de la logique pure trouveront une réalisation dans celles de la géométrie ou de l'arithmétique (et d'ailleurs aussi dans les notions intuitives), la notion de droite trouve

une réalisation dans la « trajectoire d'un rayon lumineux » ou dans le « trait tracé à la règle ».

Le but véritable de la *méthode axiomatique*, c'est de dégager les notions abstraites de la « matière » de leurs réalisations ; le but de tout *système d'axiomes* de fixer les règles selon lesquelles l'abstrait doit être traité.

Le but des lignes qui vont suivre est donc de *dégager de leurs réalisations les notions de la logique pure, ou mieux encore de construire axiomatiquement la notion même de « Logique pure »*. Il se présentera que cette façon de faire écartera tout naturellement les difficultés relatives aux antinomies.

L'esprit dans lequel notre tentative axiomatique va être entreprise étant ainsi fixé, il nous a paru inutile de spécifier chaque fois par la suite ce qui est axiome et ce qui est définition.

3. Notions fondamentales. Principes et axiomes

Passons à la construction du système axiomatique où les notions fondamentales telles que

élément ou objet logique, liaison logique, compatibilité et incompatibilité de deux ou de plusieurs liaisons, relation logique, structure logique, etc.

sont mises en relation les unes avec les autres.

Nous dirons d'une liaison établie entre deux ou plusieurs éléments qu'elle touche ceux-ci, ou qu'elle les recouvre, ou qu'elle les relie, etc.

Entre deux (ou plusieurs) éléments logiques on peut imaginer un nombre quelconque de liaisons logiques, toutes différentes entre elles. (Deux liaisons sont ou identiques ou différentes, sans qu'il y ait jamais lieu de préciser par quoi elles peuvent ne pas être identiques.)

Il est naturellement très facile d'indiquer des *réalisations* intuitives ou mathématiques qui justifient cet axiome. Par exemple : On peut former autant de groupes différents de trois objets que l'on veut, où entrent deux objets donnés. La présence simultanée de ces deux objets dans un même groupe est une liaison. (Naturellement pas une *liaison logique*.) Ou bien aussi : On peut tracer entre deux points d'un plan autant de chemins différents que l'on veut ; etc., etc.

Deux liaisons sont ou compatibles ou incompatibles (qu'elles soient ou non établies entre les mêmes éléments).

Trois liaisons sont également soit compatibles, soit incompatibles. Elles sont en particulier incompatibles si deux d'entre elles le sont. Elles pourront l'être aussi, bien que compatibles deux à deux.

En général : *Des liaisons en nombre quelconque sont incompatibles si c'est le cas pour une partie d'entre elles.*

Décréter l'incompatibilité de deux liaisons, c'est donc décréter l'incompatibilité pour toutes les combinaisons où ces deux liaisons seraient présentes. C'est là-dessus que se basera tout à l'heure le principe du libre choix des incompatibilités.

A propos de la compatibilité ou de l'incompatibilité de deux liaisons, il y a une remarque essentielle à faire.

A priori, il n'y a aucune incompatibilité entre deux liaisons différentes, tant qu'on n'a encore rien décrété, si ce n'est qu'elles ne sont pas identiques. L'incompatibilité ne peut provenir que d'une mise en rapport de ces liaisons. Tant qu'elles n'existent que de façon purement individuelle et chacune pour soi, toute possibilité de contradiction est exclue. C'est de la même façon que deux idées ne peuvent jamais entrer en opposition, tant qu'elles restent étrangères l'une à l'autre et ne sont pas associées.

Dans la pratique du raisonnement, certaines affirmations d'incompatibilité prendront une forme positive. Si, par exemple, on sait que n liaisons sont incompatibles, et que $n-1$ d'entre elles existent, on en déduira que la n^{e} n'existe pas.

Passons à la notion d'ensemble. Elle se présentera pour nous sous la forme de l'ensemble lié : c'est une collection d'éléments logiques entre lesquels on a établi un certain nombre de liaisons. Pour bien marquer la différence avec la notion habituelle d'ensemble, un ensemble lié sera appelé aussi une *structure logique*.

La construction de ces structures se fera selon les prescriptions de certains principes très simples, qu'il faut considérer comme venant compléter les règles de la logique ordinaire et du nombre, valables dans le fini.

A. Principe de libre extension. Une structure étant donnée, on peut toujours imaginer un nouvel élément différent de tous les éléments déjà existants.

Ce principe ne fait que formuler ce qu'on appelle aussi la liberté des constructions mentales.

B. Principe de libre liaison. Entre deux ou plusieurs éléments, on peut toujours imaginer une nouvelle liaison différente des liaisons déjà existantes.

La justification de ce principe est dans la remarque que nous avons faite plus haut sur la compatibilité des liaisons qui restent sans rapports entre elles. Tant que nous n'aurons pas introduit la possibilité d'identifier deux liaisons dans une structure, toute tentative d'établir une contradiction dérivée de l'existence de certaines liaisons est évidemment sans objet.

Nous appellerons *structures libres*, celles pour lesquelles on n'invoque que les deux principes précédents. Remarquons que les éléments (et les liaisons) d'une structure libre ne sont pas nécessairement en « quantité dénombrable » : les structures dénombrables, continues, etc, s'obtiendront par spécialisation (c'est-à-dire par adjonction d'axiomes) à partir des structures libres.

On pourrait appeler *homogène* une structure où aucun élément ne jouit d'une position ou d'une propriété privilégiée : Une structure de ce genre est par exemple celle où il y a, entre deux éléments quelconques, exactement deux liaisons.

Il y a une certaine *dualité* entre les deux notions d'élément logique et de liaison logique : On pourrait considérer les liaisons comme des éléments ; la propriété de deux ou de plusieurs liaisons de toucher un élément commun serait alors considérée comme établissant une liaison entre ces liaisons-éléments. En tenant compte de cette dualité, on pourrait considérer le principe *A* comme réciproque de *B* sous la forme plus précise, mais moins simple que voici :

On peut toujours imaginer que deux ou plusieurs liaisons viennent se nouer sur un nouvel élément différent de tous les éléments déjà existants.

C. Principe du libre choix des incompatibilités. On peut librement exiger l'incompatibilité de deux ou de plusieurs liaisons choisies à volonté dans une structure.

La justification de ce principe est dans la remarque que nous avons faite concernant les « conséquences » d'une incompatibilité.

Un cas spécial de ce principe est le suivant :

On peut librement supprimer d'une structure toutes les liaisons que l'on veut

qui suggère un principe analogue par dualité :

On peut librement supprimer d'une structure tous les éléments que l'on veut (et les liaisons qui les touchent).

Nous sommes ainsi conduit à la notion de *structure partielle*. On obtient une *structure partielle* par suppression de liaisons et d'éléments.

Il faut naturellement se garder de croire qu'une structure partielle est moins « ample » que la structure originelle. Mais il y a une remarque plus subtile à faire. Il ne faut pas croire que, si l'on a supprimé un élément, cet élément ne se retrouve pas dans la structure partielle. Ceci n'est pas un paradoxe : Un élément logique n'a *a priori* aucune signification et n'en reçoit une que par sa position, son insertion dans une

structure. Si, par conséquent, nous supprimons un élément et que la structure restante soit identique à la structure originelle, l'élément supprimé renaît à l'intérieur de la structure partielle. Ainsi, dans une suite semblable à celle des nombres entiers, si nous en supprimons le premier, celui-ci renaît là où était le second. Quant à savoir quand deux structures peuvent être considérées comme identiques, c'est une question que nous reprendrons après l'introduction des axiomes proprement dits.

D. Principe de contraction. On peut remplacer une structure partielle par un seul élément, sur lequel viendront se nouer toutes les liaisons qui ont été supprimées dans la construction de la structure partielle, et qui aboutissaient à un élément de celle-ci.

Plus important que ce principe sera sa réciproque qui permettra l'opération inverse :

D'. Principe d'insertion. On peut, à la place d'un élément, insérer une structure, par l'opération inverse de celle que nous venons de décrire.

Ce principe permet de construire une structure de structures, pourvu que celles-ci puissent tout d'abord être mises en état de liaison « comme » les éléments d'une structure.

La liste des principes nécessaires n'est pas encore complète. En particulier nous n'avons encore aucune règle concernant l'identification de deux liaisons.

Mais il nous paraît utile d'examiner tout d'abord sur des exemples simples quelle est la portée des principes que nous avons déjà introduits. D'ailleurs, les principes *D* et *D'* pourraient être envisagés comme conséquences des principes qui les précèdent.

4. Les axiomes de l'ordre⁴⁾ et les structures ordonnées

Nous allons examiner par quels axiomes purement restrictifs, les structures ordonnées « dans un sens » peuvent être obtenues à partir d'une structure libre. Conformément aux principes *A* et *B*, nous partirons de la structure libre que nous avons une fois déjà donnée en exemple.

Entre deux éléments quelconques a et b , nous imaginons donc deux liaisons différentes (a, b) et $(a, b)^$.*

On représente ces liaisons de façon plus commode si l'on se sert du fait que a et b peuvent être nommés dans l'ordre inverse et si l'on pose

$$(a, b) = (b, a)^* \quad \text{et} \quad (b, a) = (a, b)^*.$$

⁴⁾ Cf. *B. Russel*. Introduction to mathematical philosophy. Chap. III. The définition of order.

Nous appliquons ensuite le principe *C* du libre choix des incompatibilités.

Axiome O_1 . Les deux liaisons

$$(a, b) \quad \text{et} \quad (b, a)$$

sont incompatibles.

Considérons ensuite deux paires d'éléments *a* et *b*, et *c* et *d*. Si l'on tient compte de l'axiome 2, les combinaisons suivantes sont encore formées de deux liaisons compatibles :

$$\begin{array}{ll} (a, b) & \text{et} \quad (c, d) \\ (b, a) & \text{et} \quad (c, d) \\ (a, b) & \text{et} \quad (d, c) \\ (b, a) & \text{et} \quad (d, c). \end{array}$$

Nous allons décréter que deux de ces combinaisons sont incompatibles. Donc :

Axiome O_2 . A supposer que les liaisons de la combinaison (a, b) et (c, d) soient compatibles, les liaisons (b, a) et (c, d) de même que (a, b) et (d, c) sont à tenir pour incompatibles (tandis que (b, a) et (d, c) restent compatibles).

Considérons ensuite les trois éléments *a*, *b*, *c* et les trois paires *a* et *b*, *b* et *c*, et *c* et *a*. Supposons que les liaisons (a, b) et (b, c) soient compatibles. Les axiomes 1 et 2 ne permettent encore de rien affirmer au sujet des deux combinaisons à trois termes

$$\begin{array}{lll} (a, b) & \text{et} & (b, c) \quad \text{et} \quad (a, c) \\ (a, b) & \text{et} & (b, c) \quad \text{et} \quad (c, a). \end{array}$$

On s'en rend compte immédiatement par exemple sur la réalisation (représentation) suivante. Faisons correspondre aux quatre éléments logiques *a*, *b*, *c*, *d* les sommets d'un tétraèdre (fig. 1) et supposons que les liaisons soient réalisées par les vecteurs qu'on peut tracer entre ces points. On peut alors donner à nos deux axiomes une forme positive. L'axiome 1. dit tout d'abord que l'on aura toujours soit (a, b) soit (b, a) . En d'autres termes, une arête du tétraèdre ne portera jamais qu'un vecteur.

L'axiome 2 exige simplement que, lorsque la liaison à porter sur une arête a été choisie, celle qu'il faut porter sur toute autre arête posée ne puisse être choisie que d'une seule façon.

Ainsi par exemple, si l'on a choisi les liaisons (a, b) et (b, c) , la liaison à porter sur l'arête ac ne peut être que soit (a, c) , soit (c, a) , mais pas les deux concurremment. Nous allons exclure l'une de ces deux possibilités.

Axiome O_3 . A supposer que les liaisons (a, b) et (b, c) soient compatibles, les trois liaisons (a, b) et (b, c) et (c, a) sont incompatibles.

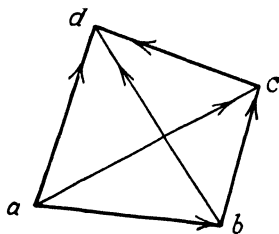


Fig. 1.

Dans la représentation, cela signifie que, s'il existe deux liaisons représentées par deux vecteurs dont le second est attaché à l'extrémité du premier, il existe aussi une liaison représentée par leur somme géométrique.

Les trois axiomes précédents peuvent être appelés les *axiomes abstraits de l'ordre*. Ils permettent de ranger n objets logiques, liés conformément aux axiomes, dans un ordre déterminé par les liaisons seulement. La chose est à peu près évidente sur la réalisation : elle est certainement vérifiée pour les trois éléments a, b, c , ces points se suivent dans l'ordre déterminé par le nombre de liaisons qui en « partent » : de a il en part deux, de b une seule et de c aucune. Introduisons le quatrième point d . Il y aura maintenant un point d'où partiront exactement trois liaisons. Si nous portons sur l'arête ad la liaison (a, d) , ce sera encore une fois le point a , et si nous avons porté la liaison (d, a) , ce serait le point d . On parvient au cas général par induction : En mettant à part le point dont partent $n-1$ liaisons et en raisonnant ensuite sur les $n-1$ éléments restants, tous les éléments sont choisis l'un après l'autre dans un ordre déterminé.

Nous dirons d'une structure qui satisfait aux trois axiomes de l'ordre qu'elle est *ordonnée*. Remarquons expressément qu'une structure qui ne satisfait encore qu'à ces axiomes n'est pas encore soumise à la catégorie du nombre, (*spécialement si celui-ci doit être transfini*). Une structure libre comme celle dont nous sommes partis n'est encore ni dénombrable, ni « de la puissance du continu » etc. et il en est de même des structures qui ne sont qu'ordonnées. La distinction en structures dénombrables ou non-

dénombrables, par exemple, exige l'intervention d'axiomes d'un autre genre. Cette remarque a son importance, car elle exclut l'affirmation que « tout ensemble » possède un nombre cardinal, fini ou transfini ». Ce nombre est, au contraire, une propriété qu'il faut encore porter dans les ensembles ordonnés, par de nouveaux axiomes restrictifs. Nous aurons encore l'occasion d'y revenir. Pour l'instant, nous allons exposer en détail au paragraphe suivant, comment les choses se présentent pour l'ensemble réalisé dans la suite des nombres entiers.

Remarquons pour finir qu'il y a deux façons opposées d'ordonner une structure en choisissant arbitrairement, entre deux éléments quelconques, l'une des deux liaisons présentes dans la structure libre dont nous sommes partis.

5. L'ensemble des nombres entiers. Les axiomes du nombre

Le titre de ce paragraphe exige une mise au point immédiate. Les nombres entiers ne sont en effet pas des éléments logiques purs. La suite de ces nombres ne fournira donc pas une réalisation absolument adéquate d'un ensemble lié. Mais on peut apercevoir dans les relations entre nombres le dessin d'un réseau de liaisons logiques. Lorsque nous parlerons de l'ensemble des nombres entiers, cela voudra donc dire que nous allons regarder cette collection comme la réalisation d'une certaine structure purement logique, le mot de réalisation ayant le sens expliqué au § 2. Cette manière de faire n'est en aucune mesure nouvelle en mathématiques ; elle s'éclaircit parfaitement si l'on cherche ce qui lui correspond dans l'axiomatique élémentaire. Le parallélisme est aussi étroit que possible : nous allons faire de la logique pure sur l'ensemble des nombres entiers comme on fait de la géométrie sur une figure.

C'est dans le même esprit que nous parlerons des axiomes de l'entier. Dans le cadre que nous traçons, la signification de ces axiomes s'éclaire très vivement. Ce sont les restrictions à apporter aux suites simplement ordonnées pour que celles-ci deviennent « numérotées ». La voie à suivre est ainsi toute tracée. La première chose à faire est d'exiger l'existence d'un *premier élément*. C'est d'ailleurs également l'objet du premier axiome du système de Peano que nous allons prendre comme terme de comparaison. En voici les 5 axiomes :

Axiome 1. Zéro est un numéro.

Axiome 2. Si a est un numéro, $a +$, c'est-à-dire « le suivant », « le successif », est aussi un numéro.

Axiome 3. Le principe d'induction sous la forme suivante : Si s est une classe comprenant le zéro et si le successif d'un numéro quelconque x qui fait partie de s , en fait également partie, s contient tous les numéros.

Axiome 4. Si les successifs de deux numéros sont égaux, ces numéros le sont aussi.

Axiome 5. Le successif d'un numéro ne peut être zéro.

Il est clair que ces axiomes doivent être remaniés pour prendre place dans notre théorie. Commençons par choisir un élément que nous nommerons 1, et demandons-nous à quelles conditions il doit satisfaire pour être le premier dans une structure. Choisissons un autre élément, l'élément 2, et choisissons encore l'ordre dans lequel la liaison (1, 2) est à conserver. Considérons ensuite deux éléments x et y , et supposons que l'on ait l'ordre $x, 1, y$. Nous supprimons les liaisons $(x, 1)$ et (x, y) , par exemple en décrétant leur incompatibilité avec (1, 2). La structure est maintenant partagée en deux structures partielles indépendantes : celle des éléments qui précèdent 1, et celle qui contient 1 et les éléments suivants. Cette construction restrictive donne lieu à l'axiome suivant qui vient prendre la place du premier axiome de Peano. (Nous nommons d'ailleurs structure ou suite dénombrée la suite réalisée par les nombres entiers.)

Axiome N_1 . La suite dénombrée possède un premier élément.

Le deuxième axiome aura pour objet de faire passer de la structure simplement ordonnée à la suite ordonnée, c'est-à-dire à la structure dans laquelle chaque élément possède un élément déterminé qui le suive immédiatement.

Jusqu'ici nous avons souvent raisonné comme si tous les éléments de la structure et toutes les liaisons étaient donnés d'avance. Il n'y avait pas d'inconvénient à le faire, mais qu'il soit bien entendu qu'une structure ne peut être décrite dans son devenir que par la façon dont elle s'engendre. Et une description de ce genre ne peut qu'indiquer les procédés selon lesquels « *aux éléments déjà existants viendront s'adjoindre encore de nouveaux éléments* ». C'est naturellement ici que doit intervenir la notion de la définition prédicative de Poincaré : Le processus de la construction doit être prédicatif en ce sens que l'adjonction de nouveaux éléments ne doit remettre en question aucune des liaisons supposées déjà établies entre les éléments supposés déjà existants.

Voyons maintenant sous cet angle le second axiome de notre suite à définir :

Axiome N_2 . Les éléments qui suivent un élément a quelconque, forment une structure qui possède elle-même un premier élément $a+$, différent de a .

Supposons donc que tous les éléments existants déjà possèdent la propriété requise, c'est-à-dire que les liaisons entre eux satisfassent aux incompatibilités axiomatiques. Soit maintenant b un nouvel élément. En exigeant certaines incompatibilités, nous pourrions l'introduire entre a et $a+$, par exemple. Par d'autres incompatibilités nous pourrions exiger que b soit le premier des éléments qui viennent après a , et ainsi de suite.

Le troisième axiome de Peano contient la notion de classe empruntée à la logique. Nous pourrions introduire cette notion sous la forme d'une structure partielle par la définition suivante :

Une structure partielle qui ne se décompose pas en deux structures partielles indépendantes est une classe dans la structure originelle. On pourra dire aussi que les éléments d'une classe possèdent un attribut déterminé.

Avec cette définition on pourrait conserver le troisième axiome en question. Mais il nous paraît préférable de faire appel à des axiomes de la même nature que le précédent. Le plus simple sera d'introduire aussi la notion de dernier élément.

Axiome N_3 . La classe des éléments qui précèdent a possède un dernier élément $a-$.

La structure est maintenant ordonnée dans les deux sens. Il suffit d'exiger finalement que la classe des éléments qui viennent avant un a quelconque soit finie. Ceci peut faire l'objet du dernier axiome que voici :

Axiome N_4 . Toute structure partielle contenue dans la classe des éléments qui précèdent a possède aussi un dernier terme ⁵⁾.

En résumé, nous définissons donc la suite dénombrée comme étant la structure libre qui satisfait aux axiomes (O_1, O_2, O_3) de l'ordre et (N_1, N_2, N_3, N_4) du nombre.

Sur la base de ces axiomes et des notions fondamentales explicitement introduites, en particulier des notions du *premier* et du *dernier*, le principe d'induction est maintenant démontrable.

On peut tout d'abord l'énoncer sous la forme suivante :

Toute structure partielle de la suite dénombrée qui commence par 1 et qui contient, en même temps qu'un élément a toujours le suivant $a+$, est identique à la suite dénombrée.

L'identité de deux structures infinies est une chose à définir : Deux

⁵⁾ Si l'on étendait la notion de structure partielle (impropre) à la structure originelle elle-même, l'axiome N_4 conduirait naturellement N_5 .

structures infinies sont identiques si elles ont été construites par une application identique des mêmes principes et si elles satisfont aux mêmes axiomes restrictifs.

Il va de soi que dans ce cas tout élément qu'on pourra attribuer à l'une pourra l'être aussi à l'autre.

Dans notre cas, la suite partielle satisfait d'elle-même aux axiomes N_1 et N_2 , et naturellement aussi à l'axiome N_4 , puisqu'une structure partielle d'une structure partielle est encore une structure partielle. Quant aux éléments qui précèdent un élément a dans la structure partielle, ils le précèdent aussi dans la suite primitive et possèdent un dernier terme d'après N_3 ou N_4 .

Enfin, il est clair que s'il ne s'était agi, dans ce qui précède, que de parvenir le plus rapidement et le plus simplement possible à la structure dénombrée, il aurait suffi de faire appel aux principes d'extension et de libre liaison.

6. La relation logique

Jusqu'ici les liaisons ne sont entrées en rapport les unes avec les autres que par leur incompatibilité éventuelle. Les seules propriétés qu'elles peuvent posséder, c'est donc d'exister ou de ne pas exister simultanément (ou aussi de toucher des éléments communs). Mais nous n'avons encore introduit aucun principe qui permette de décider quand deux liaisons, posées tout d'abord comme différentes, peuvent être regardées comme identiques ou équivalentes. Pour y parvenir, nous allons commencer par définir la *projection d'une structure sur elle-même*.

Considérons tout d'abord, par exemple, la structure dénombrée qui ne conserve que les liaisons distinguées qui joignent un élément a au suivant $a +$. On obtient une projection de la suite sur elle-même en faisant correspondre l'élément 1 à un élément quelconque x , puis l'élément $1 + = 2$ à l'élément $x +$, et ainsi de suite. En général :

Définition : Une projection d'une structure sur elle-même est une correspondance univoque et réciproque de la structure avec tout ou partie d'elle-même, la correspondance embrassant à la fois les éléments et les liaisons.

Les projections qu'une structure admet forment naturellement un quasi-groupe ⁶⁾.

Cette définition devra être encore un peu étendue dans un instant. Dans le cas qui nous occupe, ce quasi-groupe est tel que la liaison (1, 2) peut

⁶⁾ Le quasi-groupe satisfait aux axiomes de groupe excepté à celui des éléments inverses.

être amenée sur toute liaison $(a, a+)$. Si maintenant nous regardons cette *applicabilité comme la caractéristique de l'équivalence*, toutes les liaisons de la suite deviennent équivalentes à $(1, 2)$ et équivalentes entre elles.

C'est ce système de liaisons toutes équivalentes que nous définissons comme formant une relation logique ; celle de l'élément quelconque a à $a+$. Dès ici, le symbole $(a, a+)$ peut être considéré comme celui de la *relation* elle-même.

Ce premier exemple est un peu particulier, du fait que a détermine $a+$. Examinons encore la structure dénombrée où l'on a réintroduit les liaisons entre les éléments qui ne se suivent pas immédiatement. Pour décrire le quasi-groupe des projections qui interviennent ici, il est commode d'imaginer que la structure est donnée en deux exemplaires superposés. L'élément a de la liaison (a, b) sera supposé choisi dans le premier exemplaire, et b dans le second.

On peut alors « repousser » le second exemplaire « vers la droite » comme dans le premier exemple, de façon que 2 vienne en y . La liaison $(1, 2)$ tombe sur $(1, y)$, le premier élément étant resté fixe.

On peut ensuite « repousser » aussi le premier exemplaire « vers la droite » de façon que 1 vienne en x , x devant d'ailleurs précéder y . La liaison $(1, 2)$ est maintenant équivalente à la liaison (x, y) . Cette dernière est d'ailleurs quelconque dans la structure. Celle-ci est donc uniquement formée de liaisons équivalentes entre elles. Elle représente encore une fois une relation logique, qu'on écrira

$$(x, y) = x < y.$$

Le fait que la structure correspondant à la première relation est comprise dans cette dernière signifie que celle-ci est plus générale.

Examinons encore la relation réalisée arithmétiquement par deux nombres, a et b , et leur somme c . Nous l'écrirons

$$(a, b, c) \equiv a + b = c.$$

Pour obtenir cette relation, en d'autres termes pour obtenir les *axiomes de l'addition*, nous allons précisément nous servir du quasi-groupe qui sert à identifier les liaisons. Nous commençons donc par établir (principe de libre liaison) une liaison entre les éléments 1,1 et 2. Pour plus de clarté, on pourra supposer avoir 3 exemplaires de la structure dénombrée à sa disposition. Le quasi-groupe sera ensuite engendré par les opérations suivantes :

Lorsque l'exemplaire a de la structure est repoussé d'un pas vers la droite, il en est de même de l'exemplaire c .

Lorsque l'exemplaire b est repoussé d'un pas vers la droite, il en est encore une fois de même de l'exemplaire c .

Les deux éléments 1 des exemplaires a et b peuvent venir se placer en des endroits quelconques, tandis que le troisième élément de la relation est déterminé.

Les axiomes correspondant à cette façon de faire sont maintenant :

$$\begin{array}{ll} \text{Axiome } S_1 & 1 + 1 = 1 + (= 2) \\ \text{Axiome } S_2 & (a + 1) + b = (a + b) + 1 \\ \text{Axiome } S_3 & a + (b + 1) = (a + b) + 1. \end{array}$$

On pourrait naturellement retrouver des circonstances analogues dans les structures denses homogènes comme celles qui sont réalisées par la suite des nombres rationnels, par exemple, ou par le continu linéaire. Mais il n'est pas possible d'opérer de la même façon si la structure étudiée n'est pas homogène (par exemple une suite dénombrée suivie elle-même d'un ensemble dense). Mais ce n'est peut-être pas un défaut de notre point de vue, car il n'y a en réalité pas grand'chose de commun entre le fait de se suivre dans une suite dénombrée et dans une suite dense. L'identification exigerait « l'enrobage » dans une suite homogène, ce qui est d'ailleurs toujours possible.

Comme conclusion, nous voulons formuler le principe d'identification dont nous nous sommes servi et la définition de la relation logique. Il nous faut tout d'abord étendre la notion de projection d'une structure de liaisons à un système d'autant de structures superposées qu'il y a de termes dans les liaisons considérées. On peut ensuite énoncer le :

Principe des liaisons équivalentes. Deux liaisons qui, lors d'une projection (généralisée) d'une structure sur elle-même, viennent à coïncider, peuvent être regardées comme équivalentes.

La définition de la relation est maintenant la suivante :

Une relation logique est une structure de liaisons toutes équivalentes. L'existence d'une relation logique est donc liée à l'existence d'un quasi-groupe transitif de projections d'une structure sur elle-même, la transitivité devant être réalisée aussi bien pour les liaisons que pour les éléments.

7. Comparaison avec l'axiomatique actuelle

Nous allons maintenant examiner comment se présentent les axiomes de M. Zermelo dans le cadre que nous avons tracé.

La première remarque à faire concerne la notion même de l'élément : dans la théorie de M. Zermelo, l'élément doit être un ensemble, la relation

de l'ensemble-élément à l'ensemble-totalité restant dans l'état d'indétermination que Poincaré avait déjà souligné ; dans notre esquisse, au contraire, l'élément est une entité sans signification préalable et qui par exemple ne devient nombre que par sa présence dans la suite dénombrée. En revanche, la notion d'appartenance de l'élément à l'ensemble est directement fondée sur la notion intuitive de collection.

Le principe de contraction permet toutefois d'imaginer qu'une structure puisse jouer le rôle d'un élément. Mais la différence avec l'ensemble-élément n'en reste pas moins essentielle : si l'on remplace par contraction une structure par un seul élément, celui-ci ne garde plus trace du fait que sa place était occupée tout à l'heure par une structure ; sa signification est à nouveau déterminée par sa seule position dans la structure restante.

Il en résulte que les deux relations $a \varepsilon b$ et $a < b$ ne sont pas essentiellement différentes.

On voit par là que la notion d'élément logique est bien adéquate à la spéculation mathématique : elle permet de faire abstraction des propriétés « internes » d'une « pluralité » pour n'envisager que ses rapports avec d'autres « pluralités ».

Quant à la relation $a = b$ elle sera pour nous l'identité des structures a et b , dont nous avons parlé au paragraphe précédent. (Nous dirons aussi qu'elles sont superposables.)

Dans ces conditions, l'axiome 1 de M. Zermelo s'énonce comme suit :

Axiome 1. Si deux structures a et b sont superposables et si a est structure partielle d'une troisième structure A , il en est de même de b qui peut prendre la place de a .

Cet axiome est naturellement vérifié.

L'axiome 2 est également immédiat :

Axiome 2. Les éléments de deux structures différentes peuvent être réunis en une seule structure.

Il suffit en effet d'introduire des liaisons quelconques entre les éléments de l'une et les éléments de l'autre. Si par exemple, les deux structures sont ordonnées, on peut à volonté intercaler les éléments de l'une entre les éléments de l'autre.

L'axiome 3 n'est qu'une autre forme de la réciproque du principe de contraction.

Axiome 3. Si A est une structure de structures, c'est encore une structure formée à l'aide des éléments des structures partielles.

Nous reviendrons dans un instant sur les axiomes 4 et 5.

L'axiome 6 peut être envisagé de deux manières : Il peut être regardé tout d'abord comme une autre forme du principe de contraction. Mais

s'il ne s'agit pas seulement d'affirmer l'existence d'un seul ensemble « choisi », le rôle de cet axiome est en tous points semblable à celui de l'axiome 3. Il décrira comment on peut former un ensemble comprenant toutes les possibilités de choix,

Considérons le cas le plus simple de deux structures. Imaginons que a et b soient deux éléments de la première entre lesquels il existe n liaisons $(a, b)_1, (a, b)_2, \dots$, que de même a' et b' soient deux éléments pris dans la seconde entre lesquels existent les n' liaisons $(a', b')_1, (a', b')_2, \dots$. Correspondant aux deux paires a et a' , et b et b' , nous imaginons deux éléments logiques A et B entre lesquels il y ait à la fois les liaisons de la première et de la seconde structure (c'est-à-dire, puisque les liaisons n'ont aucun caractère spécifique, $n + n'$ liaisons différentes). Nous obtenons ainsi une nouvelle structure qu'on peut appeler une structure-produit des deux premières.

Cette façon de faire peut s'étendre au cas où l'on a à « choisir » les éléments dans une suite S infinie (et liée) de structures. On pourra établir dans ce cas une infinité de liaisons entre les éléments A et B . Ces liaisons formeront elles-mêmes une structure, et on les ordonnera par exemple par paquets de $n, n', n'' \dots$, liés entre eux de la même façon que les structures de la suite S .

Si par exemple, on devait construire par ce procédé le produit de la suite dénombrée par elle-même, on pourrait en représenter la structure comme dans la fig. 2. Les termes qui se suivent sur une verticale ou sur

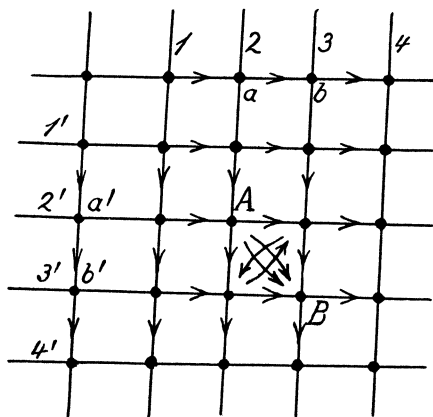


Fig. 2.

une horizontale sont liés une fois, les diagonales des carrés portent deux liaisons. On voit aussi quelles liaisons il faudrait supprimer pour que cette structure redevienne une suite dénombrée.

Nous n'avons tenu compte que des liaisons à deux termes pour « ordonner » la structure-produit. On peut traiter les liaisons à plus de deux termes de façon analogue.

L'axiome 4 est encore du même genre que le précédent. S'il ne s'agit que d'indiquer une structure qui le vérifie, la chose ne présente aucune difficulté. Par exemple on peut considérer deux sous-structures comme liées s'il existe au moins une liaison entre un élément de l'une et un élément de l'autre, ou un élément de l'une et deux éléments de l'autre, etc.

On pourrait aussi, selon le procédé de Cantor, ramener la construction d'une structure satisfaisant à cet axiome à celle d'un produit « répété ».

L'axiome 7 est remplacé par la construction axiomatique de la suite dénombrée.

Reste enfin l'axiome 5, dans lequel la notion d'attribut défini qui intervient « du dehors » et dont le domaine de signification n'a pas été délimité, est généralement considérée comme un point faible. Dans l'esquisse que nous avons tracée, nous avons placé la structure partielle avant l'attribut. L'axiome 7 est donc vérifié identiquement. Mais le rôle qu'il joue dans la construction axiomatique, ce sont les procédés d'après lesquels les structures partielles peuvent être déterminées qui en tiennent lieu.

Ces procédés peuvent être déjà distingués dans l'exemple de la suite dénombrée. Ainsi, les nombres pairs y déterminent une structure partielle parce que la notion de nombre pair peut être définie à partir des principes fondamentaux, et spécialement des axiomes restrictifs de la suite dénombrée. Il en serait de même pour toute autre propriété arithmétique. En d'autres termes et en général :

Si l'on entend par attribut bien défini une propriété de certains éléments d'une structure qui peut être construite sur la base des principes généraux et des axiomes relatifs spécialement à cette structure — et dont la validité ou la non-validité peut être constatée sans ambiguïté —, alors notre définition de la structure partielle est équivalente à l'axiome 5.

La différence est que maintenant le domaine de signification du mot attribut est délimité.

En résumé, on voit qu'il suffit de quelques modifications, dont les unes sont assez légères, mais dont la dernière est essentielle, pour que les axiomes de M. Zermelo viennent prendre place dans notre théorie. Il est vrai qu'ils s'y trouvent maintenant plutôt comme conséquences que comme axiomes.

8. Les antinomies

Nous allons maintenant faire voir que les antinomies bien connues sont écartées tout naturellement.

Commençons par le paradoxe de l'ensemble de tous les ensembles qui ne se contiennent pas eux-mêmes comme éléments.

La première remarque à faire — remarque de principe — c'est que les ensembles qui ne se contiennent pas eux-mêmes comme élément ne sont pas des éléments logiques. Ce sont des objets de la pensée caractérisés par une propriété à priori, et non par leur position dans une structure logique. Leur ensemble, dans notre théorie, n'existe pas.

Tout au plus pourrait-on former une structure qui soit ensuite réalisée par « tous » les ensembles dont nous parlons. Il faudrait trouver un principe ordinateur, selon lequel ils pourraient être mis en liaison. Il n'y a aucune raison de croire qu'un principe de ce genre soit contenu dans la notion seule « d'ensemble qui ne se contient pas lui-même comme élément ».

Ainsi, nous ne sommes pas en mesure de construire une structure correspondant à l'ensemble antinomique. Et même si nous le pouvions, le paradoxe pourrait encore être évité.

Supposons, en effet, que Σ soit une structure de ce genre. C'est dire qu'elle a été formée d'après l'axiome 3 du paragraphe précédent, ou d'après la réciproque du principe de contraction, à l'aide des structures qui ne sont pas applicables sur une partie d'elles-mêmes. Posons la question qui devrait amener le paradoxe : Σ se contient-elle elle-même comme structure partielle ?

Supposons que non : Alors, d'après la façon dont elle a été construite, elle devrait se contenir elle-même, ce qui est la contradiction bien connue.

Supposons le contraire : elle contiendrait alors une structure partielle qui se contiendrait elle-même.

Si nous devons admettre, comme dans la théorie habituelle, que cette structure ne peut être qu'un élément de Σ , nous ne pourrions éviter la seconde contradiction, et le paradoxe serait là, encore une fois. Mais la chose n'est aucunement nécessaire : il peut fort bien exister une structure partielle qui soit différente de celles qui ont servi à construire l'ensemble Σ . L'antinomie a donc disparu.

Passons à l'antinomie du plus grand cardinal, que nous considérons comme plus profonde.

Le mécanisme qui l'engendre est le suivant :

Tout ensemble E possède un nombre cardinal déterminé. Le cardinal de l'ensemble $((E))$ des sous-ensembles de E est plus grand que celui de

ce dernier. L'ensemble T de tous les ensembles ne peut avoir que le plus grand cardinal, et pourtant l'ensemble $((T))$ en a un plus grand encore.

Que pouvons nous objecter à cette façon de faire ? Il y a tout d'abord la remarque de principe qui peut être invoquée dans tous les cas semblables. C'est que les objets « ensembles » ne sont pas des éléments logiques, mais des objets doués à priori de certaines propriétés. L'ensemble de tous les ensembles est donc une construction mentale tout à fait étrangère à la théorie que nous avons développée. Tout au plus pourrions-nous parler :

soit d'une structure dont les éléments puissent être réalisés par les objets « ensembles », et les liaisons par certaines façons de ces derniers d'entrer en rapport les uns avec les autres,

soit d'une structure dont elle-même et toutes les autres structures soient des structures partielles.

Dans le premier cas, il pourrait arriver que la structure ainsi abstraite fût antinomique. Mais ce fait ne toucherait en rien notre théorie, car nous ne reconnaissons pas comme légitimes toutes les structures abstraites de tels ou tels ensembles d'objets aux propriétés les plus imprévues, mais seulement celles qui peuvent être construites sur la base de nos principes fondamentaux.

Pour reprendre la comparaison avec la géométrie élémentaire, nous sommes aussi peu dans l'obligation de reconnaître les structures abstraites de collections d'objets, que la géométrie euclidienne du plan l'est d'accepter les résultats de mesures faites sur une sphère ou sur toute autre surface. La géométrie et notre théorie échappent aux dangers que représentent pour l'une la formation de collections aux caractères antinomiques (comme le catalogue des catalogues qui ne se mentionnent pas eux-mêmes, ou comme la classe des adjectifs imprédicables de Russel), et pour l'autre les mesures sur les objets physiques, parce que l'une et l'autre ont été placées par l'axiomatisation en dehors de la « sphère de leurs réalisations respectives ».

Ceci dit « pour le principe », il nous faut reconnaître que, dans le cas qui nous occupe, nous ne sommes pas encore hors de peine, parce que le paradoxe semble renaître à propos de la seconde éventualité, à propos de la « structure de toutes les structures ». Nous pourrions d'ailleurs nous borner à considérer des structures libres, par exemple des structures où il y a exactement une liaison entre deux éléments quelconques. Supposons que l'on puisse former la somme de toutes les structures de ce genre. Il faudrait premièrement que celles-ci correspondissent aux élé-

ments d'une structure. On pourrait déjà s'arrêter ici, en remarquant que nos structures libres ne sont pas liées à priori par leur définition. Mais à défaut des liaisons imposées par les définitions, on peut en imaginer plus ou moins librement. Par exemple on pourrait dire que le fait d'envisager deux structures simultanément constitue déjà, dans la sphère des réalisations, l'équivalent d'une liaison. On est donc conduit à envisager les structures en question comme correspondant aux éléments d'une structure libre du même genre que celles dont nous nous occupons. *Acceptons* cette façon de faire.

Une fois toutes ces hypothèses admises, on pourrait former la structure-somme S en établissant entre les éléments de deux structures à additionner exactement une liaison allant de chaque élément à chaque élément. La structure-somme est alors encore une structure libre du même genre que les précédentes, et il n'en peut exister d'autres qui n'y soient pas contenues comme structure partielle.

Et maintenant nous pouvons former selon l'axiome 4 une structure dont les éléments correspondent aux structures partielles de S . En établissant une seule liaison entre les éléments représentants de deux ensembles dont au moins deux éléments sont liés, nous engendrerons encore une fois un ensemble libre $((S))$ du même genre. Tous les termes du paradoxe sont à nouveau là.

Dès lors il est clair que le nœud de la question doit se trouver dans la démonstration qui, dans la théorie de Cantor, assigne à l'ensemble de *tous* les sous-ensembles d'un ensemble quelconque un nombre cardinal supérieur au nombre cardinal de ce dernier. Pour le dénouer, il nous faut commencer par quelques remarques sur l'emploi du mot *tous*.

Examinons pour commencer le cas d'une collection finie. Pour une collection assez peu nombreuse, le mot *tous* a une signification intuitive immédiate. Dire que, dans ce cas, tous les objets sont donnés, cela veut dire qu'on peut « entrer effectivement » en possession de chacun d'eux. Mais la chose change totalement d'aspect si la collection devient suffisamment nombreuse pour qu'il ne soit plus possible d'en examiner tous les objets l'un après l'autre jusqu'au dernier. Tout comme le mot infini, le mot fini a dans ce cas une signification en partie conventionnelle. En disant que tous les objets sont donnés, on demande qu'on *admette la possibilité* de les énumérer un à un jusqu'à épuisement. Pour reprendre une expression fréquemment employée ces derniers temps dans la discussion des fondements des mathématiques, c'est un chèque en blanc sur tout

objet compris entre le premier et le dernier, mais dont on ne peut assurer d'avance si on pourra le remplir effectivement. Le sens du mot tous est alors relatif à la convention qui veut que ce chèque ne perde jamais sa valeur, c'est-à-dire aux règles axiomatiquement acceptées qui régissent l'emploi des nombres.

Quelle est la différence avec le cas de la suite dénombrée ? Le sens des mots : « Tous les éléments de la suite ont telle ou telle propriété » est ici tout engagé dans le principe d'induction, (qui est soit un axiome, comme dans le système de Peano, soit une conséquence des axiomes de la suite dénombrée) et ne porte pas plus loin que celui-ci. Dire que tous les éléments sont donnés, c'est dire simplement que l'on peut toujours passer au suivant, et c'est *accepter* que cela soit suffisant bien qu'il n'y ait pas de dernier. Par définition, la suite dénombrée est précisément celle pour laquelle cette convention (axiomatique) est valable.

Il ne sert de rien d'objecter : « Il est absurde de dire que tous les éléments d'une suite infinie, même dénombrée, ont été examinés, puisqu'il est dans la nature de l'infini de ne pas être épuisé. » En parlant ainsi, on oppose simplement l'un à l'autre le sens de tous dans une suite finie et l'autre sens de ce mot dans une suite dénombrée, ces deux significations étant en effet irréductibles l'une à l'autre.

Il n'y a pas de difficulté à raisonner de la même façon sur l'ensemble réalisé par les nombres rationnels ou par le continu des nombres réels. On postule tout d'abord les axiomes de l'ordre, puis certains axiomes restrictifs plus ou moins analogues aux axiomes du nombre qui fixent les conventions suivant lesquelles le mot tous peut être employé.

En résumé, la signification du mot tous à l'intérieur d'une structure, est relative aux axiomes que cette dernière doit (par définition) vérifier : c'est un *caractère intrinsèque* qui ne peut être sans plus reporté d'une structure sur une autre.

Ces constatations s'appliquent immédiatement à la notion de l'équivalence. Si deux structures doivent être telles que tout élément de l'une corresponde à un élément de l'autre et réciproquement, les mots : « tout élément » ont un sens fixé la première fois par les axiomes de la première structure, la seconde fois par ceux de la seconde.

Il est un cas où cette équivalence doit être admise : c'est celui où les deux structures satisfont aux mêmes axiomes. Rien ne peut dans ce cas les distinguer l'une de l'autre.

Ce sera l'*équivalence au sens restreint*. On pourra l'élargir ensuite par les définitions suivantes :

Deux structures sont équivalentes si l'une s'obtient à partir de l'autre par suppression de certaines liaisons seulement.

Deux structures sont équivalentes si ce sont deux sommes ou deux produits formés de « façon équivalente » à l'aide de structures équivalentes.

Ceci posé, envisageons deux structures libres quelconques du même genre que précédemment et supposons qu'on n'ait accepté aucun axiome restrictif (si ce n'est peut-être, pour toutes deux, les axiomes de l'ordre). D'après ce que nous venons de dire, nous n'avons aucun moyen de constater en quoi elles pourraient différer. Au contraire, d'après la définition de la structure libre, l'une quelconque peut toujours être envisagée comme faisant partie de l'autre : il faut donc les considérer comme équivalentes. Ainsi : *Deux structures libres quelconques (du même type) sont à considérer comme équivalentes.*

Ceci nous ramène aux deux structures S et $((S))$. Comme leur équivalence se confirme, c'est bien dans la démonstration indiquée qu'il nous faut rechercher la cause du paradoxe. Rappelons-en les points essentiels.

Les éléments de $((S))$ représentent les structures partielles de S . Parmi celles-ci se trouvent aussi celles qui ne comprennent qu'un élément : S est donc structure partielle de $((S))$. Nous allons partager les éléments de $((S))$ en deux classes :

la première comprendra les éléments qui sont compris dans la structure partielle à laquelle ils correspondent,

la seconde comprendra les éléments qui n'y sont pas compris.

On montre ensuite que cette seconde classe ne peut correspondre à aucun élément de $((S))$ et que, par conséquent, elle ne peut être une sous-structure de S .

Le point essentiel du raisonnement est donc celui où l'on applique le principe suivant : *De tout élément on peut décider sans ambiguïté s'il appartient ou non à un ensemble déterminé.* Ce principe forme en quelque sorte le fondement de la définition de l'ensemble dans la théorie habituelle, où les éléments d'un ensemble sont contenus dans celui-ci en tant qu'individualités possédant une propriété distinctive, sur la foi de laquelle l'élément est attribué ou non à l'ensemble.

Dans notre théorie, ce principe est faux. Un élément logique n'ayant aucune propriété a priori, il n'y a aucun sens à demander s'il fait ou non partie d'une structure. La chose est bien visible sur l'exemple des structures libres. Les principes de libre extension et de libre liaison permettent d'attribuer un élément logique quelconque à une structure libre quel-

conque. (C'est là justement la raison de la dénomination : structure libre.)

Dans notre théorie, la notion de nombre cardinal ne pourra être étendue aux structures libres : c'est ce que nous voulions déjà dire au § 4 par les mots : Les structures libres (ou seulement ordonnées) ne tombent pas encore sous la catégorie du nombre.

L'antinomie a maintenant complètement disparu.

Il est inutile d'insister encore sur la structure somme de toutes les structures, sur laquelle on ne raisonnerait pas différemment.

9. Conclusion

L'analyse du paradoxe du plus grand cardinal a montré que la source des difficultés, dans la théorie de Cantor, est en premier lieu la notion de l'objet susceptible de posséder, a priori et de par lui-même, telle ou telle propriété ; en second lieu la définition de l'ensemble comme collection d'objets ayant une propriété distinctive. C'est là ce qu'on pourrait appeler une définition *statique* de l'ensemble : les objets-éléments sont *tous* déterminés par avance et il faut admettre que, d'un objet quelconque, on saura décider s'il appartient ou non à un ensemble déterminé, mais quelconque.

Bien qu'elles réduisent d'emblée l'objet-élément à n'être plus qu'un ensemble, la théorie de M. Zermelo et les théories plus récentes qui en découlent restent complètement sur le même terrain.

A cette façon statique d'envisager un ensemble, on peut opposer la notion de la « totalité en devenir ». Toujours à propos des axiomes de M. Zermelo, Poincaré écrivait déjà : « Quand je parle de tous les points de l'espace, je veux dire tous les points dont les coordonnées sont exprimées par des nombres rationnels, ou par des nombres algébriques.... ou de toute autre manière que l'on pourra inventer. Et c'est ce *l'on pourra* qui est l'infini. »

C'est à cet ordre d'idées qu'appartient la notion de *limite*, et spécialement celle de *suite en devenir* (*werdende Folge*) de M. Weyl.

On peut dire que la plupart des difficultés de la théorie des ensembles provenaient de la contradiction « latente » entre la définition statique, et la nécessité de considérer l'infini comme étant en état de perpétuelle extension.

Pour y échapper nous avons dégagé de la *notion intuitive d'objet* la notion d'*objet purement logique*, dont les propriétés a priori se réduisent à être soit identiques, soit différents entre eux, et dont les autres propriétés lui seront conférées par une structure en devenir.

Un autre exemple typique d'objet en devenir est celui de la structure libre.

Nous croyons utile d'ajouter un mot encore sur la différence que nous avons faite entre les principes et les axiomes.

Les principes formulent les règles selon lesquelles une structure *quelconque* peut être traitée. Elles viennent compléter les règles simples, valables pour une collection finie, quant au partage de celle-ci en deux ou plusieurs collections, quant à la réunion de deux collections en une seule, quant aux possibilités de choix ou de permutations, etc., règles qu'on peut considérer comme formant la partie essentielle de la *logique ordinaire*. On pourrait donc dire que les principes ont pour objet de formuler les règles de la logique de l'infini (en dehors de la question de savoir si nous les avons déjà formulées toutes).

Les axiomes, au contraire, sont des décrets restrictifs portés sur des structures en devenir, et dont l'extension indéfinie est ainsi soumise à certaines prescriptions caractéristiques.

Ceci dit, nous pouvons énoncer de la façon suivante le but que nous nous sommes proposé d'atteindre dans ce travail : Il s'agissait de montrer

a) qu'en acceptant les règles de la logique ordinaire pour autant qu'il s'agit de structures finies,

b) qu'en les complétant par certains principes relatifs aux structures infinies,

c) et en n'opérant systématiquement que sur les « objets logiques » dont le sens et le rôle doivent être précisés par une axiomatisation *sui generis* (et d'ailleurs étroitement analogue en son principe à celle qui permet de constituer la géométrie élémentaire en science, rationnelle),

on élimine de façon toute naturelle les antinomies qui se présentaient jusqu'ici, spécialement dans la théorie des ensembles et dans la logique de l'infini.

(Reçu le 6 avril 1932)

La propagation des ondes et les caustiques

par J. HADAMARD, Paris

Le présent travail fait suite à ceux que nous avons publiés sur le principe de Huygens — ou, plus exactement, sur la „majeure de Huygens“ — pour les équations à trois variables indépendantes, et particulièrement ¹⁾ à ceux qui ont été insérés au *Journal de Mathématiques* en 1929 et aux *Acta Mathematica* en 1930 ²⁾, lesquels seront respectivement désignés par les abréviations *JM* et *A₂*.

Le premier de ces deux Mémoires était consacré au théorème d'addition intégral qui résulte de la majeure de Huygens, dans l'hypothèse où, entre leurs points extrêmes, les ondes considérées ne présentent aucune singularité, autrement dit, n'ont pas de points communs avec une caustique. Nous nous proposons, dans ce qui va suivre, d'examiner ce qui se passe lorsqu'on renonce à cette hypothèse.

L'onde rétrograde issue d'un point *o* et relative à une équation linéaire du second ordre hyperbolique

$$(E) \quad F(u) = \sum_{i,j=1,2,3} A^{ij} \frac{\partial^2 u}{\partial x^i \partial x^j} + \sum_{i=1,2,3} B^i \frac{\partial u}{\partial x^i} + Cu = \varphi$$

ou plutôt la nappe conique I'_0 qui la représente dans l'espace à trois dimensions sera donc, comme précédemment, coupée par deux surfaces successives S_1, S_2 (fig. 1) ayant toutes deux, en chacun de leurs points, une orientation d'espace et qui délimiteront avec elle deux régions bornées de l'espace. Dans la première de ces deux régions, on admettra encore, comme nous l'avons fait jusqu'ici, que tout point peut être joint au point *o* par une géodésique bien déterminée de l'élément linéaire

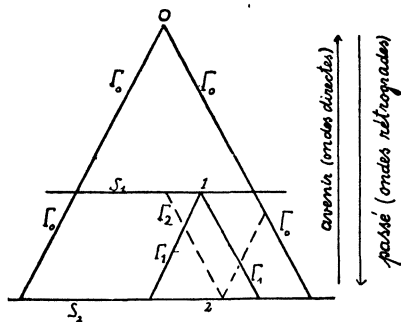


Fig. 1.

$$H(dx^1, dx^2, dx^3; x^1, x^2, x^3) = \sum H_{ij} dx^i dx^j$$

¹⁾ Voir aussi Bulletin Soc. Math. Fr., tome LII (1924), p. 141.

²⁾ Journal de Math., t. VIII (volume publié en l'honneur de MM. Appell et Picard, 1929), p. 197; Acta Math. t. LIV (volume publié en hommage à Fredholm, 1930) p. 247.

qui correspond à l'équation ³⁾, de sorte que le déterminant fonctionnel désigné par J dans nos Leçons de Yale ⁴⁾ sera différent de zéro dans la région en question, partout ailleurs qu'au sommet du conoïde et que la quantité Γ_0 , carré de la distance géodésique à ce sommet, aura partout une valeur déterminée. C'est ce que nous appellerons l'hypothèse α .

Nous continuerons également à admettre une hypothèse analogue — dite hypothèse β — en ce qui concerne la nappe de conoïde rétrograde issue d'un point quelconque 1 de S_1 , dans toute la région limitée par cette nappe et la surface S_2 et, de même ⁵⁾, pour la nappe directe de conoïde ayant pour sommet un point quelconque 2 de S_2 dans toute la région délimitée par cette nappe et la surface S_1 .

Par contre, nous avons également admis, dans nos travaux antérieurs et, en particulier, dans JM , que les choses se passaient de même pour le conoïde rétrograde de sommet 0, jusques et y compris sa rencontre avec S_2 . C'est cette dernière hypothèse — dite hypothèse γ — qui ne sera plus maintenue dans le travail actuel. Au niveau de la surface S_2 , le déterminant fonctionnel J (toujours formé à partir du point 0) sera susceptible de s'annuler, et la quantité Γ_{02} pourra cesser d'être une fonction bien déterminée et régulière des coordonnées du point 2. Si ces circonstances se produisent à l'intérieur du conoïde Γ_0 , elles ne donnent lieu, comme nous le verrons (Cf. 12), à aucune difficulté ; nous aurons, au contraire, à les étudier spécialement lorsqu'elles intéressent la nappe conoïdale elle-même, représentative de l'onde rétrograde issue de 0, c'est-à-dire lorsque celle-ci ou, plus généralement, une onde incidente ⁶⁾ quelconque (voir plus loin n° 7) présente une caustique qui rencontre la surface S_2 . Nous devons, à cet effet, utiliser l'étude géométrique qui a été faite de ce nouveau cas dans le Mémoire A_2 , Mémoire dont, en conséquence, les résultats serviront de base à ce qui va suivre.

Nous avons même à compléter sur quelques points ces remarques géométriques. Les notations de A_2 et de JM n'étant pas les mêmes, je préfère, pour plus de clarté, commencer par donner, en me plaçant dans la

³⁾ Voir nos Lectures on Cauchy's problem, Cambridge New Hawen, nos 55 à 56. Traduction française, Paris, Hermann, 1932, p. 117 — 119.

⁴⁾ Lectures on Cauchy's problem, Livre II, n° 57bis, p. 122—123 de l'édition française.

⁵⁾ Ces deux parties de l'hypothèse β sont, dans une certaine mesure (mais dans une certaine mesure seulement), équivalentes l'une à l'autre, en ce sens que le déterminant fonctionnel analogue à J , formé à l'aide des deux points 1 et 2, est le même, à un facteur près différent de zéro, quel que soit celui de ces deux points que l'on considère comme point de départ, ainsi qu'il résulte du n° 170 de nos Leçons sur le problème de Cauchy (p. 374—377 de l'édition française).

⁶⁾ Comme nous l'avons dit (JM , note de la p. 220), il s'agit d'une onde incidente rétrograde.

première d'entre elles, les indications complémentaires pour lesquelles elle convient particulièrement.

I.

1. Prenons donc tout d'abord, comme dans A_2 , pour axe des x une bicaractéristique située sur l'onde incidente I_0 ; pour plan des xy , une caractéristique régulière admettant cette bicaractéristique —, de sorte qu'on peut écrire une intégrale complète régulière

$$z = F(x, y, \alpha, \beta)$$

telle que le second membre s'annule identiquement en x, y pour $\alpha = \beta = 0$ —, l'onde incidente I_0 étant la détermination de l'intégrale générale que l'on obtient en prenant α variable et β identiquement nul et, par conséquent, en éliminant α entre les équations $z = F(x, y, \alpha, 0)$ et

$$(1) \quad \frac{\partial F}{\partial \alpha}(x, y, \alpha, 0) = 0.$$

De plus, la bicaractéristique définie par les deux équations précédentes pour $\alpha = 0$, est supposée être l'axe des x .

Une autre caractéristique se raccordant avec les premières tout le long de cet axe s'obtiendra comme enveloppe de la surface

$$z = \Phi(x, y, \alpha) = F[x, y, \alpha, g(\alpha)]$$

obtenue en remplaçant, dans l'intégrale complète, β par la fonction (nulle à l'origine ainsi que sa dérivée)

$$(2) \quad \beta = g(\alpha).$$

Entre deux surfaces de cette espèce, correspondant à deux choix différents de la fonction (2), le contact est du même ordre tout le long de la bicaractéristique et du même ordre que celui du contact entre les deux courbes (2). On peut même préciser d'une manière assez remarquable la manière dont la courbure d^2z/dy^2 dépend de x ou, plutôt, dépend simultanément de x et du choix de la fonction g . Celle-ci interviendra par sa dérivée seconde initiale $g''(0) = \mu$ et l'on aura

$$(3) \quad \frac{d^2 z}{dy^2} = \frac{A}{B + \mu}$$

les quantités A , B dépendant de x , mais non du choix de la fonction g .

En effet, la combinaison des deux formules (7) et (9) du Mémoire A_2 donne, pour $d^2 z/dy^2$, la valeur

$$(3') \quad \frac{d^2 z}{dy^2} = \frac{\partial^2 \Phi}{\partial y^2} - \frac{\left(\frac{\partial^2 \Phi}{\partial \alpha \partial y}\right)^2}{\left(\frac{\partial^2 \Phi}{\partial \alpha^2}\right)}.$$

Ici, le premier terme est nul, puisque, dans notre système de coordonnées actuel, $z = \Phi$ est représentée par un plan. D'autre part, toujours en raison de $g'(0) = 0$, la valeur de $\frac{\partial^2 \Phi}{\partial \alpha \partial y}$ se réduit à $\frac{\partial^2 F}{\partial \alpha \partial y}$, celle de $\frac{\partial^2 \Phi}{\partial \alpha^2}$ à $\frac{\partial^2 F}{\partial \alpha^2} + \frac{\partial F}{\partial \beta} g''$, d'où

$$(3'') \quad \frac{d^2 z}{dy^2} = - \frac{\left(\frac{\partial^2 F}{\partial \alpha \partial y}\right)^2}{\frac{\partial^2 F}{\partial \alpha^2} + \frac{\partial F}{\partial \beta} g''}$$

ce qui donne bien, pour la courbure en question, une expression de la forme (3).

Au reste, le même fait ressort de la théorie générale des caractéristiques. Celle-ci nous apprend ⁷⁾ que pour une équation aux dérivées partielles du premier ordre

$$A(x, y, z, p, q) = 0$$

(qu'il convient à cet effet de remplacer par l'équation

$$r \frac{\partial A}{\partial p} + s \frac{\partial A}{\partial q} + \frac{\partial A}{\partial x} + p \frac{\partial A}{\partial z} = 0$$

que l'on en déduit par une différentiation), la variation de r le long d'une bicaractéristique obéit à une équation de Riccati. En l'espèce, cette

⁷⁾ *Goursat*, Ann. Fac. Sc. Toulouse, 2^{ème} série, t. VIII (1906), p. 427 et suiv., parti, culièrement, n° 6, p. 438-9. Voir notre Cours d'Analyse, t. II, Paris, Hermann, 1930-n° 355, p. 456.

équation admettrait une solution nulle, c'est-à-dire se réduirait à une équation de Bernoulli, dont l'intégrale générale est bien de la forme (3).

Pour α différent de zéro et voisin de zéro, la valeur de d^2z/dy^2 dépendra encore homographiquement de μ : elle sera donnée par la formule (3').

2. Comparons en particulier, à ce point de vue, l'onde incidente Γ_0 avec le conoïde direct ayant pour sommet un point 2 pris sur la courbe de section τ de cette onde par la surface S_2 . Pour ce conoïde, la fonction g est donnée par l'équation

$$z_2 = F(x_2, y_2, \alpha, \beta) = F(x_2, y_2, \alpha, g),$$

x_2, y_2, z_2 étant les coordonnées de ce point 2. L'équation (1) étant encore vérifiée en ce point, tandis que $\frac{\partial F}{\partial \beta}$ est différent de zéro ⁸⁾, ceci donne encore $g' = 0$; puis g'' est donné par

$$\frac{\partial^2 F(x_2, y_2, \alpha, 0)}{\partial \alpha^2} + \frac{\partial F}{\partial \beta} g'' = 0.$$

Si le point 2 appartient à une caustique de l'onde Γ_0 , ce que nous supposons avoir lieu pour $\alpha = 0$, il satisfait à la relation

$$\frac{\partial^2 F}{\partial \alpha^2} = 0,$$

ce qui donne, pour g'' , la valeur zéro. A partir de cette position, déplaçons le point 2 sur la courbe de section τ , supposée non tangente à la bicaractéristique, α prenant alors une valeur différente de zéro que nous considérerons comme infiniment petit principal. On voit que la nouvelle valeur de g'' sera un infiniment petit du même ordre que celle de $\frac{\partial^2 F}{\partial \alpha^2}$. Or, si, dans le plan des x, y , nous considérons la famille de courbes

$$(4) \quad f(x, y, \alpha) = 0$$

⁸⁾ Nous réserverons cette dénomination à la courbe dont il s'agit, par opposition aux autres traces (trace de Γ_0 sur S_1 ; traces de conoïdes ayant leurs sommets sur S_1 ou sur S_2) dont elle est l'analogue d'après sa définition, en raison du rôle spécial qu'elle jouera dans ce qui va suivre : en particulier, elle seule aura des singularités.

⁹⁾ A_2 , note de la p. 250.

(en l'espèce, f représente la dérivée $\frac{\partial F}{\partial \alpha}$), donnant l'axe des x pour $\alpha = 0$, de sorte que l'on aura, pour α voisin de zéro,

$$(4') \quad f. = y f_0 + \alpha f_1 + \alpha^2 f_2,$$

le coefficient f_1 de α étant nul au point de contact avec l'enveloppe de la famille (4) (point que nous pourrions par exemple prendre comme origine) et si, d'autre part, les deux autres coefficients f_0, f_2 sont différents de zéro au même point, la forme de l'équation $f = 0$ montre que, sur un chemin issu de l'origine et non tangent à l'axe des x (de sorte que x est un infiniment petit de l'ordre de y au moins), y est un infiniment petit d'ordre 2 exactement en α et

$$\frac{\partial f}{\partial \alpha} = \frac{1}{f_2} \sqrt{f_1^2 - 4 f_0 f_2 y}$$

un infiniment petit du premier ordre. Appliquant ceci à la projection de notre figure de l'espace sur notre plan des $x y$, nous voyons que $\frac{\partial^2 F}{\partial \alpha^2}$ et, par conséquent, $g''(\alpha)$ seront des infiniment petits du premier ordre exactement. Il en sera de même dès lors, d'après l'expression (3) dans laquelle le numérateur $\left(\frac{\partial^2 \Phi}{\partial \alpha \partial y}\right)^2$ est différent de zéro ¹⁰⁾, pour la différence qui existe, sur la surface S_1 , entre les courbures des traces de l'onde et du conoïde direct ayant pour sommet le point 2 variable sur la courbe de section τ .

Quant aux deux hypothèses $f_0 = 0, f_2 = 0$, la première est exclue d'une manière analogue à ce qui a été dit ¹¹⁾ dans A_2 , comme impliquant, au point 2, une singularité de la bicaractéristique ; la seconde, soit $\frac{\partial^3 F}{\partial \alpha^3} = 0$, signifierait (Cf. A_2 , nos 3, 6) que la surface S_2 passe par un point de rebroussement de la caustique.

3. Nous ferons maintenant porter notre comparaison sur le conoïde direct ayant pour sommet le point 2 de l'axe des x — appartenant ou non à la caustique — et sur le conoïde rétrograde ayant pour sommet le point 1 d'abscisse un peu inférieure ¹²⁾ à celle de la surface-écran S_1 . Si, tout

¹⁰⁾ A_2 , p. 253.

¹¹⁾ Page 253. La remarque faite en cet endroit concerne le point 1, mais s'applique sans modification au point 2.

¹²⁾ Le sens positif sur l'axe des x est supposé comme dans A_2 , le sens rétrograde, celui qui va de S_1 à S_2 .

d'abord, le point 2 est lui-même très voisin de S_1 , mais du côté positif, les traces de ces deux conoïdes sur S_1 seront deux petites ovals extérieures l'une à l'autre. Lorsque nous déplacerons ensuite continuellement le point 2 le long de l'axe des x , dans le sens positif, jusqu'à l'amener sur S_2 , il n'arrivera pas que la valeur correspondante de g'' (o) devienne égale à celle qui convient au conoïde de sommet 1 : car, comme on le voit en raisonnant sur ce dernier conoïde comme nous l'avons fait sur Γ_0 , cela signifierait qu'il admettrait une caustique passant par le point 2, contrairement à l'hypothèse β . Il n'arrivera pas non plus que les traces des deux conoïdes soient tangentes en un second point, car ils auraient alors une seconde bicaractéristique commune ¹³⁾ joignant le point 1 au point 2, ce qui serait encore contraire à l'hypothèse β . Ces deux traces resteront donc, jusqu'au bout, tangentes *extérieurement* l'une à l'autre.

Dans ces conditions portons, dans S_1 , sur une ligne partant de l'origine et non tangente au plan des xy — par exemple sur l'axe des z , si S_1 a pour équation $x=0$ —, et cela dans le sens positif, qui sera supposé dirigé vers l'intérieur de la petite aire découpée par le conoïde de sommet 1, une petite longueur qui sera considérée comme l'infiniment petit principal. D'après l'hypothèse β , l'extrémité $\bar{1}$ de cette longueur pourra être jointe au point 2 par une géodésique faisant avec le conoïde, au sommet 2 de ce conoïde, un angle du même ordre que la distance du point $\bar{1}$ à la trace du conoïde en question, c'est-à-dire du premier ordre. Une telle géodésique sera d'ailleurs extérieure au conoïde (géodésique *d'espace*) et donnera, par conséquent, pour la quantité H , une valeur négative et infiniment petite du premier ordre seulement. Comme la valeur de H est constante le long de cette géodésique, celle-ci sera également, au point $\bar{1}$, extérieure au conoïde correspondant et fera, avec la direction de toute géodésique $\bar{\mathfrak{L}}$ issue de $\bar{1}$ et intérieure à ce conoïde, un angle qui ne pourra être infiniment petit d'ordre supérieur au premier. Donc encore, le point $\bar{2}$, trace de la géodésique $\bar{\mathfrak{L}}$ sur S_2 , sera à une distance du point 2 qui — toujours d'après l'hypothèse β — ne pourra être infiniment petite d'ordre supérieur au premier.

Au point 2, sur la surface S_2 que nous pouvons supposer être aussi un plan $x=\text{const.}$, le sens positif sur l'axe des z sera encore dirigé vers l'intérieur du conoïde de sommet 1. Il sera également dirigé vers l'intérieur du conoïde de sommet $\bar{1}$, infiniment voisin du premier. Ces deux

¹³⁾ Par un élément de contact sur S_1 passent deux bicaractéristiques, mais on a vu dans *J. M.* (10) comment il y a lieu de choisir entre elles, et ce choix est manifestement le même pour les deux conoïdes, puisqu'il s'agit de directions toutes deux très voisines de la bicaractéristique primitive.

conoïdes ne sauraient se couper au voisinage du point z considéré : car leur intersection serait le sommet d'un conoïde direct infiniment voisin de Γ_2 et dont la trace sur S_1 passerait par les points 1 et $\bar{1}$, alors que toute corde d'une telle trace, infiniment petite et infiniment voisine du point 1 , ne peut avoir qu'une direction infiniment voisine de celle de l'axe des y . Dès lors la ligne $\gamma_{\bar{1}}$, trace sur S_2 du conoïde de sommet $\bar{1}$, sera du côté des z positifs par rapport à la trace γ_1 du conoïde de sommet 1 , c'est-à-dire sera tout entière intérieure à ce conoïde ; et il en sera de même du point $\bar{2}$.

La géodésique $1\bar{2}$ sera donc une géodésique de temps : elle correspondra à une valeur de H positive et infiniment petite du premier ordre seulement. Donc la trace sur S_1 du conoïde de sommet $\bar{2}$ passera dans la région $z > 0$, à une distance infiniment petite du premier ordre (et non d'ordre supérieur) du point 1 .

On voit que ces considérations, qui peuvent paraître évidentes par elles-mêmes dépendent de la seule hypothèse β et ne sont pas influencées par les singularités possibles de Γ_0 .

Ces points notés, nous reprendrons, en les simplifiant un peu ¹⁴⁾, les notations de JM .

II.

4. Conformément à ce qui a été fait dans JM et dans nos autres travaux précédents sur le principe de Huygens, nous continuerons à prendre comme point de départ la formule ¹⁵⁾

$$(F) \quad 2\pi u_0 = \iint_{S_1} v_{01} u'_1 dS_1 - \frac{d}{dv} \iint_{S'_1} v_{01} u_1 dS_1 + \iint_{S_1} L_1 v_{01} u_1 dS_1 \\ = (a) - (b) + (c)$$

qui résout le problème de Cauchy, en exprimant la valeur de l'inconnue u au point 0 en fonction des valeurs u_1 de cette inconnue et des valeurs u'_1 de sa dérivée transversale en tous les points de la surface S_1 , la portion

¹⁴⁾ L'onde Γ_0 , dont l'équation, dans la notation de A_2 , était de la forme $z = y^2 \phi(x, y)$, aura pour équation, dans le système de coordonnées qui sera considéré au n° II, $y = 0$; sur cette surface, notre bicaractéristique initiale (axe des x du Mémoire A_2) sera définie par l'équation complémentaire $x = 0$. Enfin, l'écran S_1 ($x = x_1$ dans la notation de A_2) pourra être pris, dans ce qui va suivre, comme plan des x, y , en définissant, par exemple, z comme une distance transversale à S_1 , mesurée avec la forme métrique H .

¹⁵⁾ Nous nous dispenserons, pour simplifier, d'affecter de l'indice 1 la lettre v_1 pour désigner la transversale à S_1 , la notation v_2 étant toujours employée pour désigner la transversale à S_2 .

utile ¹⁶⁾ de cette surface, celle à laquelle est étendue l'intégration, étant celle qui est comprise à l'intérieur du conoïde caractéristique I'_0 de sommet o . Comme dans le Mémoire *JM*, les trois termes de cette formule seront respectivement désignés par (a) , (b) , (c) , ce dernier terme étant toujours d'une étude plus simple que les premiers, de sorte que les résultats qui le concernent peuvent être écrits sans qu'il y ait lieu d'insister sur leur démonstration. La quantité

$$v_{01} = \frac{V_{01}}{\sqrt{I_{01}}}$$

est la solution élémentaire, solution de l'équation (E) par rapport aux coordonnées du point o et de son adjointe (ε) par rapport aux coordonnées du point 1 . Dans le terme (b) , on doit différentier par rapport à v l'intégrale double

$$(5) \quad \iint_{S'_1} v_{01} u_1 dS_1$$

étendue successivement à la portion ¹⁶⁾ de S_1 comprise dans le conoïde I_0 et à la portion analogue d'une surface S'_1 menée à une distance transversale constante très petite v de S_1 , les valeurs de u étant prises les mêmes en deux points correspondants 1 et $1'$ de S_1 et de S'_1 , c'est-à-dire en deux points qui sont projection transversale l'un de l'autre. La même convention s'applique à l'élément de surface dS_1 ; au contraire, la quantité v doit être calculée pour les points o et 1 dans un cas, pour les points o et $1'$ dans l'autre. Le terme (a) peut d'ailleurs, comme nous l'avons fait dans *JM*, se mettre également sous la forme d'une dérivée par rapport à v , celle de l'intégrale

$$(5 \text{ bis}) \quad \iint_{S_1} v_{01} u_1 dS_1$$

avec la convention, contraire à la précédente, que la valeur de u est prise différente au point 1 et au point $1'$, tandis que la valeur de v est la même en ces deux points (ainsi que l'élément de surface).

¹⁶⁾ Comme dans nos travaux précédents, ce sera cette région qui sera plus spécialement désignée par la notation S_1 , la même convention étant faite en ce qui regarde S_2 .

5. Il n'est pas sans intérêt pour nous de préciser une manière dont peut s'effectuer, dans le terme (b), la différentiation de l'intégrale (5) sans que la singularité de la solution élémentaire le long du conoïde caractéristique donne lieu à aucune difficulté. C'est à quoi nous arriverons aisément en considérant un point arbitraire de S_1 comme défini par la géodésique L qui le joint au point o . En vertu de l'hypothèse α faite sur la situation relative de ce point et de la surface S_1 , les deux paramètres ξ, η qui définissent la direction d'une géodésique au départ de o pourront être considérés comme coordonnées curvilignes sur S_1 , l'élément de cette dernière surface étant de la forme

$$(6) \quad K d\xi d\eta$$

avec $K \neq 0$. Il est clair que, dans un pareil système de coordonnées, la trace du conoïde caractéristique de sommet o aura la même équation, quelle que soit la surface de section S_1 .

6. Plus précisément encore, nous pouvons écrire

$$H = \omega_3^2 - \omega_1^2 - \omega_2^2,$$

les polynômes

$$\omega_i = \omega_i(dx; x)$$

étant du premier degré en dx^1, dx^2, dx^3 , avec des coefficients fonctions régulières des x , au moins dans une petite région autour de la position considérée du sommet du conoïde o . Si alors nous posons

$$\xi = \frac{\omega_1(\dot{x}; x)}{\omega_3(\dot{x}; x)}, \quad \eta = \frac{\omega_2(\dot{x}; x)}{\omega_3(\dot{x}; x)}$$

où les $\dot{x}^i$ sont des valeurs de dx^i/ds au sommet, nous pourrons, à chaque système de valeurs de ξ, η , faire correspondre des valeurs déterminées (et non plus seulement des rapports mutuels) des x — par exemple, en prenant

$$\dot{x}^i = \frac{dx^i}{\sqrt{\sum (dx^i)^2}},$$

ces valeurs des $\dot{x}$ étant fonctions régulières de ξ, η et des x . Dans ces conditions, la quantité Γ_{01} sera le produit d'un facteur différent de 0

tant sur S_1 que sur S'_1 par la quantité $1 - \xi^2 - \eta^2$, toujours la même quelles que soient non seulement la surface de section S_1 , mais même la position du sommet o . Ce facteur étant le seul qui donne lieu à singularité dans la quantité

$$v_{01} = \frac{V_{01}}{\sqrt{1 - \xi^2 - \eta^2}},$$

une différentiation quelconque effectuée sur v_{01} n'élèvera pas l'ordre de la singularité puisqu'elle n'affectera que le numérateur régulier V_{01} . Il ne se présentera donc aucune difficulté dans la différentiation de l'intégrale (5) par rapport à v et même, un nombre de fois quelconque, par rapport aux coordonnées du sommet du conoïde [telle qu'on aura à l'appliquer à la quantité u_1 définie par la formule (F_1) , pour la substituer dans (F) (Cf. n° 8)].

7. Nous avons, jusqu'ici, considéré uniquement une onde rétrograde émanant d'un point fixe de l'espace-temps. Mais des circonstances plus générales peuvent se présenter, par exemple dans le cas d'une réflexion. Analytiquement, ce phénomène se traduit par la résolution d'un « problème mixte » et introduit ¹⁷⁾, outre la fonction v , une autre fonction (v) , singulière sur l'onde réfléchie, c'est-à-dire sur une caractéristique qui n'a plus de point conique, laquelle caractéristique découpera, sur une surface telle que S_1 , une certaine région intérieure. Il y aura encore lieu de différentier par rapport à v (c'est-à-dire par rapport à la substitution de S'_1 à S_1) une intégrale double étendue à une région de cette espèce : la méthode précédente permettra d'éviter toute difficulté dans cette différentiation si l'on a, dans la portion d'espace intérieure à la caractéristique dont il s'agit et au moins au voisinage de S_1 , défini des coordonnées curvilignes ξ, η, s satisfaisant aux conditions suivantes :

$\xi = \text{const}, \eta = \text{const}$, ces deux constantes satisfaisant à une relation

$$(7) \quad \Psi(\xi, \eta) = 0,$$

représentera une bicaractéristique située sur I_0 ;

pour $\Psi(\xi, \eta) > 0$, ces mêmes équations $\xi = \text{const.}, \eta = \text{const.}$ définiront une géodésique $\mathfrak{L}$ orientée dans le temps ou même de longueur nulle, ligne qui, ne pouvant être nulle part tangente à la surface d'espace S_1 , sera supposée la couper à l'intérieur de l'aire délimitée par I_0 .

¹⁷⁾ Voir Le problème de Cauchy (Paris Hermann), Appendice II, 223.

Dans toute région où l'onde incidente (onde issue d'un point d'univers unique o ou onde provenant de réflexions ou de réfractions antérieures) sera dépourvue de caustique, une pareille congruence $\mathfrak{L}$ et de pareilles coordonnées pourront être définies.

Remarque. — Lorsqu'on opérera comme il vient d'être dit dans ce n^o ou dans l'un des deux précédents, on devra, dans la différentiation de u_1 , ainsi que dans celle du facteur K qui figure dans l'expression (6) de dS_1 , tenir compte de ce que le calcul de ces deux quantités introduit la projection transversale, sur S_1 , d'un point arbitraire $\mathbf{r}'$ de S'_1 . Soient x^1, x^2, x^3 les coordonnées de cette projection ; $\frac{\delta x_i}{\delta \nu}$, leurs dérivées lorsque le point $\mathbf{r}'$ ainsi projeté variera en fonction de ν sans changement de ξ, η (de sorte que $\mathbf{r}'$ sera l'intersection de S'_1 avec la géodésique ou la ligne $\mathfrak{L}$ qui viennent d'être considérées). On devra avoir

$$(8) \quad \sum_i \pi_i \frac{\delta x^i}{\delta \nu} = 0,$$

les π_i étant les paramètres directeurs du plan tangent à S_1 ; et, d'autre part, d'après la relation qui existe entre les points $\mathbf{r}$ et $\mathbf{r}'$,

$$\frac{\delta x^i}{\delta \nu} + \frac{dx^i}{d\nu} = \varrho \frac{\partial x^i}{\partial s} = \varrho \dot{x}^i,$$

les dérivées $\dot{x}^i$ étant, cette fois, prises au point $\mathbf{r}'$ et ϱ désignant la dérivée $\frac{ds}{d\nu}$: cette dernière quantité ϱ se déterminera en portant les valeurs précédentes dans (8), son coefficient dans l'équation ainsi obtenue étant différent de zéro puisque la ligne $\xi = \text{const.}, \eta = \text{const.}$ n'est pas tangente à S_1 . Les dérivées finies et régulières ainsi obtenues pour les coordonnées spatiales de la projection permettront de calculer les dérivées des coordonnées curvilignes correspondantes sur S_1 , dérivées que l'on pourra, à leur tour, différentier par rapport à ξ, η , de manière à en déduire la dérivée, par rapport à ν , du déterminant fonctionnel et du facteur K qui permet d'exprimer, à l'aide de $d\xi d\eta$, la projection transversale de l'élément superficiel dS'_1 décrit par le point $\mathbf{r}'$. Ce sont ces dérivées, toutes régulières, qui serviront à différentier l'intégrale (5) dans les conditions indiquées.

8. Nous aurons, comme dans JM , à combiner la formule (F) avec la formule analogue

$$F_1) \quad 2\pi u_1 = \iint v_{12} u_2' dS_2 - \frac{d}{dv_2} \iint v_{12} u_2 dS_2 + \iint L_2 v_{12} u_2 dS_2,$$

dans laquelle le point o est remplacé par un point 1 de S_1 (ou par un point $1'$ de S'_1), S_1 étant remplacé par S_2 ou S'_2 ; et, comme dans ce Mémoire, ceci conduira à introduire, outre les traces de l'onde incidente I'_0 sur S_1 ou sur S'_1 :

la nappe rétrograde I'_1 de conoïde ayant pour sommet un point 1 ou $1'$, et sa trace γ_1 sur S_2 ou S'_2 ;

la nappe directe I'_2 de conoïde ayant pour sommet un point 2 ou $2'$ de S_2 ou de S'_2 , et sa trace γ_2 sur S_1 ou S'_1 .

Dans les hypothèses où nous nous plaçons, les nappes conoïdales de ces deux catégories, limitées à S_2 ou S'_2 , l'autre à S_1 ou S'_1 , ne présenteront pas de caustiques.

La condition que la trace de la seconde d'entre elles sur S_1 ait des points intérieurs communs avec celle de I'_0 définit la région de S_2 ou de S'_2 intérieure à l'onde. Lorsque le point 2 sera sur l'onde elle-même, les deux traces seront tangentes extérieurement¹⁸⁾. Elles seront tangentes intérieurement lorsque le point 2 appartiendra à ce que nous avons appelé dans notre travail précédent, l'« onde interne » et que nous préférons, dans celui-ci, désigner par le terme plus topique de « pseudo-onde ». Comme il a été expliqué antérieurement, en effet, cette pseudo-onde ne doit intervenir qu'en apparence, et ne donner lieu, tous calculs faits, à aucun phénomène particulier.

9. Les circonstances que nous venons de mentionner sont celles qui se présentaient dans nos Mémoires antérieurs. Mais dans nos hypothèses actuelles, des circonstances nouvelles vont apparaître, conformément aux résultats obtenus dans A_2 . Lorsque le point 2 appartiendra à une caustique de l'onde, les deux traces auront entre elles un contact du second ordre¹⁹⁾.

¹⁸⁾ Ceci reste valable même lorsqu'on abandonne l'hypothèse γ (quoique, dans certains cas de figure, tels que ceux qui sont représentés sous les nos IV—VI sur la figure 3, les deux aires découpées dans S_1 par I'_0 et par I'_2 pénètrent l'une dans l'autre au voisinage du point de contact), en ce sens que, sur une ligne menée par ce point et non tangente aux traces, les deux segments respectivement intérieurs aux aires en question n'ont d'autre point commun que leur extrémité.

Au contraire, si le point 2 était sur la pseudo-onde, les deux segments dont il s'agit seraient de même sens et se recouvriraient.

¹⁹⁾ Si S_2 passait par un point de rebroussement de la caustique (Cf. A_2 , 3), les deux traces correspondantes auraient un contact du troisième ordre. Cette disposition étant exceptionnelle, nous ne jugeons pas utile d'en entreprendre l'étude.

Il résulte également de A_2 que la surface S_2 (du moins lorsqu'elle commence à s'écarter de S_1 : voir plus loin n° 10) ne peut pas rencontrer la caustique sans rencontrer également une ligne double de la caractéristique Γ_0 . Il y a lieu de reprendre ici la discussion des relations de contact entre les deux traces en faisant varier le point 2 non plus le long d'une bicaractéristique, mais le long de la section τ de l'onde par la surface S_2 , laquelle comporte (au moins) un point double et deux points de rebroussement.

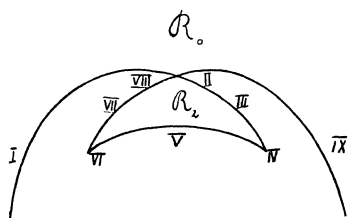


Fig. 2.

Cette section étant représentée fig. 2, nous avons représenté schématiquement, fig. 3, les formes successives de la courbe γ_2 , trace de Γ_2 , la trace de Γ_0 étant figurée²⁰⁾ par l'axe des x . Sur un premier arc décrit par le point 2, ce point étant, par exemple, en I (fig. 3), les deux traces tangentes sont

dans la position relative normale, c'est-à-dire celle qu'on rencontrerait seule s'il n'y avait pas de caustique²¹⁾. La première particularité qui se présente est nécessairement, comme

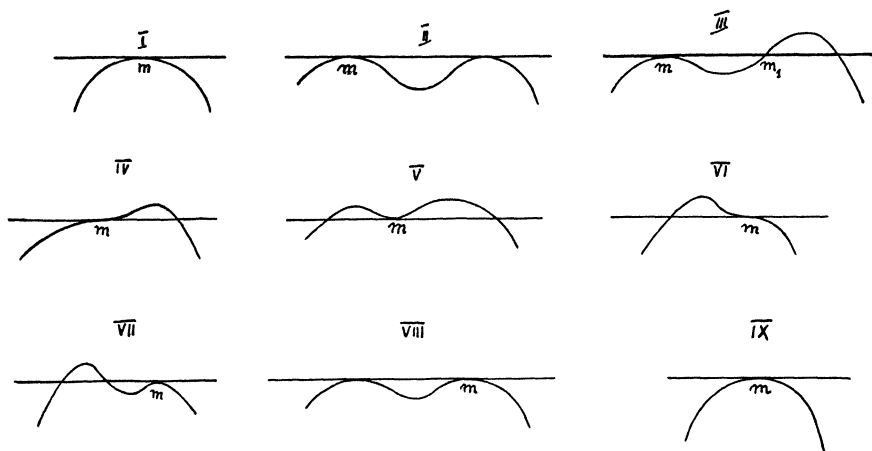


Fig. 3.

nous l'avons constaté dans A_2 , la présence d'un second point de contact (cas (II), fig. 3): ceci correspond au premier passage du

²⁰⁾ Les figures successives (I)–(IX) doivent être en réalité envisagées comme décalées les unes par rapport aux autres, le point m se déplaçant continuellement (de gauche à droite dans la disposition adoptée dans nos dessins).

²¹⁾ Il se peut (Cf. 10) que de tels points I n'existent pas, aucune position du point 2 sur la courbe de la section τ ne donnant lieu à des traces complètement extérieures l'une à l'autre. Mais cette circonstance ne peut pas se présenter lorsque S_2 , s'écarter progressivement de S_1 , commence à rencontrer la caustique.

point mobile 2 par le point double II de la section. Puis (position III du point mobile) la trace γ_2 , toujours tangente extérieurement en un premier point de contact m , pénètre par ailleurs dans la région intérieure $y > 0$. On arrive ainsi à un point de rebroussement IV de la section, l'osculation des deux traces étant atteinte par le fait que le premier point d'intersection m_1 de ces deux lignes dans le cas (III) vient coïncider avec m , pendant que la dérivée seconde d^2y/dx^2 en m , primitivement négative, augmente jusqu'à la valeur zéro. Cette dérivée seconde devenant ensuite positive, la disposition devient celle du cas (V), c'est-à-dire que la seconde trace γ_2 a deux régions différentes communes avec la première. Les déformations se succèdent ensuite dans l'ordre inverse. Le point d'intersection le plus à droite des deux traces venant à son tour coïncider avec m , on a une nouvelle osculation, correspondant au second point de rebroussement VI ; puis, à nouveau, pour toute position telle que VII, une seule aire intérieure commune aux deux traces ; puis, nouveau passage VIII par le point double (de sorte que la figure (VIII) n'est autre que la figure (II), à la dénomination près du point m) après quoi on retrouve, en IX, la position normale.

Il est à remarquer que la courbe de section que nous avons tracée fig. 2 ($\mathcal{R}_0$, région extérieure) aurait pu aussi, semble-t-il, être conçue avec la disposition représentée fig. 2^{bis}, différant de la première par le fait que le triangle compris entre le point double et les deux points de rebroussement est *extérieur*. Mais la discussion précédente permet de décider d'une façon certaine entre ces deux hypothèses et d'éliminer la seconde d'entre elles. Dans cette dernière, en effet, il y aurait, au voisinage des points de rebroussement, des points entièrement extérieurs à l'onde : pour de telles positions du point 2, les deux traces seraient entièrement extérieures l'une à l'autre. Or c'est ce qui ne peut se produire au voisinage des points IV ou VI, en vertu de la forme des figures (IV) et (VI) qui correspondent à ces points. Au contraire, les figures en question montrent qu'il existera, sur S_2 , des points — en l'espèce ceux qui sont compris dans le triangle $\mathcal{R}_2$ (fig. 2) délimité ainsi entre trois arcs de la courbe τ — « doublement intérieurs » à l'onde, c'est-à-dire tels que la trace du conoïde direct ayant pour sommet l'un d'eux ait deux régions différentes intérieures à γ_0 .

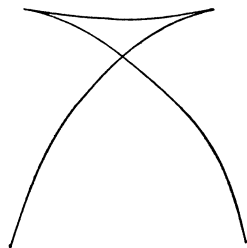


Fig. 2^{bis}.

10. Tous ces détails peuvent être aisément vérifiés sur l'exemple des ondes cylindriques ordinaires, c'est-à-dire sur les modèles en fils de surfaces à pente constante ou, ce qui revient au même, sur les figures de courbes parallèles à une courbe fixe. Nous figurons un peu plus loin (fig. 5) la disposition correspondant à une onde dont le front, à l'instant 1²²) coïncide avec une ellipse de demi-axes a , b et qui — comme il est nécessaire pour trouver une caustique — se propage, lorsqu'on fait décroître le temps t , vers l'intérieur de la courbe, de sorte que la région du plan S_1 qualifiée d'intérieure à l'onde sera la région extérieure à l'ellipse.

S_2 sera également un plan $t = \text{const.}$, de sorte que la section de l'onde par S_2 s'obtiendra en portant, à partir du pied de chaque normale à l'ellipse, une longueur constante L dans le sens intérieur. Conformément à ce qui a été dit au n° 7 nous serons conduits à faire passer non seulement par les points situés sur l'ellipse, mais par les points extérieurs. c'est-à-dire intérieurs à notre onde, des géodésiques — c'est-à-dire des droites — $\mathfrak{L}$ projetées, par exemple, sur les plans $t = \text{const.}$ suivant les normales à l'ellipse. Si, comme il a été admis au n° 7, nous supposons encore ces géodésiques de longueur nulle, c'est-à-dire de même pente que celles qui sont issues des points de la courbe, nous voyons que la correspondance entre les régions intérieures à l'onde situées sur S_1 et sur S_2 se fera, en projection sur $t = \text{const.}$, par segments de longueur constante L portés sur les normales à l'ellipse. Cette correspondance sera univoque à partir de S_1 (la normale qui sert à la réaliser étant, de convention expresse, la plus courte, celle qui ne passe pas à l'intérieur de l'ellipse entre son point de départ et son pied); mais elle ne le sera pas à partir de S_2 , dès que L dépassera le rayon de courbure minimum de l'ellipse (rayon de courbure au sommet du grand axe, soit $\frac{b^2}{a}$): à partir de ce moment, l'aire triangulaire $\mathcal{R}_2$ prendra naissance et tout point situé à l'intérieur de cette aire²³⁾ aura, dans S_1 , deux homologues différents. La trace γ_2 correspondant à un tel point, c'est-à-dire le cercle de rayon L ayant ce point pour centre, coupera la conique en quatre points de manière à déterminer, extérieurement à elle, deux aires distinctes.

Ajoutons que les choses se passent d'une manière un peu différente si l'on fait encore croître la constante L . Pour des valeurs suffisamment grandes de cette quantité, le point double de la courbe de section disparaît, la région extérieure que nous avons précédemment dénommée $\mathcal{R}_0$ cessant d'exister: au-

²²⁾ Sous le point de vue auquel nous nous plaçons, comme dans tout ce qui précède, il s'agit d'une propagation rétrograde, c'est-à-dire vers le passé.

²³⁾ La même propriété appartiendra d'ailleurs (toujours pour $L > b^2/a$) à d'autres points de la figure. En fait, la forme de la „surface de Riemann“ qui correspond à S_1 de par la transformation considérée est facile à indiquer immédiatement, car elle n'est autre que celle de la surface d'onde Γ_0 elle-même (convenablement arrêtée à la courbe de section τ , puisque S_1 est arrêtée à l'ellipse). Cette surface affecte donc, au voisinage du point de rebroussement de la caustique, la forme décrite dans A_2 (n° 6): elle recouvre trois fois le triangle $\mathcal{C}$ compris entre les deux branches de la développée et l'arc IV—VI de la courbe de section, deux fois la région $\mathcal{R}_2$. Mais nous n'avons à tenir compte, en ce moment, que du cas où deux images, sur S_1 , d'un même point de S_2 sont près de se confondre, et ce cas est celui que nous considérons dans le texte.

trement dit, aucun point 2 n'est centre d'un cercle de rayon L entièrement intérieur à l'ellipse (c'est-à-dire extérieur à l'onde). C'est ce qui se produit évidemment à partir du moment où L dépasse le demi petit axe ²⁴⁾. Dans ce cas, S_2 peut rencontrer la caustique sans rencontrer de ligne double de la surface.

II. C'est sous le bénéfice de ces données géométriques que nous allons reprendre (Cf. JM) l'étude d'une intégrale telle que

$$(9) \quad I = \iint v_{01} v_{12} dS_1$$

et de sa différentiation par rapport à v . Je me propose même de reprendre cette étude dès le commencement, c'est-à-dire tout d'abord, comme dans JM , en l'absence de caustiques : la méthode de calcul employée dans JM peut en effet être notablement simplifiée et, pour le cas primitivement traité où les caustiques n'interviennent pas, les résultats deviennent à peu près intuitifs moyennant une simple interversion de l'ordre des intégrations, la première d'entre elles ayant lieu par rapport à la variable que nous avons appelée y .

Rappelons que nous avons, dans JM , appelé *fonction-unité* toute fonction régulière non nulle au voisinage duquel nous nous plaçons, de sorte que toute puissance positive ou négative, entière ou fractionnaire d'une telle unité est encore une unité. Ajoutons que de telles fonctions-unités, lorsqu'elles figurent comme coefficients dans l'équation d'une courbe, se comportent, au point de vue de la forme de celle-ci, comme des constantes. Les courbes

$$ay + bx^2 = 0, \quad ay + bx^3 = 0, \quad ay^2 + bx^3 = 0,$$

où a, b , fonctions de x et de y , sont des unités, ont sensiblement au voisinage de l'origine, la première la forme d'une parabole ; la seconde, celle d'une parabole cubique ; la troisième, celle d'une parabole semi-cubique ; etc.

III^{bis} . Dans JM , nous avons été conduits à écrire les équations de nos deux traces, au moins lorsqu'elles sont tangentes ou près d'être tangentes, sous la forme ²⁵⁾

$$(10) \quad y = 0, \quad y = Y(x)$$

²⁴⁾ Pour $L = b$, la courbe de section τ présente un auto-contact, au centre de l'ellipse.

²⁵⁾ La notation Y n'a pas le même sens que dans JM .

et, par conséquent, par un usage convenable des facteurs-unités, l'intégrale (9) sous la forme

$$(9') \quad \iint \frac{f(x, y) dx dy}{\sqrt{\pm y(y-Y)}}$$

la fonction f contenant, indépendamment de x , y et la fonction Y , indépendamment de x , les paramètres ν , ν_2 et les coordonnées du point 2 sur la surface S_2 ou S'_2 . Au voisinage du contact extérieur, c'est-à-dire lorsque le point 2 sera voisin de l'onde incidente tout en lui étant intérieur, le signe à prendre sous le radical, dans la formule (q') , sera — et l'aire d'intégration sera définie par

$$0 \leq y \leq Y(x),$$

x variant lui-même dans un intervalle (x', x'') dont les extrémités ont pour abscisses deux racines de l'équation $Y(x) = 0$. Or on peut écrire

$$(11) \quad f(x, y) = f(x, 0) + y f_1(x, y) = f_0(x) + y f_1(x, y)$$

f_1 étant elle-même une fonction régulière, donc bornée ainsi que les différentes dérivées que nous aurons à en considérer. Réduisant d'abord cette expression de f à son premier terme, nous voyons immédiatement que la quadrature par rapport à y donne le résultat $\pi f_0(x)$ et, par conséquent, la quadrature double, le résultat

$$\pi \int_{x'}^{x''} f_0(x) dx$$

Si la fonction f_0 contient les paramètres (principalement le paramètre ν) par rapport auxquels on veut différentier, la dérivée correspondante sera, par hypothèse elle-même régulière et son intégrale dans l'intervalle (x', x'') , restera bornée, ou sera même infiniment petite avec l'amplitude de l'intégrale. Nous n'aurons donc à considérer que les termes aux limites

$$(12) \quad \pi \left[f_0(x'') \frac{dx''}{d\nu} - f_0(x') \frac{dx'}{d\nu} \right]$$

et nous voyons que la dérivée $\frac{dI}{d\nu}$ ne pourra devenir infinie que (au plus) de l'ordre des dérivées

$$(13) \quad \frac{dx'}{d\nu}, \quad \frac{dx''}{d\nu}.$$

Quant au second terme de l'expression (11), en vertu de la formule

$$\int_0^Y \sqrt{\frac{y}{Y-y}} dy = \frac{\pi}{2} Y,$$

sa différentiation n'introduira aucun terme aux limites analogue à (12) et (puisque les quantités sous les signes d'intégration sont supposées régulières) donnera un résultat nécessairement borné, dont nous n'aurons pas à nous occuper. On voit d'ailleurs qu'il en serait encore de même si l'on soumettait la portion correspondante d'intégrale à un nombre quelconque de différentiations successives.

12. Ceci s'applique encore lorsque les deux traces qui limitent l'aire d'intégration sont franchement sécantes, c'est-à-dire sécantes sans être au voisinage du contact. On peut alors, moyennant une transformation ponctuelle qui ne change pas l'axe des x , mettre leurs équations sous la forme (10) et, d'autre part, les deux dérivées (12) seront finies. Un calcul très analogue s'applique d'ailleurs aux cas où l'une des traces est entièrement et « franchement » intérieure à l'autre et limite seule l'aire d'intégration, de sorte que l'on a ainsi, sous forme manifestement bornée, la dérivée, par rapport à ν ou à tout autre paramètre, d'une intégrale du type (9) toutes les fois que le point 2 est franchement intérieur à l'aire utile S_2 .

13. On pourrait appliquer une marche analogue au cas du contact intérieur, qui correspond, comme nous l'avons dit, à celui où le point 2 est sur la pseudo-onde, cas dont nous savons d'ailleurs a priori qu'il ne donne

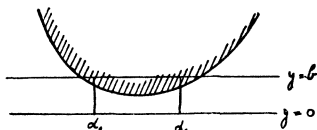


Fig. 4.

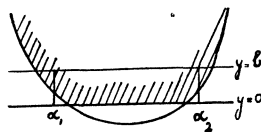


Fig. 4 bis.

lieu à aucune singularité. Au voisinage d'un tel contact intérieur, les deux traces peuvent être considérées comme ayant l'une des deux dispositions relatives représentées fig. 4 ou fig. 4^{bis}, l'aire d'intégration étant du

côte ombré sur l'une ou l'autre de ces figures et le signe à prendre sous le radical, dans la formule (9'), étant $+$. Comme dans *JM*, nous limiterons l'aire par une parallèle $y = b > 0$ à l'axe des x , toute la partie située au delà de cette parallèle donnant un résultat régulier, et par deux parallèles $x = \alpha_1$, $x = \alpha_2$ à l'axe des y situées de part et d'autre du point de contact. L'ordonnée b et les abscisses α_1 , α_2 étant supposées constantes, toute la partie de l'aire d'intégration extérieure au rectangle ainsi délimité donnera un résultat régulier. D'autre part, puisque la courbe $y = Y(x)$ est supposée passer au point considéré p pour les valeurs initiales des paramètres dont elle dépend, les deux côtés $x = \alpha_1$, $x = \alpha_2$ parallèles à l'axe des y auront pu être choisis assez voisins du point p pour que les ordonnées correspondantes de la courbe soient inférieures à b , le côté du rectangle qui appartient à la droite $y = b$ étant, par conséquent, entièrement compris dans l'aire d'intégration, et cela pour toutes les valeurs des paramètres suffisamment voisines des valeurs initiales. La primitive de $\frac{1}{\sqrt{y(y-Y)}}$ étant

$$2 \log (\sqrt{V \pm y} + \sqrt{V \pm (y - Y)}),$$

le résultat de l'intégration correspondante effectuée entre la limite supérieure b et une limite inférieure égale à la plus grande des deux quantités 0 , Y est, en toute hypothèse

$$2 \log (\sqrt{b} + \sqrt{b - Y}) - \log |Y|.$$

Cette quantité doit être multipliée par $f_0(x) dx$ et intégrée de α_1 à α_2 . Le premier terme, différentié ou non par rapport aux paramètres, ne donne lieu à aucune singularité, de sorte que toute la question porte sur la différentiation de l'intégrale

$$\int_{\alpha_1}^{\alpha_2} \log |Y(x)| dx.$$

Le calcul étant ainsi dirigé, on voit qu'il s'appliquerait, non seulement aux deux dispositions des figures 4 et 4^{bis}, mais aussi (Cf. *JM*, 9) aux

Pour les paramètres ξ, η introduits aux nos 5, 6 et qui, dans le cas d'une onde rétrograde incidente émanée d'un point fixe 0, définissent une géodésique issue de ce point, nous pourrons prendre les coordonnées x, y du point en lequel cette géodésique (ou, dans le cas du n° 7, la ligne $\mathfrak{L}$ de la congruence considérée) rencontre S_1 (de sorte que le premier membre de (7) se réduira à η). Ces mêmes quantités ξ, η serviront de coordonnées sur S_2 , d'après la méthode du n° 7. Il est bien entendu que la correspondance entre ce système de coordonnées et la position d'un point de S_2 n'est pas nécessairement biunivoque. Elle l'est dans le cas qui a fait l'objet de *JM*; elle ne le sera plus si nous admettons l'intervention des caustiques. Nous pourrons toujours exprimer, en fonction de ξ, η , des coordonnées curvilignes λ_1, λ_2 choisies sur S_2 et aptes à représenter cette surface d'une manière régulière; mais, dans le cas des caustiques, le déterminant fonctionnel de ces expressions sera susceptible de s'annuler et, avec lui, le facteur K_2 qui figure dans l'expression de l'élément superficiel

$$(6 \text{ bis}) \quad dS_2 = K_2 d\xi d\eta$$

de la surface.

C'est, en particulier, ce qui aura lieu aux points de rebroussement de la courbe de section τ , numérotés IV et VI sur la figure 3, puisque la courbe représentée sur cette figure correspond, de par la correspondance dont il s'agit, à la trace I_0 sur S_1 , trace qui est dépourvue de singularité.

15. La méthode indiquée au n° 11^{bis} amène, nous l'avons dit, à une forme très simple l'évaluation de l'intégrale (9) et de sa dérivée pour le cas traité dans *JM*, celui où il n'y a pas de caustique et où, par conséquent, pour chaque position du point 2 sur la courbe de section τ , les deux traces sont tangentes au premier ordre seulement. Le point de contact ayant une abscisse x précisément égale, dans nos notations, à la coordonnée ξ qui définit la position du point 2 sur τ , la quantité Y qui représente l'ordonnée d'une de ces traces (l'autre ordonnée étant nulle) sera (pour un contact extérieur) de la forme

$$(13) \quad Y = -A(x - \xi)^2,$$

le coefficient A , fonction régulière de x , étant positif et non nul — donc ce que nous avons appelé une unité — pour x voisin de ξ .

Si, à partir de la position que nous venons de considérer, on donne au point 2 un petit déplacement quelconque, la nouvelle expression de l'ordonnée Y deviendra

$$(14) \quad Y = -A(x - \xi)^2 + A_1(x - \xi) + A_0 = -A(x - \xi)^2 + b_1\eta(x - \xi) + b_0\eta,$$

le coefficient A ayant, en général, changé de valeur et les coefficients A_1 , A_0 étant également fonctions de ξ , η , mais contenant nécessairement η en facteur, puisque, pour toute valeur de ξ , Y a, lorsque η est nul, la forme (13).

15^{bis}. Dans le cas que nous considérons actuellement, le coefficient b_0 de η dans la formule (14), c'est-à-dire la dérivée $\frac{\partial Y}{\partial \eta}$ pour $x = \xi$, $\eta = 0$, est nécessairement positif. En effet, une valeur positive très petite de η correspond à un point z intérieur à notre aire S_2 , auquel correspond une trace γ_2 qui passe à l'intérieur de γ_0 , de sorte que le coefficient en question ne peut être négatif ; il ne peut non plus être nul, c'est-à-dire que la trace γ_2 du nouveau conoïde de sommet z ne pourra pas contenir un point ayant avec notre point $(\xi, 0)$ une distance infiniment petite du second ordre (par rapport à η pris comme infiniment petit principal, ξ restant invariable) : c'est ce que nous avons vu au n° 3.

16. Voyons maintenant comment l'expression (14) sera modifiée lorsque nous ferons varier v , c'est-à-dire lorsque nous couperons les caractéristiques I_0 et I_2 non plus par la surface S_1 , mais par la surface S'_1 à distance transversale constante de la première. Dans le système de coordonnées x, y adopté, la nouvelle trace de l'onde incidente I_0 aura encore pour équation $y = 0$ (et, en projection transversale sur S_1 , $y = v$). Celle du conoïde I_2 sera tangente à la première au point $x = \xi$, $y = 0$ et aura, par conséquent, une équation de forme analogue à (13). Pour η différent de zéro, elle sera complétée, comme tout à l'heure, par des termes en η , soit

$$(14') \quad \bar{Y} = -\bar{A}(x - \xi)^2 + \bar{b}_1 \eta (x - \xi) + \bar{b}_0 \eta$$

et, en projection transversale sur S_1 , par le terme additif v .

Les nouveaux coefficients $\bar{b}_1$, $\bar{b}_0$ (ce dernier positif) sont, aux facteurs numériques près, ceux de la formule de Taylor : ils contiennent (en général) η , v , mais non x , qui y est remplacé par la valeur ξ . Au contraire, le coefficient $\bar{A}$ contient en général (régulièrement) x . Par contre, puisque nous continuons à admettre l'hypothèse γ , ce coefficient est une unité, qui admettra même une borne inférieure positive pour toutes les positions du point z le long de la ligne τ et dans le voisinage de cette ligne.

Comme nous l'avons vu dans JM (13), les deux traces sur S'_1 ne sont pas à combiner entre elles, mais avec les traces primitives, après projection transversale sur S_1 . Dans le calcul du terme (aa) interviendra l'intégrale

$$(15) \quad \bar{I}_{02} = \iint v_{01} v_{12} dS_1 = \iint \frac{f(x, y) dx dy}{\sqrt{y(\bar{Y} + v - y)}},$$

étendue à l'aire comprise entre les lignes

$$(16) \quad y = 0, \quad y = \bar{Y} + v;$$

pour le terme $(b a)$, l'intégrale

$$(15') \quad \bar{I}_{02} = \iint v_{01} v_{12} dS_1 = \iint \frac{\bar{f}(x, y) dx dy}{\sqrt{(y-v)(Y-y)}}$$

étendue à l'aire comprise entre les lignes

$$(16') \quad y = v, \quad y = Y.$$

17. D'après ce qui a été établi au n° 11, les dérivées des intégrales (15), (15') par rapport à v , sont du même ordre de grandeur que celles des abscisses des points communs aux courbes (16) ou aux courbes (16'), c'est-à-dire des racines de chacune des équations

$$(17) \quad \bar{Y} + v = -\bar{A}(x - \xi)^2 + \bar{b}_1 \eta (x - \xi) + \bar{b}_0 \eta + v = 0,$$

$$(17') \quad Y - v = -A(x - \xi)^2 + b_1 \eta (x - \xi) + b_0 \eta - v = 0.$$

Pour différencier les fonctions implicites ainsi définies, on dispose de deux méthodes classiques :

ou bien appliquer le théorème des fonctions implicites, soit, pour l'équation (17) par exemple,

$$(18) \quad \frac{dx}{dv} = \frac{1 + \frac{\partial \bar{Y}}{\partial v}}{2 \bar{A}(x - \xi) - (x - \xi)^2 \frac{\partial \bar{A}}{\partial x} + \bar{b}_1 \eta} = \frac{1 + \frac{\partial \bar{Y}}{\partial v}}{(x - \xi) \left[2 \bar{A} - (x - \xi) \frac{\partial \bar{A}}{\partial x} \right] + \bar{b}_1 \eta}$$

ou bien imaginer que l'on forme, si l'on peut, l'expression explicite de x .

C'est, en réalité, la seconde de ces deux méthodes que nous emploierons. Mais la première va nous permettre de simplifier le calcul et de le rendre possible ²⁷⁾.

Les premiers membres des équations (17), (17') et, par conséquent, les racines que nous avons à étudier, dépendent en effet de ν de deux manières, à savoir par la présence du terme explicite $\pm \nu$ et par l'intervention du paramètre ν dans les divers coefficients. Au numérateur de (18), cette deuxième dépendance se traduit par des termes tous infiniment petits : en les négligeant, on n'altérera donc point sensiblement la dérivée cherchée.

Le dénominateur de (18) ne fait entrer la variabilité des coefficients en fonction de ν , x que par l'intermédiaire du terme en $\frac{\partial \bar{A}}{\partial x}$. Ce dernier terme, à son tour, en raison de la présence du facteur $(x - \xi)$, est infiniment petit vis-à-vis de la quantité $\bar{A}$, que nous savons être bornée inférieurement.

Ainsi, dans la différentiation qui nous intéresse, nous pouvons traiter tous les coefficients $\bar{A}$, $\bar{b}$ comme des constantes, en commettant simplement ainsi sur le résultat une erreur relative infiniment petite, sous la condition de faire subir une altération convenable, elle aussi infiniment petite, au coefficient $\bar{A}$. Les évaluations que nous nous proposons d'obtenir pour la dérivée n'en seront pas affectées si elles ne font intervenir que la borne inférieure de $\bar{A}$.

Or la résolution de l'équation du second degré (17) donne

$$x = \frac{\bar{b}_1 \eta \pm \sqrt{\bar{b}_1^2 \eta^2 + 4 \bar{A} (\eta + \nu)}}{2 \bar{A}}$$

et, en dérivant par rapport à ν et faisant $\nu = 0$, on trouve immédiatement un résultat de l'ordre de grandeur de $\frac{1}{\sqrt{\eta}}$, ce qui est bien le résultat que nous devons attendre (puisque la somme algébrique des dérivées que nous venons d'évaluer ²⁸⁾ doit fournir, à un terme borné près, la valeur

²⁷⁾ L'emploi du théorème de factorisation permettrait de surmonter la même difficulté. (Voir plus loin n° 19).

²⁸⁾ Les infinis qui interviennent dans les deux premiers termes sous le signe $\int \int$ de (19) s'ajoutent.

Le contraire a lieu au niveau de la pseudo-onde, par un mécanisme semblable à celui qui intervenait dans notre travail des *Acta Mathematica*, tome 49, (n° 21, p. 243).

de v_{02}) et qui suffit à justifier, comme nous l'avons vu dans JM, les transformations effectuées dans ce Mémoire et à établir la formule

$$(19) \quad v_{02} = \frac{1}{2\pi} \left(\frac{d\bar{I}_{02}}{d\nu} - \frac{d\underline{I}_{02}}{d\nu} + \mathcal{F}_{02} \right) \quad \mathcal{F}_{02} = \iint_{S_1} L_1 v_{01} v_{02} dS_1$$

18. Abandonnons maintenant l'hypothèse γ et admettons par conséquent que la courbe de section τ présente un ou plusieurs points de rebroussement, intersections de S_2 avec une caustique. Les propriétés géométriques de la figure seront alors celles qui ont été décrites au n° 9. Nous continuerons d'ailleurs à rapporter les points de S_2 aux coordonnées ξ, η considérées au n° 14; seulement celles-ci n'auront plus, avec les points qu'elles représentent, une correspondance parfaitement biunivoque: un point de S_2 pourra avoir suivant les cas, soit un, soit plusieurs systèmes de coordonnées.

Cette circonstance ne change d'ailleurs, en général, rien d'essentiel aux calculs précédents ou aux considérations de JM. Les points 2 qui auront deux systèmes de coordonnées ξ, η seront les sommets de conoïdes interceptant à leur intérieur, sur S_1 , des régions ayant en commun avec celle qui est intérieure à l'onde incidente, deux aires distinctes: une intégrale telle que $\bar{I}_{02}, \underline{I}_{02}$ ou $\mathcal{F}_{02}$ devra alors être étendue successivement à ces deux aires; sa différentiation fera intervenir non plus deux, mais quatre quantités de la forme (13); mais celles-ci continueront à relever de l'étude précédente.

Les seuls points qui exigent une étude nouvelle sont les points de rebroussement. Nous considérerons, par exemple, le point IV de la fig. 2, qui pourra évidemment être supposé correspondre à l'origine des coordonnées dans le plan des ξ, η , pendant que le point m correspondant sera l'origine des coordonnées sur S_1 . Nous limiterons l'aire d'intégration, sur S_1 , à sa portion intérieure à un petit rectangle fixe ayant pour base, sur l'axe des x , un segment (ξ_1, ξ_2) pris autour de l'origine.

La formule de Taylor utilisée pour représenter l'ordonnée Y devra, pour ξ voisin de zéro, comporter un terme de plus, soit

$$(20) \quad Y = H(x - \xi)^3 + A(x - \xi)^2,$$

x ne figurant plus, cette fois, dans le coefficient A , mais dans le coefficient H . Ce dernier sera différent de zéro même pour $\xi = 0$, sans quoi ($A_2, 3$) le point IV considéré serait un point de rebroussement de la

caustique, et nous excluons le cas où S_2 passerait par un tel point de rebroussement. Si nous avons pris comme sens positif, sur l'axe des x , le sens de gauche à droite sur la figure 3, le coefficient en question sera positif pour toutes les valeurs de ξ qui nous intéressent actuellement.

Le coefficient A (courbure au point de contact) s'annulera au contraire, pour $\xi = 0$. Ce zéro, par rapport à la variable indépendante ξ , est simple ²⁹⁾, en vertu de ce qui a été dit au n° 2. Nous écrirons donc

$$Y = H(x - \xi)^3 + a\xi(x - \xi)^2,$$

le nouveau coefficient a étant une unité et étant d'ailleurs positif, puisque d'après les figures (III)—(V) (fig. 3) la concavité de la ligne γ_2 passe du côté négatif au côté positif de l'axe des x lorsque ξ passe par la valeur zéro en croissant.

Conformément à ce qui a été dit au n° 1, l'équation qui vient d'être écrite conserve sa forme, non seulement pour $\nu = 0$, mais ³⁰⁾ lorsqu'on fait varier ν .

Pour η différent de zéro, mais ν nul, la valeur de Y s'écrira

$$(20') \quad Y = H(x - \xi)^3 + A(x - \xi)^2 + b_1\eta(x - \xi) + b_0\eta, \\ A = A(\xi, \eta) = a\xi + b\eta.$$

Grâce aux remarques faites au n° 3, nous pouvons affirmer que le résultat du n° 15^{bis} subsiste dans les conditions actuelles, c'est-à-dire que le coefficient b_0 est encore différent de zéro et même positif.

Si enfin l'on fait également varier ν , on aura (en projection sur S_1 et avec modifications des divers coefficients), pour équation de la trace du conoïde Γ_2 ,

$$y = \bar{Y} = \bar{H}(x - \xi)^3 + \bar{A}(x - \xi)^2 + \bar{b}_1\eta(x - \xi) + \bar{b}_0\eta, \\ \bar{A} = \bar{A}(\xi, \eta, \nu) = \bar{a}\xi + \bar{b}\eta,$$

de sorte que les deux équations en x seront

$$(21) \quad \bar{Y} + \nu = \bar{H}(x - \xi)^3 + \bar{A}(x - \xi)^2 + \bar{b}_1\eta(x - \xi) + \bar{b}_0\eta + \nu = 0,$$

$$(21') \quad Y - \nu = H(x - \xi)^3 + A(x - \xi)^2 + b_1\eta(x - \xi) + b_0\eta - \nu = 0.$$

²⁹⁾ S_2 , surface d'espace, n'est certainement pas tangente à la bicaractéristique.

³⁰⁾ Le zéro de la fonction $A(\xi)$ serait, au contraire, variable avec ν_2 .

Dans ces deux équations, on peut simplifier l'expression des coefficients A , $\bar{A}$ de $(x - \xi)^2$ par un changement de variable de la forme

$$(22) \quad \xi + \eta \varphi(\xi, \eta)$$

qui laisse fixes les points de l'axe des ξ . Il est clair, en effet, qu'on peut choisir un tel changement de variable de manière à réduire A à $a \xi$ et $\bar{A}$ à $\bar{a} \xi$. Dans le second cas, les coefficients du changement de variable dépendent de v ; mais, cette dépendance étant régulière, elle n'apporte pas d'obstacle aux calculs qui vont suivre.

19. Les divers coefficients des équations (21), (21') sont encore des fonctions de ξ , η , de v (pour la première d'entre elles) et — pour $\bar{H}$, H — de x . Mais, comme au n° 17, nous pouvons, dans le calcul de dx/dv , faire abstraction de cette dépendance et considérer les H , A , B , b comme des constantes.

Toutefois, en vue de la suite que nous aurons à donner à ce calcul aux nos 20, 22, il est préférable de ne pas user de cette faculté en ce qui regarde la dépendance de H (ou de $\bar{H}$) par rapport à x et d'appliquer, relativement à cette dernière variable, le théorème de factorisation. Donc, dans l'équation (21'), nous remplacerons H par son développement suivant les puissances de $(x - \xi)$, soit

$$H + H'(x - \xi) + H''(x - \xi)^2 + \dots$$

où cette fois, H , H' , H'' , ... sont indépendants de x et, opérant suivant la méthode de M. Goursat³¹⁾, nous mettrons le premier membre de (21') sous forme du produit d'un facteur unité par le polynôme

$$II(x) = (x - \xi)^3 + U_2(x - \xi)^2 + U_1(x - \xi) + U_0$$

les quantités U_1 , U_2 , U_3 étant déterminées par les équations

$$\begin{aligned} b_0 \eta - v - H U_0 - H' U_0 U_2 - H'' U_0 (U_2^2 - U_1) - \dots &= 0, \\ b_1 \eta - H U_1 - H' (U_0 - U_1 U_2) + H'' (U_0 U_2 + U_1^2 - U_1 U_2^2) + \dots &= 0, \\ a \xi - H U_2 - H' (U_1 - U_2^2) - H'' (U_0 - 2 U_1 U_2 + U_2^3) - \dots &= 0. \end{aligned}$$

³¹⁾ Bull. Soc. Math. Franç., T. XXXVI (1908), p. 209. Nous sommes revenus sur ce sujet dans le même recueil (T. XLVII, 1919, Comptes rendus des séances, p. 44).

L'examen des termes du premier ordre de ces équations montre que U_0 , U_1 , U_2 peuvent s'écrire

$$\begin{aligned}
 U_0 &= \frac{b_0 \eta - \nu}{H} + \dots \\
 (23) \quad U_1 &= c_1 \eta + h_1 \nu & (h_1 = \frac{H'}{H^2} + \dots, \quad c_1 = \frac{b_1}{H} - \frac{b_0 H'}{H^2} + \dots) \\
 U_2 &= \frac{a \xi}{H} + h_2 \nu & (h_2 = \frac{H''}{H^2} - \frac{H'}{H} h_1 + \dots)
 \end{aligned}$$

où, toutefois dans la forme donnée à U_2 , on suppose encore effectuée une transformation convenable du type (22). Les termes remplacés par des points sont d'ordre supérieur ; et il est encore à noter que ceux qui figurent dans U_0 contiennent tous en facteur celui qui est explicitement écrit.

Nous traiterons l'équation $II(x) = 0$ à la manière classique, en faisant d'abord disparaître le terme du second degré par le changement d'inconnue

$$(24) \quad x = -\frac{U_2}{3} + X.$$

La nouvelle équation sera

$$X^3 + PX + Q = 0$$

avec

$$(25) \quad P = U_1 - \frac{U_2^2}{3} = c_1 \eta + h_1 \nu - \frac{1}{3} \left(\frac{a \xi}{H} + h_2 \nu \right)^2$$

$$\begin{aligned}
 (25') \quad Q &= U_0 - \frac{U_1 U_2}{3} + \frac{2}{27} U_2^3 = c_0 \eta - h_0 \nu + \frac{2}{27} \left(\frac{a \xi}{H} + h_2 \nu \right)^3 \\
 & \quad (c_0 = \frac{b_0}{H} + \dots, \quad h_0 = \frac{1}{H} + \dots)
 \end{aligned}$$

(le terme en $U_1 U_2$ pouvant être incorporé aux termes en η et en ν), de sorte que le discriminant

$$(26) \quad \Delta' = -\Delta = -(4P^3 + 27Q^2)$$

se réduit, pour $\nu = 0$, à

$$(26') \quad \Delta_0' = -\Delta_0 = -c_0 \eta R, \quad R = 27 c_0 \eta + \frac{4 a^3 \xi^3}{H^3} + \dots$$

Le calcul relatif à l'équation (21) différerait uniquement du précédent par le changement de signe de ν et le changement de H, A, b_1, b_0 en $\bar{H}, \bar{A}, \bar{b}_1, \bar{b}_0$.

L'équation admet trois racines réelles pour $\Delta < 0$ et une seule pour $\Delta > 0$.

Comme il fallait s'y attendre, la valeur initiale Δ_0 du discriminant, celle qui correspond à $\nu = 0$, renferme η en facteur. L'autre facteur, égalé à zéro, représente également, sur S_2 , un lieu de points z tels que la trace γ_2 correspondante soit tangente à l'axe des x : autrement dit, il donne à nouveau ³²⁾ une partie de la courbe de section τ , qui, nous l'avons vu, a, dans le plan des $\xi \eta$, deux arcs images passant à l'origine. Le fait que ce second lieu ait avec le premier, à l'origine, un contact du second ordre ressort bien de l'interprétation que nous venons d'en donner. Prenons, par exemple, le cas, déjà considéré au n° 10, de la propagation d'ondes cylindriques ordinaires avec front initial suivant une courbe régulière quelconque telle qu'une ellipse, c'est-à-dire la transformation ponctuelle par segments de longueur constante L normaux à la courbe. Nous avons figuré (fig. 5) l'ellipse, sa développée, projection de la caustique, et sa transformée, c'est-à-dire la courbe τ . Au voisinage du point de rebroussement IV de cette dernière, c'est-à-dire du centre de courbure en un point m de l'ellipse où le rayon de courbure est L , une normale en un point m' , distant de m d'un arc qui sera pris comme infiniment petit principal, passera à une distance du point IV qui sera infiniment petite du second ordre et, par conséquent, le segment intercepté sur cette normale par la courbe τ sera, en vertu des propriétés connues du point de rebroussement, un infiniment petit du troisième ordre : or c'est ce segment qui est à reporter, avec un décalage égal à L , pour obtenir celui qui est compris entre l'ellipse, front initial, et la seconde courbe qui a même image qu'elle dans la figure S_2 .

Le même fait apparaît à l'inspection de la fig. 3 [figures partielles (III) — (IV)]. Si, en effet, on tient compte de ce que Y est une fonction croissante de η , on voit que l'unique possibilité d'une trace γ_2 coupant l'axe des x en trois points voisins de l'origine se présente pour ξ voisin de zéro et négatif, $b_0 \eta$ étant inférieur à la valeur de Y qui correspond à $\eta = 0$, laquelle est de l'ordre de ξ^3 .

³²⁾ La trace de pseudo-onde satisfait à la même condition, mais ne passe pas au voisinage de l'origine.

La région (voisine de l'origine) dans laquelle devra être situé le point (ξ, η) pour que les trois racines de l'équation cubique soient réelles est celle qui est comprise entre l'axe des x et la ligne infiniment voisine $R=0$.

20. Pour ν voisin de zéro, mais différent de zéro, la ligne $\Delta=0$ du plan des $x y$ aura une forme infiniment peu différente de la précédente. En tout cas, la forme de l'expression (26), où le coefficient de η^2 est essentiellement différent de zéro et négatif, montre que si l'on a tracé une parallèle à l'axe des x à une distance positive, suffisamment petite, mais choisie une fois pour toutes, β de cet axe et pris sur cette droite, également une fois pour toutes, un segment (ξ_1, ξ_2) suffisamment petit, la quantité Δ sera positive sur ce segment pour toutes les valeurs suffisamment petites de ν . Pour de telles valeurs de ν et pour toute valeur de ξ comprise entre ξ_1 et ξ_2 , le nombre des racines de l'équation $\Delta=0$ voisines de zéro pourra être de zéro ou de deux ; mais β sera toujours supérieur à la plus grande d'entre elles.

La forme de la ligne $\Delta=0$ dépend de la situation relative des deux lignes $P=0$, $Q=0$, dont la première (Cf. II) est sensiblement une parabole (du moins si c_1 n'est pas nul) et dont la seconde peut, au point de vue de son intersection avec la première, être assimilée à une parallèle à l'axe des ξ , savoir $\eta = \frac{h_0 \nu}{c_0}$. Ces deux lignes se coupent ou non au voisinage de l'origine (pour les petites valeurs de ν) suivant les signes comparés de ν et du coefficient $k = \frac{c_1 h_0}{c_0} + h_1$. Dans le premier cas, les deux points d'intersection seront,

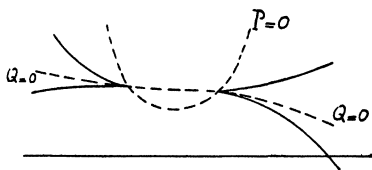


Fig. 6.

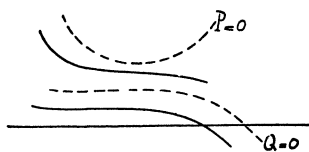


Fig. 6 bis.

d'après la forme de l'expression (26), des points de rebroussement de la ligne en question (fig. 6). Il résulte de ce qui précède que, pour ν tendant vers zéro, l'une des deux branches aboutissant à un tel rebroussement tendra vers un segment de l'axe des ξ et l'autre vers un arc de la courbe $R=0$.

Si, au contraire, les deux courbes $P=0$, $Q=0$ sont sans point commun dans la région considérée, la ligne $\Delta=0$ comprend (fig. 6 bis) deux branches entièrement séparées, de part et d'autre de $Q=0$.

Ces deux figures 6, 6 bis devraient, naturellement, être échangées l'une avec l'autre si l'on raisonnait sur l'équation (21); et, d'autre part, on ne doit, dans

chaque cas, en retenir que la partie utile, c'est-à-dire celle qui est située dans la région $\eta \geq 0$.

On formera d'ailleurs l'équation du second degré qui donne, dans le plan des $\xi \eta$ et au voisinage de l'origine, les deux ordonnées de la courbe $\Delta = 0$, en appliquant (par rapport à η) le théorème de factorisation à l'expression (26). Il est préférable d'introduire l'inconnue Q , en fonction de laquelle (25') permet de développer η : reportant dans (25), P prend la forme $P = M Q + N$, où M est une série entière en Q , pendant que

$$(27) \quad N = \left(\frac{c_1 h_0}{c_0} + h_1 \right) v - \frac{1}{3} \left(\frac{a \xi}{H} + h_2 v \right)^2 + \dots$$

est le terme indépendant de Q , c'est-à-dire la quantité obtenue en remplaçant η par sa valeur en fonction de ξ , v tirée de l'équation $Q = 0$. Dans l'expression $\Delta = 27 Q^2 + 4 (M Q + N)^3$, le théorème en question fera apparaître le facteur

$$(28) \quad Q^2 + s_1 Q + s_0$$

où s_0, s_1 contiennent respectivement N^3 et N^2 en facteur, le premier avec un coefficient initialement égal à 1.

Les abscisses des points de rebroussement sont donnés par l'équation $N = 0$, et il apparaît bien qu'elles sont réelles ou imaginaires suivant le signe du premier terme de (27).

21. La quantité (26) est toujours négative et la racine réelle unique lorsque les racines du polynôme du second degré (28) sont imaginaires.

Si, au contraire, les racines de ce polynôme et, par conséquent, les valeurs correspondantes η', η'' de η sont réelles, les racines de l'équation du troisième degré sont toutes trois réelles lorsque η est compris entre η' et η'' . Pour des valeurs de η satisfaisant à cette condition, nous aurons, conformément à la méthode de résolution classique, à introduire un angle ω par les deux relations concordantes

$$Q = 2 \left(\frac{-P}{3} \right)^{3/2} \cos \omega, \quad \sqrt{\Delta'} = 2 (-P)^{3/2} \sin \omega,$$

ce qui donne (en différentiant, comme nous savons en avoir le droit, sans tenir compte de la variabilité des coefficients)

$$\frac{d\omega}{d\nu} = \frac{3\sqrt[3]{3}}{\sqrt[3]{\Delta'}} \frac{dQ}{d\nu}$$

après quoi, les racines de l'équation seront

$$(29) \quad X = \sqrt[3]{\frac{-P}{3}} \cos \frac{\omega + 2j\pi}{3} \quad (j = 0, 1, 2).$$

Ici encore, la dérivée de P n'a pas à entrer en ligne de compte ; et, puisque la dérivée $\frac{dQ}{d\nu}$ est exactement ou sensiblement égale à 1, on voit que les dérivées des racines (29) sont de l'ordre de $\sqrt[3]{\frac{-P}{\Delta'}} \nu$, c'est-à-dire de

$$(30) \quad \frac{\sqrt[3]{-P}}{\sqrt[3]{(\eta - \eta')(\eta'' - \eta)}}$$

La contribution de cette région du plan des $\xi \eta$ (pour une valeur déterminée de ν) à l'intégrale destinée à représenter v_{02} s'obtient en multipliant par

$$K_2 d\xi d\eta$$

et intégrant : cette contribution est finie, comme on le voit sans même avoir besoin de tenir compte du numérateur $\sqrt[3]{-P}$ de l'expression (28) ni du facteur K_2 (que nous avons vu s'annuler pour $\xi = \eta = 0$), puisque l'intégrale $\int_{\eta'}^{\eta''} \frac{d\eta}{\sqrt[3]{(\eta - \eta')(\eta'' - \eta)}}$ a la valeur π . La convergence est, de plus, uniforme pour toutes les valeurs (suffisamment petites) de ν .

22. Dans la région $\Delta > 0$, nous aurons, pour notre équation du troisième degré, la racine unique

$$X = \sqrt[3]{-\frac{Q}{2} + \frac{1}{6\sqrt[3]{3}}\sqrt[3]{\Delta}} + \sqrt[3]{-\frac{Q}{2} - \frac{1}{6\sqrt[3]{3}}\sqrt[3]{\Delta}} = \sqrt[3]{T_1} + \sqrt[3]{T_2}$$

T_1, T_2 étant les racines de l'équation du second degré

$$T^2 + QT - \frac{P^3}{27} = 0.$$

L'un quelconque des deux termes ainsi écrits a pour dérivée (en vertu de l'équation précédente, et en continuant à ne considérer comme variable que le terme v explicitement décrit)

$$\frac{1}{3} T^{-\frac{2}{3}} \frac{dT}{dv} = \frac{1}{3} \frac{\sqrt[3]{T}}{2T+Q} = \pm \sqrt[3]{3} \frac{\sqrt[3]{T}}{\sqrt[2]{A}},$$

dérivée que nous aurons à intégrer soit dans l'intervalle $(0, \beta)$ (si les ordonnées η', η'' de la courbe $A=0$ sont imaginaires), soit dans la partie de cet intervalle extérieure à (η', η'') . Ici encore, il nous suffira de noter que la dérivée que nous venons de former ne contient d'autre dénominateur que $\sqrt[2]{A}$ ou, ce qui revient au même, la racine carrée du polynôme (28). Or la primitive, par rapport à η , ou (ce qui revient au même, à des facteurs bornés près) par rapport à Q , de

$$\frac{1}{\sqrt[2]{Q^2 + s_1 Q + s_0}}$$

est

$$\pm \log \left(Q + \frac{s_1}{2} \pm \sqrt{Q^2 + s_1 Q + s_0} \right)$$

expression où les signes se correspondent, mais peuvent toujours être choisis de manière que, dans l'argument du logarithme, les deux termes aient le même signe. Si l'on prend cette précaution, le logarithme portera sur une quantité bornée supérieurement (son maximum ayant lieu pour $\eta = \beta$ ou pour $\eta = 0$) et dont la borne inférieure ne sera autre que la racine carrée du discriminant D du polynôme (28). Or $D = \pm (s_1^2 - 4s_0)$ ne diffère que par un facteur unité de N^3 , de sorte que nous n'avons finalement à opérer que sur l'intégrale $\int \log. N d\xi$, laquelle est bien convergente, et cela uniformément pour toutes les valeurs de v .

De cette uniforme convergence, tant dans la région $A > 0$ que dans la région $A < 0$, il s'ensuit, comme on sait, que la différentiation par rapport à v , sous les signes de quadrature, est légitime, et c'est tout ce qui nous était nécessaire pour étendre au cas actuel les conclusions des nos 13—15 de *JM*.

23. Les remarques des nos 5—7 permettent de simplifier les considérations de *JM* (16) relatives aux termes obtenus en prenant, dans la for-

mule (F_1) , le terme (b) . On va voir que ces termes n'introduisent aucune difficulté nouvelle et, pour toute disposition de la figure, relèvent de ce qui a été dit ci-dessus.

Dans la formule (F) , le terme (a) est la dérivée, par rapport à v , de l'intégrale (5^{bis}) étendue au domaine fixe S_1 (le point $1'$ étant considéré comme défini par sa projection transversale sur S_1 et sa distance transversale v à S_1); le terme (b) est la dérivée de l'intégrale (5) , laquelle est également étendue à un domaine fixe si l'on définit le point $1'$ par v et les coordonnées ξ, η du n° 5 ou du n° 7. Moyennant ces choix de variables, les deux quantités à intégrer ne contiennent que des singularités indépendantes de v .

Dans les deux quantités en question, nous avons à remplacer u_1 ou $u_{1'}$, uniquement par la partie (b) de sa valeur (F_1) , laquelle est une dérivée par rapport à v_2 . Donc les deux termes considérés seront de la forme

$$(31) \quad (ab) = \frac{d}{dv} \iint \frac{\partial}{\partial v_2} \Phi_1(x, y, v, v_2) dx dy, \quad (bb) = \frac{d}{dv} \iint \frac{\partial}{\partial v_2} \Phi_2(\xi, \eta, v, v_2) d\xi d\eta,$$

les intégrales étant étendues à des domaines fixes et les différentiations se faisant dans les conditions classiques. Dès lors, l'interversion des différentiations, soit

$$(ab) = \frac{dJ_1}{dv_2}, \quad (bb) = \frac{dJ_2}{dv_2}$$

avec

$$J_1 = \frac{d}{dv} \iint \Phi_1(x, y, v, v_2) dx dy, \quad J_2 = \frac{d}{dv} \iint \Phi_2(\xi, \eta, v, v_2) d\xi d\eta$$

se fera également dans les conditions classiques. Or, les intégrales J_1, J_2 ne diffèrent de celles qui ont été traitées dans ce qui précède que par le changement de u'_2 en u_2 et de S_2 en S'_2 . La question est donc toute résolue à leur égard, et l'ensemble des termes (ab) , (bb) , (cb) donnera

$$-\frac{d}{dv_2} \iint u_2 v_{02} dS_2$$

v_{02} , étant l'expression (19) dans laquelle le point 2 est remplacé par $2'$.

Ce raisonnement s'applique tant au cas traité dans JM qu'à celui qui vient de nous occuper.

24. La même manière de procéder permet de combler une lacune que nous avons laissée subsister dans un précédent Mémoire des *Acta Mathematica* ³³⁾ consacré aux équations à quatre variables indépendantes. Nous avons résolu pour ces dernières, dans le travail dont il s'agit, le problème correspondant à celui que nous avons traité dans *JM*, c'est-à-dire que nous avons exprimé les parties constituantes $\mathfrak{D}_{02}$ et V_{02} de la solution élémentaire relative aux points 0 et 2 en fonction des quantités analogues $\mathfrak{D}_{01}$, V_{01} , $\mathfrak{D}_{12}$, V_{12} , moyennant intégration sur S_1 par rapport aux coordonnées du point 1. Mais nous avons à cet effet raisonné exclusivement sur les termes contenant les valeurs de la fonction φ , second membre de l'équation donnée, dans la région T_2 comprise entre S_1 et S_2 .

Les termes contenant les valeurs de u'_2 (ou encore de $L_2 u_2$) le long de S_2 donnent lieu à des calculs exactement identiques aux premiers.

Mais il n'en est pas de même de ceux qui contiennent les valeurs de u_2 (sans facteur L) et pour lesquels les opérations semblent au premier abord plus compliquées.

En réalité, ce dernier cas se ramène au précédent par une marche tout analogue à celle que nous venons d'indiquer pour les équations à trois variables. Il suffit, en ce qui concerne le passage de S_2 à S_1 , de prendre la formule de résolution du problème de Cauchy sous la forme (29^{bis}) de nos Leçons de Yale ³⁴⁾, soit, puisque $m = 4$, $m_1 = 2$,

$$2 \pi u_1 = (a) + (d) + (c') + (f') + (b') + (e');$$

$$(a) = - \iiint V_{12} \varphi dT, \quad (d) = \iiint V_{12} \varphi \frac{dT}{d\gamma},$$

$$(c') = - \iiint V_{12} (u'_2 + L_2 u_2) dS_2, \quad (f') = \iint V_{12} (u'_2 + L_2 u_2) \frac{dS_2}{d\gamma}$$

$$(b') = \frac{d}{d\gamma_2} \iiint V_{12'} u_2 dS_2, \quad (e') = - \frac{d}{d\gamma_2} \iint V_{12'} u_2 \frac{dS_2}{d\gamma}$$

Les termes (a) , (d) sont les mêmes que ceux qui figurent, avec la même notation, dans les formules (2) et (2') du Mémoire A_1 (Mémoire des *Acta Mathematica* qui vient d'être cité). Les termes (c') , (f') se déduisent

³³⁾ Tome XLIX, 1926.

³⁴⁾ N° 144, p. 311 de l'édition française. La notation des divers termes est aussi voisine que possible de celle de A_1 ; mais nous les avons rangés dans un ordre plus logique et mieux adapté au raisonnement actuel.

des précédents par le changement de φ en $u'_2 + L_2 u_2$, de T_2 en S_2 , des intégrations quadruples en intégrations triples et des intégrations triples en intégrations doubles. Mais, de plus, on passe également suivant une loi très simple de ces termes aux termes restants (b') , (e') : on doit remplacer $u'_2 + L_2 u_2$ par u_2 et, d'autre part, appliquer la différentiation d/dv_2 en dehors des signes de quadrature. Telle est l'expression u_1 à laquelle, pour arriver à u_0 , on doit appliquer, le long de S_1 , la formule de résolution, que l'on peut d'ailleurs également prendre sous la forme précédente, les termes (b') , (e') et aussi, comme précédemment, (c') et (f') (pour la partie qui ne contient pas L) seront des dérivées par rapport à v . Ici encore, les intégrales ainsi différenciées étant relatives à des domaines fixes si on applique à (b') et à (e') la méthode du n° 5 ou du n° 7. Dans ces conditions, si, après avoir formé les termes

$$(32) \quad (c'a), (c'd), (f'a), (f'd), (b'a), (b'd), (e'a), (e'd),$$

dont le calcul a été indiqué dans A_1 et, d'une manière tout identique, $(c'c')$, $(c'f')$, $(f'c')$, $(f'f')$, $(b'c')$, $(b'f')$, $(e'c')$, $(e'f')$, on passe aux termes

$$(c'b'), (c'e'), (f'b'), (f'e'), (b'b'), (b'e'), (e'b'), (e'e'),$$

ceux-ci seront de forme tout analogue à (31) ; et c'est en vertu des règles classiques que l'on pourra faire sortir la différentiation $\frac{d}{dv_2}$ des signes de quadrature, en l'intervertissant, au besoin, avec la différentiation par rapport à v .

Moyennant cette transformation, les expressions soumises à la différentiation en v_2 sont identiques aux termes (32), au changement près de $u'_2 + L_2 u_2$ en u_2 , et relèvent, par conséquent des mêmes calculs.

En un mot, la valeur de $\mathcal{V}_{02}$, fournie par la formule (14) (n° 19) du Mémoire cité A_1 et la valeur de V_{02} donnée aux n° 20 du même Mémoire permettent de construire la valeur u_0 , solution du problème de Cauchy, à partir des données portées par S_2 , absolument comme le permettaient $\mathcal{V}_{01}$, V_{01} à partir des données portées sur S_1 . C'est ce qui ne ressortait pas des calculs du Mémoire en question.

(Reçu le 8 juin 1932)

Les probabilités continues „en chaîne“

par Maurice FRÉCHET, Paris

Résumé

Le présent mémoire est un essai de mise au point en ce qui concerne la détermination exacte des champs de validité des résultats antérieurement connus de la théorie des probabilités continues dites « en chaîne ».

Sommaire

Résumé, Sommaire, Introduction ; page 175.

Position du problème: Diffusion, p. 176. Cas rectiligne p. 176, cas de l'espace p. 179; cas général, p. 179.

Intervention de la continuité: Continuité éventuelle des bornes, 181; Remarques, 182; Continuité éventuelle des $p^{(n)}(F)$, 183; ... des $P^{(n)}(E, F)$ 184.

Introduction de la régularisation: Définition du cas régulier, 187; exemple, 189; définition du cas quasi-régulier, 191; cas régulier, 194; une condition suffisante plus simple, 197; condition nécessaire pour le cas quasi-régulier, 200; simplifications, 201; condition suffisante, 201; nature de la convergence, 205; précisions, 206. Cas positivement régulier, 208. Condition de M. Hostinsky, 211. Convergence des moyennes arithmétiques, 214. Comportement des probabilités, 215.

Valeur de la densité limite $P(F)$: Cas quasi-régulier, 216. Cas régulier, 220. Cas de la densité-limite constante, 221. Cas du domaine infini, 224. Exemple de réalisations de (T') , 224. Expressions de la densité-limite, 225. Calcul de $s(E, F)$, 226. Remarques, 228.

Valeurs moyennes, fréquences et dispersion: Valeur moyenne d'une variable aléatoire, 228. Exemple, 233; effet de la condition (T') , 235; moyennes arithmétiques, 236. Fréquences moyennes, 236. Dispersion d'une variable aléatoire, 238; retour à l'exemple, 241. Dispersion des fréquences, 242. Références bibliographiques, 245.

Introduction

Markoff a appelé événements « en chaîne » des événements fortuits tels que, dans une suite d'épreuves, la réalisation de l'un d'eux dépende du résultat de l'épreuve précédente. Il a étudié particulièrement le cas où les événements possibles considérés sont en nombre fini. Divers auteurs,

— à sa suite, ou indépendamment, — ont étudié le même problème dans le cas où ces événements sont en nombre infini: dans le cas des probabilités continues. On trouvera un excellent résumé de leurs recherches dans un petit ouvrage dû à M. Hostinsky (I)¹).

Dans ces travaux, on s'est d'abord préoccupé, comme il est légitime, d'aller de l'avant et d'arriver aux applications. Toutefois, il nous a paru qu'il était maintenant utile de *consolider les résultats acquis en les revisant au point de vue de la rigueur mathématique, puis de procéder à leur extension*. Nous nous sommes consacrés à cette tâche dans notre cours du Premier Semestre 1931—32. Et nous publions ici une partie de ce cours.

L'étude critique des cas de validité des propriétés antérieurement énoncées nous a conduit à introduire des distinctions qui, malheureusement, ne simplifient pas l'exposition, mais qui serrent de plus près la vérité. Telles sont les notions de cas positivement régulier, de cas quasi-régulier. En ne limitant pas les recherches au cas des densités continues et des domaines bornés, on a aussi introduit des complications nouvelles, mais des complications auxquelles on ne peut échapper, même dans les applications.

Il n'est guère possible de résumer ici les résultats nouveaux contenus dans ce travail. On ne pourra s'en rendre compte qu'en le comparant en détail avec les travaux antérieurs sur le même sujet.

Position du problème

Diffusion. L'un des problèmes physiques où s'est présenté naturellement la conception des événements «en chaîne» est celui de la diffusion. On considère un liquide comme formé de molécules soumises à des chocs incessants. Une première approximation consiste à admettre que la probabilité de la position B de la molécule après un choc ne dépend que de B et de la position A qu'avait la molécule lors du choc précédent. Lorsque le nombre des positions possibles de la molécule n'est pas supposé limité, il faut faire intervenir au lieu du point B , un intervalle, de la façon que nous allons préciser.

Cas rectiligne. Prenons pour commencer le cas simple d'une molécule dont le mouvement est rectiligne. Nous supposerons qu'il y a une probabilité déterminée pour qu'une molécule occupant une abscisse x vienne

¹) On a reporté à la fin du mémoire les titres complets des publications distinguées dans le texte par des chiffres romains.

prendre, après n chocs, une position comprise dans un intervalle déterminé (y', y'') . Nous admettrons que cette probabilité ne dépend que des positions initiale et finale — c'est-à-dire de x, y', y'' — et du nombre de chocs qui ont produit ce déplacement, indépendamment du nombre et de l'effet des chocs qui l'ont précédé. La probabilité en question devrait se mettre sous la forme d'une intégrale de Stieltjes étendue à l'intervalle $y' y''$. Mais en abordant le cas des probabilités continues, on rencontre plusieurs difficultés qui ne se présentent pas dans la régularisation des probabilités discontinues. Les intégrales qui représentent des probabilités nécessairement finies peuvent s'étendre à des régions illimitées et à des fonctions infinies en certains points. L'étude de la régularisation étant elle-même assez récente, il sera peut-être préférable de ne pas aborder toutes les difficultés à la fois; nous nous contenterons donc de l'étudier dans le cas où les probabilités en question ont des densités de probabilité généralement finies. Plus précisément, nous supposerons que la probabilité ci-dessus se représente par une intégrale ordinaire. Son élément différentiel dépendra naturellement de x , de n et de la variable d'intégration y . En résumé, nous supposerons que la probabilité en question soit représentable sous la forme

$$\int_{y'}^{y''} P^{(n)}(x, y) dy.$$

On exprime ce fait d'une manière brève, mais peu rigoureuse, en disant que la probabilité élémentaire pour que la molécule passe, après n chocs, de l'abscisse x à une abscisse comprise entre y et $y + dy$ est $P^{(n)}(x, y) dy$. On aura évidemment $P^{(n)}(x, y) \geq 0$. Et, puisque après n épreuves il est certain que la molécule partant de x se trouvera quelque part, on a

$$\int_{-\infty}^{+\infty} P^{(n)}(x, y) dy = 1.$$

Dans le cas particulier où la molécule reste dans un vase de dimensions finies, y devra rester dans un intervalle fixe (a, b) et on pourra se contenter d'écrire

$$\int_a^b P^{(n)}(x, y) dy = 1.$$

Pour simplifier les notations on posera

$$p(x, y) = P^{(1)}(x, y).$$

Pour aller de la position x à une position située dans l'intervalle $y, y + dy$ à la $n + m^{\text{ème}}$ épreuve, il faut aller, à la $n^{\text{ème}}$ épreuve, à une certaine position qu'on peut situer dans un intervalle $u, u + du$ et, de là, à la position située entre y et $y + dy$ à la $n + m^{\text{ème}}$ épreuve. Plaçons-nous dans l'hypothèse où ce dernier déplacement a la même probabilité que si les deux positions dernières successives étaient prises avant la première épreuve et à la $m^{\text{ème}}$. Alors, si l'on suppose en outre que les $P^{(n)}(x, y)$ sont des fonctions continues de x et de y , on aura, en faisant varier seulement u ,

$$P^{(m+n)}(x, y) dy = \int_{-\infty}^{+\infty} P^{(n)}(x, u) du P^{(m)}(u, y) dy$$

en vertu du théorème des probabilités composées; d'où la relation d'itération

$$P^{(m+n)}(x, y) = \int_{-\infty}^{+\infty} P^{(n)}(x, u) P^{(m)}(u, y) du.$$

On a, en outre, pour chaque entier n

$$\int_{-\infty}^{+\infty} P^{(n)}(x, y) dy = 1$$

puisque la molécule partant de x doit bien arriver en quelque position y . On voit que, connaissant $p(x, y)$, la relation d'itération permet, en prenant $m = 1$, avant et après y avoir permuté m et n , de calculer successivement tous les $P^{(n)}(x, y)$ par l'une ou l'autre des formules

$$P^{(n+1)}(x, y) = \int_{-\infty}^{+\infty} P^{(n)}(x, u) p(u, y) du$$

$$P^{(n+1)}(x, y) = \int_{-\infty}^{+\infty} p(x, u) P^{(n)}(u, y) du.$$

Cas de l'espace. Plus généralement, supposons que la molécule puisse se déplacer dans l'espace à trois dimensions, alors on fera intervenir la probabilité élémentaire $P^{(n)}(A, B) dv_B$ pour que la molécule primitivement en A se trouve à la $n^{\text{ième}}$ épreuve dans un volume dv_B entourant le point B . Plus précisément, on étudiera le cas où la probabilité pour que la molécule, primitivement en A , se trouve à la $n^{\text{ième}}$ épreuve dans un volume R est représentée par une intégrale triple de la forme

$$\int_R P^{(n)}(A, B) dv_B.$$

Cas général. On peut encore généraliser et considérer le cas d'un système dépendant de k paramètres qui ne sont plus nécessairement de nature géométrique, mais dont la connaissance définit l'état du système. On pourra désigner par E et F deux états du système et par v une région de l'espace à k dimensions. Et l'on pourra étudier le cas où la probabilité pour que le système passe, en n épreuves, de l'état E à l'un quelconque des états F appartenant à v est représentée par une intégrale multiple de la forme

$$\int_v P^{(n)}(E, F) d\tau_F$$

c'est-à-dire de la forme

$$\int_v \Phi^{(n)}(q_1, q_2, \dots, q_k; u_1, u_2, \dots, u_k) du_1 \dots du_k$$

$q_1, \dots, q_k; u_1, \dots, u_k$ désignant les paramètres qui définissent respectivement E et F . En raisonnant comme plus haut, on voit qu'on devra avoir

$$\int \dots \int_v \Phi^{(n)}(q_1, \dots, q_k; u_1, \dots, u_k) du_1 \dots du_k = 1$$

ou, sous une forme plus condensée

$$(T) \quad \int_v P^{(n)}(E, F) d\tau_F = 1$$

et de même

$$(I) \quad P^{(m+n)}(E, F) = \int_v P^{(n)}(E, G) P^{(m)}(G, F) d\tau_G$$

en désignant par V la région de l'espace à k dimensions formée par l'ensemble des états possibles du système (V peut être bornée ou illimitée). Enfin, on aura évidemment la condition $p(E, F) \geq 0$, d'où, d'après (I), $P^{(n)}(E, F) \geq 0$.

Puisque les $P^{(n)}(E, G)$ sont ≥ 0 et vérifient (T), $P^{(m+n)}(E, F)$ est une moyenne « pondérée » des $P^{(m)}(G, F)$. Si donc on nomme $P^{(m)}(F)$ et $p^{(m)}(F)$ les bornes finies ou non de $P^{(m)}(G, F)$ quand G parcourt V , on aura

$$0 \leq p^{(m)}(F) \leq p^{(m+n)}(F) \leq P^{(m+n)}(F) \leq P^{(m)}(F).$$

Lorsque n croît indéfiniment $p^{(n)}(F)$ et $P^{(n)}(F)$ ont des limites déterminées, finies ou non, $p(F)$ et $P(F)$ et l'on a

$$0 \leq p(F) \leq P(F).$$

Nous pourrions prendre pour V un ensemble très général. Nous supposerons seulement que ce soit un *domaine*, entendant par là que V est formé par la réunion des états intérieurs à V et de leurs états limites. C'est ce qui a lieu pour toutes les figures simples. Les deux conséquences qui nous seront utiles par la suite sont les suivantes. D'une part, V sera « fermé » au sens de la théorie des ensembles c'est-à-dire comprendra tous les états limites d'une suite d'états de V . D'autre part, pour tout état E de V , et quel que soit $\eta > 0$, il existe un état F intérieur à V et à distance ²⁾ de $E < \frac{\eta}{2}$. Donc, il y aura une sphère v de centre F et de rayon assez petit pour qu'elle soit formée uniquement d'états de V , tous à distance de E inférieure à η . En appelant *mesure* de V l'intégrale $\int_V d\tau$, on voit ainsi que pour tout état E de V (même pour un état appartenant à la frontière de V) et pour tout nombre $\eta < 0$, il existe une partie v de V à distance de E de moins de η et dont la mesure est positive.

Nous allons maintenant étudier les conséquences qu'on peut déduire de l'hypothèse particulière que l'une des densités itérées $P^{(n)}(E, F)$ est uniformément continue.

²⁾ Si E, F sont déterminés respectivement par les paramètres $q_1, \dots, q_k; u_1, \dots, u_k$, on pourra appeler distance de E à F la quantité $\sqrt{(q_1 - u_1)^2 + \dots + (q_k - u_k)^2}$, par exemple.

Intervention de la continuité.

Continuité éventuelle des bornes. Faisons d'abord un raisonnement général s'appliquant aux bornes inférieure $\varphi(F)$ et supérieure $\Phi(F)$, quand F est fixe, d'une fonction $\Phi(E, F)$ uniformément continue quand E, F varient simultanément sur V , et partout ≥ 0 sur V .

Alors $\varphi(F)$ est partout finie et ≥ 0 et on a

$$\Phi(E, F) \geq [\Phi(E, F) - \Phi(E, F_1)] + \varphi(F_1).$$

Or, pour tout $\varepsilon > 0$ il existe un nombre $\eta > 0$ et indépendant de E tel que

$$|\Phi(E, F) - \Phi(E, F_1)| < \varepsilon \text{ pour } FF_1 < \eta.$$

On a donc

$$\Phi(E, F) \geq -\varepsilon + \varphi(F_1)$$

et puisque le second membre est indépendant de E

$$\varphi(F) \geq -\varepsilon + \varphi(F_1).$$

$$\text{On a de même } \varphi(F_1) \geq -\varepsilon + \varphi(F)$$

et finalement

$$|\varphi(F) - \varphi(F_1)| \leq \varepsilon \text{ pour } FF_1 < \eta.$$

Ainsi, que V soit borné ou non, $\varphi(F)$ est uniformément continue sur V . Et même, puisque tout couple ε, η convenant pour $\Phi(E, F)$ convient à $\varphi(F)$, on voit que $\varphi(F)$ est, en ce sens, au moins aussi continu que $\Phi(E, F)$.

Le résultat correspondant pour $\Phi(F)$ est moins simple, car $\Phi(F)$ peut être infini. Toutefois, si l'on suppose $\Phi(E, F)$ non seulement uniformément continu, mais encore borné sur V , des raisonnements analogues aux précédents montrent que $\Phi(F)$ sera aussi borné et uniformément continu sur V . C'est, en particulier, ce qui aura lieu pour toute fonction $\Phi(E, F) \geq 0$ et continue quand E et F varient sur un domaine fini V .

En passant au cas général, supposons qu'il y ait au moins un état F_0 où $\Phi(F_0)$ soit fini. Comme on a, pour $FF_0 < \eta$

$$\Phi(E, F) = [\Phi(E, F) - \Phi(E, F_0)] + \Phi(E, F_0) \leq \varepsilon + \Phi(F_0)$$

quel que soit E , on aura

$$\Phi(F) \leq \varepsilon + \Phi(F_0)$$

et en particulier $\Phi(F)$ sera aussi fini.

Ainsi, étant donnés deux états F, F_0 de V , à distance $< \eta$, $\Phi(F)$ et $\Phi(F_0)$ sont en même temps infinis, ou bien tous deux finis et on a

$$|\Phi(F) - \Phi(F_0)| < \varepsilon.$$

Dans la partie V' de V où $\Phi(F)$ est fini, $\Phi(F)$ est donc uniformément continu. On peut d'ailleurs observer que si U est une partie de V qui est d'un seul tenant, U appartient entièrement à V' ou appartient entièrement à $V - V'$. Par hypothèse, si F, F' appartiennent à U , alors, à tout $\eta > 0$ correspond un nombre fini d'états $F_1, F_2, \dots, F_s$ de U tels de $FF_1, F_1F_2, \dots, F_sF'$ soient tous $< \eta$. Si donc, par exemple, F appartient à V' , alors $\Phi(F)$ est fini, donc $\Phi(F_1)$ est fini; $\Phi(F_1)$ étant fini, il en est de même de $\Phi(F_2)$, etc. Finalement $\Phi(F')$ est aussi fini.

Remarques. I. On peut observer, en suivant la démonstration de plus près que le résultat subsiste si on remplace l'uniforme continuité de $\varphi(E, F)$ quand E, F varient simultanément par une condition moins stricte, « l'égale continuité » des fonctions de F , $\varphi(E, F)$ qui correspondent chacune à un état E déterminé. Autrement dit, il suffit de supposer qu'à tout $\varepsilon > 0$, correspond η tel que

$$|\varphi(E, F) - \varphi(E, F_0)| < \varepsilon \text{ pour } FF_0 < \eta$$

E variant sur V . Cette remarque sera utile plus loin (p. 183).

II. Supposons qu'on sache que $\Phi(F)$ est fini presque partout sur V , circonstance qui se présentera souvent plus loin. Alors $V - V'$ est vide ou de mesure nulle: Si $V - V'$ n'était pas vide, il y aurait un état F appartenant à $V - V'$ et pour toute valeur de $\eta > 0$, la sphère de centre F , rayon η , contiendrait comme on l'a vu p. 180 une sphère appartenant à V et par suite n'appartenant pas entièrement à $V - V'$ qui est de mesure nulle.

Il y aurait donc un état F_0 de V où $\Phi(F_0)$ est fini et à distance $< \eta$ de F . Par suite, $\Phi(F)$ serait fini contrairement à l'hypothèse. Ainsi dans ce cas $V' \equiv V$ c'est-à-dire que: si $\Phi(F)$ est fini presque partout sur V , $\Phi(F)$ est fini partout sur V et y est uniformément continu.

Un cas intéressant, que nous rencontrerons souvent par la suite est celui où $\varphi(E, F)$ est ≥ 0 et est majoré, quels que soient E, F sur V , par une fonction $\psi(F)$ sommable sur V . Alors, l'ensemble sur lequel $\psi(F)$ est infini, est vide ou de mesure nulle. Il en est donc de même de $\Phi(F)$. Par suite, $\Phi(F)$ est fini et uniformément continu sur tout le domaine V .

Continuité éventuelle des $p^{(n)}(F)$. Supposons que l'une au moins $P^{(m)}$ des fonctions $P^{(n)}(E, F)$ soit uniformément continue. Ou même, plus généralement, supposons que les fonctions de F , $P^{(m)}(E, F)$ correspondant respectivement aux divers états E soient « également continues » (p. 182) en F . Alors, à tout ε positif correspond un nombre η tel que $|P^{(m)}(E, F) - P^{(m)}(E, F_1)| < \varepsilon$ pour $FF_1 < \eta$, E variant sur V , ε et η étant indépendants de E . Alors

$$|P^{(s+m)}(E, F) - P^{(s+m)}(E, F_1)| \leq \int_V |P^{(s)}(E, G)| |P^{(m)}(G, F) - P^{(m)}(G, F_1)| d\tau_G \leq \varepsilon$$

pour $FF_1 < \eta$.

On voit que pour E, n fixes, $P^{(n)}(E, F)$ sera une fonction de F uniformément continue et que les fonctions de la famille de fonctions de F obtenue en faisant varier E et n ($\geq m$) seront aussi « également continues ». Il en résulte, d'après la p. 182, que l'on aura

$$|p^{(n)}(F) - p^{(n)}(F_1)| < \varepsilon$$

pour $n > m$ et $FF_1 < \eta$. En passant à la limite pour F, F_1 fixes, on aura aussi

$$|p(F) - p(F_1)| < \varepsilon.$$

En résumé: que le domaine V soit fini ou non, il suffit que l'une des probabilités itérées $P^{(n)}(E, F)$ soit uniformément continue en E, F sur V , pour que les $p^{(n)}(F)$, à partir d'un certain rang, et leur limite $p(F)$ soient chacune uniformément continues sur V et même y soient dans leur ensemble « également continues ».

En ce qui concerne les bornes supérieures $P^{(n)}(F)$, le résultat précédent subsistera si l'une au moins des probabilités $P^{(n)}(E, F)$ qui est supposée uniformément continue est en même temps bornée sur V et alors les $P^{(n)}(F)$ et $P(F)$ seront même « également » bornées à partir d'un certain rang. Il sub-

sistera aussi quand l'une des probabilités $P^{(n)}(E, F)$ sans être bornée est majorée par une fonction $\psi(F)$ sommable sur V .

Enfin, dans le cas général, soit V_r' l'ensemble sur lequel $P^{(r)}(F)$ est fini et V' l'ensemble sur lequel $P(F)$ est fini: V_r' appartient à V' . Si $F, F_1 < \eta$ et si F, F_1 appartiennent à V' , alors pour r assez grand F, F_1 appartiennent à V_r' . Or pour $r > m$, les $P^{(n)}(E, F)$ sont des fonctions de F telles que

$$|P^{(r)}(E, F) - P^{(r)}(E, F_1)| < \varepsilon.$$

D'après ce qu'on a vu, p. 182

$$|P^{(r)}(F) - P^{(r)}(F_1)| < \varepsilon$$

pour r assez grand et par suite, à la limite,

$$|P(F) - P(F_1)| \leq \varepsilon.$$

Ainsi, dans le cas général, $P^{(r)}(F)$ est uniformément continu sur V_r' et $P(F)$ est uniformément continu sur tout l'ensemble V' où $P(F)$ est fini.

Dans le cas particulier où $P(F)$ est fini presque partout, soient F, F_1 deux états quelconques de V , à distance $< \eta$. Alors, si $P(F_1)$ est fini, il en est de même de $P^{(r)}(F_1)$ pour r assez grand. $P^{(r)}(E, F)$ étant, en supposant $r > m$ «également continu» en F quand E varie, il en résulte d'après la p. 182 que $P^{(r)}(F)$ est aussi fini et par suite que $P(F)$ est aussi fini. Ainsi V' contient tout état F à distance $< \eta$ d'un état F_1 de V' et $V - V'$ est de mesure nulle. D'après le raisonnement de la p. 182 il en résulte que $V - V'$ est vide ou que $V' \equiv V$. En résumé, si l'une des fonctions $P^{(n)}(E, F)$ est uniformément continue quand E, F varient sur V , alors, que V soit limité ou non, si l'on est certain que $P(F)$ est fini presque partout sur V , on peut affirmer que $P(F)$ est fini et uniformément continu partout sur V . Nous savions déjà qu'il en est de même pour $p(F)$.

Continuité éventuelle des $P^{(n)}(E, F)$. Par contre, le raisonnement précédent ne prouve pas que si l'une des fonctions $P^{(n)}(E, F)$ est uniformément continue sur V , il en soit de même pour les autres à partir d'un certain rang. Il y a cependant des cas simples où cette conclusion est légitime.

Par hypothèse, pour tout $\varepsilon > 0$ il existe un nombre η tel que pour $EE_1 < \eta$, $FF_1 < \eta$ on ait

$$\begin{aligned} & |P^{(m)}(E_1, F_1) - P^{(m)}(E, F)| < \varepsilon. \text{ On a donc} \\ & |P^{(s+m)}(E_1, F_1) - P^{(s+m)}(E, F)| \leq \\ & \int_V P^{(s)}(E_1, G) |P^{(m)}(G, F_1) - P^{(m)}(G, F)| d\tau_G \\ & + \int_V |P^{(m)}(E_1, G) - P^{(m)}(E, G)| P^{(s)}(G, F) d\tau_G. \end{aligned}$$

Si $P^{(s)}(G, F)$ est une fonction de G sommable sur V , le second membre est $< \varepsilon [1 + \int_V P^{(s)}(G, F) d\tau_G]$.

Nous aurons alors plusieurs cas :

I. Si V est borné, alors $P^{(m)}(E, F)$ supposée continue sur V y a une borne supérieure finie μ et il en est de même pour $P^{(s)}(E, F)$ pour $s \geq m$. On a donc pour $n \geq 2m$, $EE_1 < \eta$, $FF_1 < \eta$.

$$|P^{(n)}(E, F) - P^{(n)}(E_1, F_1)| < \varepsilon [1 + \mu \text{ mes. } V].$$

Ainsi, quand le domaine V est borné il suffit que l'une des fonctions $P^{(n)}(E, F)$ soit continue quand E, F parcourent V — et par suite uniformément continue — pour que toutes ces fonctions soient à partir d'un certain rang, chacune uniformément continue et même, dans leur ensemble, «également continues».

II. Ne supposons plus le domaine V borné; alors si l'une au moins des fonctions $P^{(n)}(E, F)$ est uniformément continue, et si, en outre, l'intégrale $\int_V P^{(n)}(G, F) d\tau_G$ est, pour n assez grand, inférieure à un nombre indépendant de F et de n , la conclusion précédente subsiste. La deuxième condition sera satisfaite en particulier dans le cas où est vérifiée une certaine condition T_1' dont nous reconnâtrons plus loin (p. 222) l'importance.

Cette dernière condition sera aussi satisfaite dans le cas où $P^{(m)}(E, F)$ est non seulement uniformément continu mais encore borné supérieurement sur V et même avec une borne inférieure positive a . Alors, on a

$$P^{(n+m)}(E, F) = \int_V P^{(m)}(E, G) P^{(n)}(G, F) d\tau_G > a \int_V P^{(n)}(G, F) d\tau_G.$$

Or, puisque $P^{(m)}(E, F)$ a une borne supérieure M il en est de même de $P^{(n+m)}(E, F)$ et l'on a

$$\mathcal{J} = \int_V P^{(n)}(G, F) d\tau_G < \frac{M}{a}$$

et ceci quels que soient n et F . Il en résulte encore que $P^{(m)}(E, F)$, $P^{(m+1)}(E, F)$, ... seront des fonctions «également continues» et «également» bornées.

Si l'intégrale $\mathcal{F} = \int_V P^{(s)}(G, F) d\tau_G$ a, pour s assez grand, une borne simplement indépendante de F , mais variable avec s et si $P^{(m)}(G, F)$ est seulement supposée uniformément continue sur V , alors chacune des fonctions $P^{(n)}(E, F)$ sera, pour n assez grand, uniformément continue sur V .

Enfin, dans le cas plus général où $P^{(m)}(G, F)$ étant uniformément continue sur V , on suppose simplement que, quel que soit s assez grand $P^{(s)}(G, F)$ est une fonction de G sommable sur V quand F est en F_0 , alors $P^{(n)}(E, F)$ sera simplement, pour n assez grand, une fonction continue de E et de F pour le couple d'états E_0, F_0 où E_0 est arbitraire.

Notons d'ailleurs que, dans bien des cas, l'étude directe des fonctions $P^{(n)}(E, F)$ permettra d'établir leur continuité, à partir d'un certain rang, sans recourir aux propositions précédentes. Tel est le cas de l'exemple de la p. 189.

Introduction de la régularisation

On a
$$p^{(m)}(F) \leq P^{(m)}(E, F) \leq P^{(m)}(F).$$

Les bornes $p^{(m)}(F)$ et $P^{(m)}(F)$ ne peuvent jamais (p. 180) s'éloigner l'une de l'autre quand m croît. En général, elles vont se rapprocher quand le nombre m des épreuves augmente. En sorte qu'en général l'accroissement du nombre des épreuves a pour effet de diminuer l'amplitude des oscillations possibles — quand on envisage plusieurs positions possibles de E — des densités de probabilités. C'est en cela que consiste, à un premier point de vue, la régularisation des probabilités. L'étendue limite de l'oscillation sera mesurée par $P(F) - p(F)$. Il est d'ailleurs clair que cette régularisation sera d'autant plus marquée que $P(F) - p(F)$ qui est ≥ 0 sera plus petit et qu'elle atteindra son maximum quand $P(F) - p(F)$ sera nul.

C'est la recherche des cas où cette circonstance se produira qui va nous occuper maintenant.

Supposons d'abord que $p(E, F)$ ou plus généralement l'un, $P^{(m)}(E, F)$, des $P^{(n)}(E, F)$ a une borne supérieure μ indépendante de E et de F , au moins quand ceux-ci varient sur V . Alors $P^{(m)}(F) \leq \mu$ et $P^{(n)}(E, F) \leq \mu$ pour $n > m$ et enfin $P(F) \leq \mu$.

Notons, d'autre part, que la loi de formation des $P^{(n)}(E, F)$ définie par la condition (I) de la page 179, est la même que celle des noyaux itérés de l'équation intégrale de Fredholm

$$X(M) = f(M) + \int_V p(E, F) X(F) d\tau_F.$$

La théorie de cette équation sera utilisée plusieurs fois par la suite. Dès maintenant, nous pourrions y renvoyer pour la démonstration de plusieurs propriétés. Il faudra toutefois observer que les démonstrations en sont faites, généralement en supposant V borné et que leur extension au cas d'une région V illimitée n'est pas toujours immédiate. Elle conduirait même parfois à des énoncés inexacts si ceux-ci n'étaient pas convenablement appropriés à ce cas. Par exemple on ne pourra pas, dans ce cas, considérer toute fonction continue comme bornée, intégrable et uniformément continue; et une fonction uniformément continue dans une région illimitée n'y est nécessairement ni bornée, ni intégrable. Par contre, quand la région V est bornée, on sait (II, p. 343) que si le noyau $p(E, F)$ y est continu, il en est de même de tous les noyaux itérés $P^{(n)}(E, F)$. Cette propriété qui va nous servir immédiatement a été établie plus haut, p. 185, comme cas particulier de propositions plus générales.

Définition du cas régulier. Plaçons-nous d'abord, pour simplifier, dans les *hypothèses suivantes*: la fonction $p(E, F)$ est uniformément continue quand E et F varient simultanément dans V et la région V est bornée. Alors $p(E, F)$ a une borne supérieure μ et les fonctions $P^{(n)}(E, F)$ sont uniformément continues et au plus égales à μ . Du fait que $P^{(n)}(E, F)$ est uniformément continue par rapport à l'ensemble de E, F , résulte que $P^{(n)}(F)$ et $p^{(n)}(F)$ sont aussi uniformément continues et d'ailleurs $\leq \mu$. Ceci étant les inégalités

$$\dots p^{(n)}(F) \leq p^{(n+1)}(F) \leq \dots \leq p(F) \leq P(F) \leq \dots \leq P^{(n+1)}(F) \leq P^{(n)}(F) \leq \dots$$

$$(I) \quad 0 \leq |P^{(n)}(E, F) - p(F)| \leq |P^{(n)}(F) - p^{(n)}(F)|$$

$$\text{et} \quad P(F) - p(F) = \lim_{n \rightarrow \infty} [P^{(n)}(F) - p^{(n)}(F)]$$

montrent que, dans le cas où on aurait $p(F) = P(F)$ la suite des noyaux itérés $P^{(n)}(E, F)$ aurait une limite déterminée $p(F)$ indépen-

dante de E . Ce serait là une circonstance bien remarquable : la densité de probabilité du passage en n épreuves de l'état E à l'état F deviendrait, pour n croissant, de plus en plus indépendante de l'état initial E .

L'égalité

$$P(F) = p(F)$$

se traduit par l'égalité

$$\lim_{n \rightarrow \infty} [P^{(n)}(F) - p^{(n)}(F)] = 0.$$

Dans les hypothèses actuelles la convergence du crochet est nécessairement uniforme. Car il en est ainsi pour toute suite non croissante et convergente de fonctions $\varphi^{(n)}(F)$ continues sur une région V bornée (et, comme on le supposera naturellement toujours ici, fermée, c'est-à-dire comprenant ses éléments limites).

En effet si $\varphi(F)$ est la limite de $\varphi_n(F)$, si l'on pose $D_n(F) = \varphi_n(F) - \varphi(F)$ et si on désigne par M_n le maximum de $D_n(F)$ sur V , M_n ne peut croître quand n croît. Donc M_n a une limite $M \geq 0$. Il s'agit de montrer que $M = 0$. Dans le cas contraire, l'ensemble S_n des états F de V où $D_n(F) \geq \frac{M}{2} > 0$ comporterait au moins un état F_n . De la suite des F_n on pourrait tirer une suite $F_{n_1}, F_{n_2}, \dots$ convergeant vers un état F' de V . Comme $F_{n_p}, F_{n_p+1}, \dots$ appartiendraient à S_{n_p} , F' appartiendrait à S_{n_p} ; on aurait donc $D_{n_p}(F') \geq \frac{M}{2} > 0$ alors que la suite des $D_m(F')$ tend vers 0.

Dans ces conditions, en vertu de $\bullet(1)$, $P^{(n)}(E, F)$ converge uniformément vers $p(F)$ quand E, F varient indépendamment sur V . Par suite, on a

$$\int_V P(F) d\tau_F = \int_V p(F) d\tau_F = \lim_{n \rightarrow \infty} \int_V P^{(n)}(E, F) d\tau_F = 1.$$

Réciproquement, si, dans les mêmes hypothèses sur $p(E, F)$ et V , les intégrales de $P(F)$ et de $p(F)$ sont égales à l'unité, on aura

$$\int_V [P(F) - p(F)] d\tau_F = 0$$

et par suite, la fonction $P(F) - p(F)$ étant continue et ≥ 0 sur V y sera nécessairement nulle partout.

Car, si $\psi(F)$ est une fonction continue dans une région V , bornée ou non telle que $\int_V \psi(F) dF = 0$ avec $\psi(F) \geq 0$ et s'il existait un état F_0 de V où $\psi(F_0) > 0$, alors, il existerait (p. 180) une région v de mesure positive, appartenant à V et toute entière assez voisine de F_0 pour que l'on ait sur v , $\psi(F) \geq \frac{\psi(F_0)}{2}$, d'où $0 = \int_V \psi(F) dF \geq \int_v \psi(F) dF \geq \frac{\psi(F_0)}{2} \int_v dF \geq 0$ d'où $\int_v dF = 0$, contrairement à la définition de v .

Alors, la suite non croissante de fonctions continues $P^{(n)}(F) - p^{(n)}(F)$ qui converge uniformément sur V vers $P(F) - p(F)$ convergera vers 0 et finalement $P^{(n)}(E, F)$ converge uniformément sur v vers une limite $p(F)$ indépendante de F .

Dans les hypothèses actuelles, les trois conditions suivantes sont donc équivalentes :

1° $P^{(n)}(E, F)$ converge uniformément quand E et F varient sur V vers une fonction limite indépendante de E .

2° $p(F) = P(F)$ sur V .

3° $\int_V p(F) d\tau_F = \int_V P(F) d\tau_F = 1$.

L'équivalence des trois propriétés ne subsiste plus dans le cas général comme va le montrer l'exemple suivant. Nous serons donc amené à ne retenir que l'une de ces conditions pour définir le cas régulier dans les hypothèses les plus générales

Exemple. Supposons le système matériel défini par un seul paramètre numérique et revenons aux notations de la p. 178. Considérons le cas particulier où $p(x, y) = \frac{1}{\sqrt{\pi}} e^{-(x-y)^2}$. On aura bien

$$p(x, y) \geq 0, \int_{-\infty}^{+\infty} p(x, y) dy = 1.$$

Or la relation d'itération fournit aisément l'expression

$$P^{(n)}(x, y) = \frac{1}{\sqrt{n\pi}} e^{-\frac{(x-y)^2}{n}}$$

que d'ailleurs on vérifie plus facilement encore.

Les bornes $P^{(n)}(y)$ et $p^{(n)}(y)$ quand x varie, de $P^{(n)}(x, y)$ sont évidemment

$$P^{(n)}(y) = \frac{1}{\sqrt{n\pi}} \text{ et } p^{(n)}(y) = 0$$

de sorte que $\int_{-\infty}^{+\infty} p^{(n)}(y) dy = 0$, et $\int_{-\infty}^{+\infty} P^{(n)}(y) dy$ est infinie quel que soit n .

Pourtant $P^{(n)}(y)$ et $p^{(n)}(y)$ tendent vers la même limite $P(y) = p(y) = 0$. Et même la convergence de l'un et de l'autre est uniforme. Dans ce cas $P^{(n)}(x, y)$ converge vers une limite $p(y) = 0$ et cela uniformément quand x, y varient indépendamment. Et les intégrales $\int_{-\infty}^{+\infty} P(y) dy$ et $\int_{-\infty}^{+\infty} p(y) dy$ sont toutes deux finies et égales. Mais leur valeur commune n'est pas l'unité.

En revenant au cas d'un nombre quelconque de variables, on voit que cet exemple fournit l'apparent paradoxe suivant. Il peut arriver : que $P^{(n)}(E, F)$ converge vers une limite indépendante de E et cela uniformément quand E et F varient arbitrairement sur V , qu'en outre $P(F) = p(F)$, qu'enfin les intégrales $\int_V p(F) d\tau_F$ et $\int_V P(F) d\tau_F$ existent et soient égales et cependant que leur valeur commune soit $\neq 1$ c'est-à-dire que

$$\int_V P(F) d\tau_F \neq \lim_{n \rightarrow \infty} \int_V P^{(n)}(E, F) d\tau_F.$$

Naturellement cet ensemble de circonstances ne peut se présenter que si V est illimité.

Ainsi donc, dans le cas général les conditions 1°, 2°, ne sont plus équivalentes à 3°.

Or, si la condition 1° est celle qui semble fournir dans les applications la propriété la plus importante, son utilité serait bien diminuée si 3° n'était pas réalisée. Car, d'après 1° il y aurait bien une fonction $p(F)$ qui serait limite de densités de probabilités. Mais c'est seulement si 3° était réalisée que cette limite serait elle-même une densité de probabilité (bien entendu, on a toujours $p(F) \geq 0$, $P(F) \geq 0$).

Or, dans le cas général, nous allons voir que la condition 3° a pour conséquence des propriétés qui sont presque les mêmes que 1° et 2° et qui conservent tout l'essentiel de l'utilité de la condition 1° pour les

applications. Cela nous permettra de définir le cas régulier — quand on n'impose à $p(E, F)$ et à V que les conditions (P) , (T) , (I) — comme celui où 3° est réalisé.

Commençons d'abord par une observation qui s'applique qu'on soit ou non dans le cas régulier.

Si $p(E, F)$ n'est pas borné, alors, comme on suppose toujours $\int_V P^{(n)}(E, F) d\tau_F = 1$, la fonction $p^{(n)}(F)$ qui est ≥ 0 et $\leq P^{(n)}(E, F)$ sera aussi une fonction *sommable* sur V (c'est-à-dire dont l'intégrale sur V est déterminée et finie) telle que $\int_V p^{(n)}(F) d\tau_F \leq 1$. Et, puisque $p^{(n)}(F)$ tend sans décroître vers $p(F)$ (finie ou non), $p(F)$ est aussi sommable sur V et $\int_V p(F) d\tau_F \leq 1$ (III, p. 120). Par suite, non seulement $p(F)$ ne peut être infini partout, mais $p(F)$ ne peut être infini que sur un ensemble vide ou de mesure nulle.

Appelons mesure d'un ensemble S d'états F l'intégrale $\int_S d\tau_F$. Alors si S est l'ensemble des états F de V où $p(F) \geq A > 0$, on aura $1 \geq \int_V p(F) d\tau_F \geq A \int_S d\tau_F$, d'où $\int_S d\tau_F \leq \frac{1}{A}$. L'ensemble ω (qui est peut-être vide) où $p(F)$ est infini, est compris dans S . Sa mesure est donc inférieure à $\frac{1}{A}$, pour tout nombre positif A . Elle est donc nulle.

Définition du cas quasi-régulier. Nous avons dit qu'on appellerait cas régulier le cas où les deux intégrales $\int_V p(F) d\tau_F$ et $\int_V P(F) d\tau_F$ existent et sont égales à l'unité. Ce cas est compris dans le cas plus général que nous examinerons d'abord, sous le nom de cas quasi-régulier, où l'on suppose seulement que ces intégrales existent et ont la même valeur finie. D'ailleurs on a vu que $p(F)$ est toujours sommable sur V et que $\int_V p(F) d\tau_F \leq 1$. Si donc on n'est pas dans le cas régulier la valeur commune de ces deux intégrales sera < 1 .

On a dans le cas quasi-régulier

$$\int_V [P(F) - p(F)] d\tau_F = 0 \quad \text{avec} \quad P(F) - p(F) \geq 0.$$

Or si une fonction $\theta(F)$ positive ou nulle sur V a une intégrale nulle sur V alors — qu'elle soit continue ou non, bornée ou non sur V — elle y est nulle « presque partout », c'est-à-dire que l'on a $\int_{\nu} d\tau_F = 0$, en désignant par ν l'ensemble des états F de V où $\theta(F) \neq 0$.

Car, soit v_A l'ensemble des états F où $\theta(F) > A > 0$. On a :

$$0 = \int_V \theta(F) d\tau_F \geq A \int_{v_A} d\tau_F \geq 0 \quad \text{d'où} \quad \int_{v_A} d\tau_F = 0$$

la « mesure » $\int_{v_A} d\tau_F$ de v_A étant nulle quel que soit A , l'ensemble v qui est la réunion de $v_1, v^{1/2}, \dots, v^{1/n}, \dots$ aura aussi une mesure nulle.

Donc $P(F) - p(F)$ est nul presque partout sur V , c'est-à-dire, sauf peut-être sur un ensemble w_1 de mesure nulle. Comme $p(F)$ ne peut être infini que sur un ensemble w_2 de mesure nulle, on voit que $P(F)$ et $p(F)$ sont finis et égaux sur V sauf peut-être sur un ensemble de mesure nulle w formé de la réunion de w_1 et w_2 . La relation

$$0 \leq |P^{(n)}(E, F) - p(F)| \leq P^{(n)}(F) - p^{(n)}(F)$$

montre alors que $P^{(n)}(E, F)$ converge vers une limite $p(F)$ indépendante de E — et cela uniformément pour F fixe quand E varie sur V — sauf peut-être quand F est sur w .

Ceci va nous permettre de donner une définition du cas quasi-régulier, moins condensée mais plus intuitive que celle qui consiste dans l'égalité de deux intégrales d'ailleurs non directement données.

Nous appellerons cas *quasi-régulier* le cas où la densité de probabilité $P^{(n)}(E, F)$ converge — sauf peut-être quand F appartient à un certain ensemble w de mesure nulle — vers une limite $\pi(F)$ finie et indépendante de E , la convergence ayant lieu uniformément pour F fixe quand E varie sur V (mais en cessant d'exiger l'égalité $\int_V \pi(F) d\tau_F = 1$).

Dans ce cas, pour F fixe, en dehors de w , et pour ε positif donné, il existe un entier N tel que

$$\pi(F) - \varepsilon < P^{(n)}(E, F) < \pi(F) + \varepsilon$$

pour $n > N$. D'où $\pi(F) - \varepsilon \leq p^{(n)}(F) \leq P^{(n)}(F) \leq \pi(F) + \varepsilon$ pour $n > N$. Autrement dit, en dehors de w , $p^{(n)}(F)$ et $P^{(n)}(F)$ — qui tendent toujours vers $p(F)$ et $P(F)$ — tendent vers la limite finie $\pi(F)$. Donc en dehors de w , $p(F)$ et $P(F)$ sont finis et égaux. D'ailleurs, dans tous les cas, l'intégrale $\int_V p(F) d\tau_F$ existe et est ≤ 1 . Donc $\int_V P(F) d\tau_F$ existe aussi et on a

$$\int_V P(F) d\tau_F = \int_V p(F) d\tau_F \leq 1.$$

(Si l'on n'est pas dans le cas régulier la valeur commune de ces intégrales sera < 1). La réciproque a été établie plus haut et complète l'identité des deux définitions du cas quasi-régulier.

On peut même obtenir une sorte de convergence uniforme relativement non seulement à E mais encore à F . Pour cela, considérons une partie *bornée* arbitraire mais fixe W de V et, pour ε positif arbitraire, désignons par S_n l'ensemble des F de W où $P^{(n)}(F) - p^{(n)}(F) > \varepsilon$. L'ensemble S commun aux S_n étant évidemment compris dans w est de mesure nulle. D'autre part S_{n+1} appartient à S_n . On peut donc écrire

$$S_1 = S + (S_1 - S_2) + (S_2 - S_3) + (S_3 - S_4) \dots\dots\dots$$

d'où, en posant $\int_G d\tau_F = \text{mes. } G$.

$$\text{mes. } S_1 = [\text{mes. } S_1 - \text{mes. } S_2] + [\text{mes. } S_2 - \text{mes. } S_3] + \dots\dots\dots$$

Le second membre est donc une série convergente. Pour η positif donné, il existera un rang q tel que le reste de cette série de rang q soit $< \eta$. Donc $\text{mes. } S_q < \eta$. Jusqu'ici ε et η étaient arbitraires. Donnons à ε et η une même valeur $\frac{\omega}{2^n}$ en prenant successivement pour n les valeurs 1, 2, 3 ... ; alors q prendra une suite de valeurs $q_1, q_2 \dots$ qu'on peut supposer croissantes et S_q deviendra successivement un des termes $\sigma_1, \sigma_2, \dots$ d'une certaine suite d'ensembles. La mesure de σ_n sera inférieure à $\frac{\omega}{2^n}$ et celle de l'ensemble $T_n = \sigma_n + \sigma_{n+1} + \dots\dots$ sera inférieure à $\frac{\omega}{2^{n-1}}$. Comme σ_n est l'ensemble des F de W où $P^{(q_n)}(F) - p^{(q_n)}(F) > \varepsilon_n$, on voit que sur W , on aura en dehors de T_n

$$P^{(r)}(F) - p^{(r)}(F) \leq \varepsilon_{n+s}$$

pour $r \geq q_{n+s}$ et $s = 1, 2 \dots$. Autrement dit $[P^{(r)}(F) - p^{(r)}(F)]$ converge uniformément vers zéro sur W en dehors d'un ensemble T_n de mesure aussi petite que l'on veut.

En résumé, dans le cas quasi-régulier, la densité de probabilité $P^{(n)}(E, F)$ converge, quand n croît, vers une limite $P(F)$ indépendante de E et cela uniformément quand d'une part E varie arbitrairement sur V , quand d'autre part et simultanément F varie arbitrairement sur une partie bornée quelconque W de V sauf peut-être sur une partie T de W dont on peut supposer la mesure aussi petite que l'on veut.

(Il doit être rappelé que, si, pour tout nombre positif θ , on peut supposer T de mesure inférieure à θ , il n'en résulte pas qu'on puisse supposer T de mesure nulle. Car, T varie avec θ ; or, en appelant t la partie commune à une infinité d'ensembles T il y aura bien dans W convergence en dehors de t mais non convergence nécessairement uniforme).

Les raisonnements ci-dessus et leur conclusion subsisteraient si on prenait pour W une partie de V qui soit illimitée pourvu qu'on suppose sa mesure finie.

Cas régulier. — Cherchons la condition pour que le cas quasi-régulier soit en même temps le cas régulier.

On a toujours³⁾

$$\lim_{n \rightarrow \infty} \int_V p^{(n)}(F) dF = \int_V p(F) dF \leq 1 = \int_V P^{(n)}(E, F) dF.$$

Et, dans le cas quasi-régulier

$$(2) \quad \int_V p(F) dF = \int_V P(F) dF.$$

Pour qu'on soit dans le cas régulier quand la condition (2) est remplie, il faut et il suffit

$$\lim_{n \rightarrow \infty} \int_V [P^{(n)}(E, F) - p^{(n)}(F)] dF = 0.$$

Soit alors ε un nombre positif arbitraire il y aura un entier q tel que, pour $n > q$ on ait

$$\int_V [P^{(n)}(E, F) - p^{(n)}(F)] dF < \frac{\varepsilon}{2}.$$

Pour toute partie v de V on aura donc, si $n > q$

$$\int_v P^{(n)}(E, F) dF < \frac{\varepsilon}{2} + \int_v p^{(n)}(F) dF < \frac{\varepsilon}{2} + \int_v p(F) dF$$

et si v est une partie de V telle que $\int_v p(F) dF < \frac{\varepsilon}{2}$, on aura

$$(3) \quad \int_v P^{(n)}(E, F) dF < \varepsilon.$$

³⁾ Pour abrégier, nous remplacerons dans la suite la notation $d\tau_F$ par la notation plus simple dF .

La fonction $p(F)$ étant, en tout cas, sommable sur V , y est, comme on sait, (III, p. 109), « absolument continue » c'est-à-dire qu'à chaque nombre positif ω correspond un nombre η_0 tel que l'on ait $\int_v p(F) dF < \omega$ pour

toute partie v de V de mesure $< \eta_0$. En prenant $\omega = \frac{\varepsilon}{2}$, on voit qu'on aura, pour $n > q$, $\int_v P^{(n)}(E, F) dF < \varepsilon$. D'ailleurs, chacune des fonctions $P^{(1)} \dots P^{(q)}$, étant sommable sur V , y sera absolument continue, c'est-à-dire qu'il existera des nombres $\eta_1 \dots \eta_q$ tels que, pour mes. $v < \eta_k$, on ait $\int_v P^{(k)}(E, F) dF < \varepsilon$. Finalement, si η est le plus petit des nombres

$\eta_0, \eta_1, \dots, \eta_q$, on voit que, pour toute partie v de V de mesure $< \eta$, on aura $\int_v P^{(n)}(E, F) dF < \varepsilon$, quel que soit n , ε, η, v étant indépendants de n . C'est ce que nous exprimerons en disant que l'absolue continuité des fonctions de F , $P^{(n)}(E, F)$ doit être, pour E fixe, *uniforme* quand n varie. Considérons maintenant deux cas. Supposons d'abord V borné ou plus généralement de *mesure finie*. Alors, cette condition nécessaire d'uniformité est suffisante. En effet, soit ε arbitraire > 0 , il y a, par hypothèse, pour E fixe, un nombre $\eta > 0$ tel que pour toute partie v de V de mesure $< \eta$, on ait $\int_v P^{(n)}(E, F) dF < \frac{\varepsilon}{2}$. D'autre part, d'après la p. 193, puisqu'on est déjà dans le cas quasi-régulier, il y a une partie T de V de mesure inférieure à η et telle que $P^{(n)}(E, F)$ converge uniformément vers $p(F)$ quand F varie hors de T . Dès lors, on a

$$\int_T P^{(n)}(E, F) dF < \frac{\varepsilon}{2} \text{ et il existe un nombre } q \text{ tel que}$$

$$\int_{V-T} |P^{(n)}(E, F) - p(F)| dF < \frac{\varepsilon}{2} \text{ pour } n > q. \text{ D'où}$$

$$(4) \quad \begin{cases} \int_v p(F) dF \geq \int_{V-T} p(F) dF > \int_{V-T} P^{(n)}(E, F) dF - \frac{\varepsilon}{2} \\ = 1 - \int_T P^{(n)}(E, F) dF - \frac{\varepsilon}{2} > 1 - \varepsilon. \end{cases}$$

Pour tout $\varepsilon > 0$ on a donc

$$1 - \varepsilon < \int_V p(F) dF = \int_V P(F) dF \leq 1$$

on est bien dans le cas régulier.

On observera que si la condition d'uniformité de l'absolue continuité de $\int_V P^{(n)}(E, F) dF$ a été prouvée nécessaire pour chaque état E fixe de V , la démonstration qu'elle est suffisante suppose simplement qu'elle ait lieu pour au moins un état E de V .

Si V n'est pas de mesure finie, cette uniformité n'est plus suffisante comme le montrerait l'exemple de la p. 189. On doit compléter par une autre, cette condition d'uniformité de la continuité absolue des intégrales $\int_V P^{(n)}(E, F) dF$.

Pour cela, rappelons que, par définition de l'intégrale d'une fonction $\Phi(F)$ sur un ensemble non borné V

$$\int_V \Phi(F) dF = \lim_{s \rightarrow \infty} \int_{W_s} \Phi(F) dF$$

où $W_1, W_2, \dots, W_s, \dots$ est une suite de parties bornées de V telles que W_s appartienne à W_{s+1} et que V soit identique à la réunion des W_s . Posons $v_s = V - W_s$. Dès lors, pour chaque entier n , il existe un entier s_n , tel que $\int_{v_s} P^{(n)}(E, F) dF < \varepsilon$ pour $s \geq s_n$ et un entier s' tel

que $\int_{v_s} p(F) dF < \frac{\varepsilon}{2}$ pour $s \geq s'$. Or, si l'on est dans le cas régulier, on aura en vertu de l'inégalité (3)

$$\int_{v_s} P^{(n)}(E, F) dF < \varepsilon$$

pour $s \geq s'$ et $n > q$, q étant indépendant de s . Appelons alors s'' le plus grand des nombres $s_1, \dots, s_q, s'$, on aura

$$\int_{v_{s''}} P^{(n)}(E, F) dF < \varepsilon$$

pour toute valeur entière de n , s'' étant indépendant de n . D'où

$$\int_V P^{(n)}(E, F) dF - \int_{W_s} P^{(n)}(E, F) dF < \varepsilon$$

pour $s > s''$, quel que soit l'entier arbitraire n . C'est ce que nous exprimerons en disant que, dans le cas régulier, la convergence, exprimée par la formule

$$\lim_{s \rightarrow \infty} \left[\int_V P^{(n)}(E, F) dF - \int_{W_s} P^{(n)}(E, F) dF \right] = 0$$

des intégrales à limites infinies $\int_V P^{(n)}(E, F) dF$, est uniforme quand n varie.

Réciproquement, supposons — toujours dans le cas quasi-régulier — que non seulement l'absolue continuité mais la convergence de chaque intégrale à limites infinies $\int_V P^{(n)}(E, F) dF$ soient uniformes quand n varie, et ceci au moins pour un état E fixé. Alors, pour tout $\varepsilon > 0$ il existera une partie bornée W de V telle qu'en posant $v = V - W$, on ait à la fois

$$\int_v p(F) dF < \varepsilon, \quad \int_v P^{(n)}(E, F) < \varepsilon$$

et ceci quel que soit n . Or, en vertu de l'uniformité de l'absolue continuité de $\int_V P^{(n)}(E, F) dF$ et, par suite, de celle de $\int_W P^{(n)}(E, F) dF$, le raisonnement fait plus haut dans le cas où V est borné conduira à remplacer dans le cas actuel les inégalités (4) par

$$\int_W p(F) dF > \int_W P^{(n)}(E, F) dF - \varepsilon = 1 - \int_v P^{(n)}(E, F) dF - \varepsilon > 1 - 2\varepsilon.$$

On a donc pour tout $\varepsilon > 0$

$$1 \geq \int_V p(F) dF > 1 - 2\varepsilon$$

et on est bien encore dans le cas régulier.

Une condition suffisante plus simple. Il y a un cas particulier où on peut obtenir une condition suffisante plus simple en utilisant un théorème (III, p. 120) d'après lequel, si les fonctions d'une suite non croissante de

fonctions $\Phi^{(n)}(F) \geq 0$ sont sommables sur V^* , il en est de même de leur limite $\Phi(F)$ et on a

$$(5) \quad \int_V \Phi(F) dF = \lim_{n \rightarrow \infty} \int_V \Phi^{(n)}(F) dF.$$

Les fonctions $P^{(n)}(F)$ majorent les fonctions $P^{(n)}(E, F)$, sommables quand F varie sur V . Il n'en résulte pas que les $P^{(n)}(F)$ soient sommables. L'intégrale $\int_V P^{(n)}(F) dF$ peut être infinie, soit parce que $P^{(n)}(F)$ est infinie sur un ensemble de mesure non nulle, soit même quand $P^{(n)}(F)$ est finie et même borné, si V est non borné, comme le montre l'exemple de la p. 189.

Par contre, si l'une des fonctions $P^{(n)}(F)$ est sommable sur V , il en sera de même de $P^{(n+1)}(F)$, $P^{(n+2)}(F)$... et l'on aura une relation analogue à (5). D'ailleurs, on aura aussi dans ce cas

$$\int_V P^{(n)}(F) dF \geq \int_V P^{(n)}(E, F) dF = 1$$

d'où $\int_V P(F) dF \geq 1$.

En résumé, qu'on soit ou non dans le cas quasi-régulier, s'il existe une fonction $\varphi(F)$ sommable sur V , qui majore — quels que soient E et F sur V — l'une au moins des fonctions $P^{(n)}(E, F)$ — et, par suite, si l'une au moins des $P^{(n)}(F)$ est sommable sur V — alors $P(F)$ est sommable sur V et on a

$$\int_V P(F) dF \geq 1.$$

Quand une telle fonction $\varphi(F)$ n'existe pas, deux circonstances aussi opposées que possible peuvent se présenter: l'intégrale $\int_V P(F) dF$ peut être infinie, ou au contraire < 1 et même nulle comme dans l'exemple de la page 189.

Quand la fonction $\varphi(F)$ existe, on a à la fois

$$\int_V \varphi(F) dF \leq 1 \leq \int_V P(F) dF.$$

⁴) Nous entendons ici le mot sommable pour intégrable au sens de M. Lebesgue (III p. 107). D'ailleurs, dans tout ce mémoire, nous supposons mesurables au sens de M. Lebesgue toutes les fonctions employées.

Si donc on se trouve dans le cas quasi-régulier ces deux intégrales sont égales et par suite sont égales à l'unité. Nous obtenons ainsi la condition suffisante plus simple annoncée :

Pour que le cas quasi-régulier se trouve être en même temps le cas régulier, il suffit qu'il existe une fonction $\varphi(F)$ sommable sur V qui majore quels que soient E et F sur V , mais pour un même rang n , l'une au moins des densités itérées $P^{(n)}(E, F)$.

Un cas simple où l'existence de la fonction φ est assurée est celui où V est borné (ou de mesure finie) et où, en outre $p(E, F)$, — ou l'une au moins des $P^{(n)}(E, F)$ — a une borne supérieure finie quand E, F varient indépendamment sur V . Plus particulièrement encore, c'est ce qui aura lieu quand, V étant borné, $p(E, F)$ — ou l'une au moins des $P^{(n)}(E, F)$ — est continue quand E, F varient sur V .

Une autre simplification à déduire de l'existence de la fonction φ concerne le mode de convergence de $P^{(n)}(E, F)$ vers $P(F)$. Nous avons vu que dans le cas quasi-régulier cette convergence qui a lieu presque partout est uniforme sur chaque partie bornée W de V , quand on excepte de W un ensemble peut-être vide mais qu'on peut supposer de mesure aussi petite que l'on veut. *Lorsqu'il existe une fonction $\varphi(F)$ sommable sur V et majorant l'une des fonctions $P^{(n)}(E, F)$, le même résultat est valable en prenant pour W toute la région V , que cette dernière soit bornée ou non.*

En effet, dans ce cas, pour n assez grand ($n > q$), $P^{(n)}(F)$ est sommable sur V et l'on a

$$0 = \lim_{n \rightarrow \infty} \int_V [P^{(n)}(F) - p^{(n)}(F)] dF.$$

Si $A_\varepsilon^{(n)}$ est l'ensemble des F de V où $P^{(n)}(F) - p^{(n)}(F) > \varepsilon$ on a

$$\int_V [P^{(n)}(F) - p^{(n)}(F)] dF > \varepsilon \text{ mes. } A_\varepsilon^{(n)} \geq 0$$

donc mes. $A_\varepsilon^{(n)}$ tend vers 0 avec $\frac{1}{n}$. Soit alors ω un nombre positif arbitraire; pour $\varepsilon = \frac{1}{q}$ il y a un nombre $n_q > q$ tel que mes. $A_\varepsilon^{(n)} < \frac{\omega}{2^q}$ pour $n \geq n_q$. Soit enfin A la réunion des ensembles $A_\varepsilon^{(n)}$ pour $\varepsilon = \frac{1}{q}$,

$n = n_q$, $q = 1, 2, \dots$. Sa mesure est $< \sum_q \frac{\omega}{2^q} = \omega$ et l'on aura, en dehors de A

$$|P^{(n)}(E, F) - p(F)| \leq P^{(n)}(F) - p^{(n)}(F) \leq \frac{1}{q}$$

pour $n = n_q$ et par suite pour $n \geq n_q$. Dès lors $P^{(n)}(E, F)$ converge vers $p(F)$ uniformément quand E est arbitraire sur V et quand F varie sur la région obtenue en enlevant de V un ensemble A , peut-être vide, mais en tout cas de mesure inférieure à un nombre positif ω choisi arbitrairement.

Conditions nécessaires pour le cas quasi-régulier. Cherchons à déterminer des conditions nécessaires pour que $P(F)$ et $p(F)$ soient finis et égaux presque partout sur V .

I. *Condition (A).* Deux cas sont à distinguer: 1° $P(F)$ est nul presque partout sur V ; alors il en est de même pour $p(F)$ et aucune condition nécessaire n'est à envisager. 2° $P(F)$ n'est pas presque partout nul. En désignant par v_ε l'ensemble sur lequel $P(F) > \varepsilon$, nous sommes dans le cas où v_0 est de mesure non nulle. Or v_0 étant formé de la réunion des ensembles v_ε où $\varepsilon > 0$ serait de mesure nulle s'il en était ainsi de tous ces v_ε . Dès lors: il existe au moins une valeur positive ε telle que v_ε soit de mesure positive.

Or, on a vu (p. 193) que dans le cas quasi-régulier pour toute partie W bornée de V et tout nombre $\omega > 0$, $P(F) - p^{(n)}(F)$ converge uniformément vers zéro sur W sauf, peut-être, sur un ensemble T de mesure $< \omega$. En particulier, il existe un nombre ν tel que pour $n \geq \nu$

$$p^{(n)}(F) > P(F) - \frac{\varepsilon}{2}$$

sur $W - T$. Prenons, en particulier, pour W l'ensemble v_ε ou une partie de v_ε qui soit bornée et de mesure positive et pour ω la moitié de la mesure positive de v_ε ou de cette partie: alors l'ensemble $w = W - T$ sera aussi de mesure positive et l'on aura sur w

$$p^{(n)}(F) > \frac{\varepsilon}{2} \quad \text{pour } n > \nu.$$

Autrement dit, il existe dans ce cas une partie w de V qui est bornée et de mesure positive et telle que, à partir d'un certain rang ν , la fonction $P^{(n)}(E, F)$ reste, quels que soient E sur V , F sur w , supérieure à un nombre positif indépendant de E , de F et de n .

En particulier, il existe au moins un rang ν tel que $p^{(\nu)}(F)$ ne soit pas nul presque partout. On dira, dans ce cas que la condition (A) est réalisée.

Cette condition sera plus commode à employer parmi les conditions suffisantes, que la précédente. Mais elle lui est équivalente. Car, si $p^{(\nu)}(F) > 0$ sur un ensemble de mesure positive, alors il y a, au moins un nombre $\eta > 0$ tel que la mesure de l'ensemble où $p^{(\nu)}(F) > \eta$ soit positive. Dès lors, sur ce même ensemble on aura, pour $n \geq \nu$, $p^{(n)}(F) > \eta$ d'où $P^{(n)}(E, F) > \eta$.

II. *Condition nécessaire* (B). Une autre condition nécessaire pour le cas quasi-régulier est évidemment que $P(F)$ ne puisse être infini que sur un ensemble de mesure nulle, puisqu'il en est ainsi pour la fonction sommable $p(F)$, et que $p(F)$ et $P(F)$ sont finis et égaux presque partout.

Simplifications. Si l'une au moins des fonctions $P^{(n)}(E, F)$, soit $P^{(\nu)}(E, F)$, est uniformément continue quand E et F varient sur V , la condition (A) se simplifie.

Si elle a lieu, il existe au moins un entier ν et un état L tel que $p^{(\nu)}(L) > 0$. Réciproquement, s'il en est ainsi, on peut d'abord supposer $\nu > r$ puisque $p^{(n)}(L)$ est une fonction non décroissante de n . Alors, comme on a vu (p. 183) que $p^{(\nu)}(F)$ est continue, il existe au moins un nombre $\eta > 0$, tel que $p^{(\nu)}(F) > \frac{1}{2} p^{(\nu)}(L) > 0$ pour $LF < \eta$. On a supposé (p. 180) le domaine V , tel qu'il existe au moins une sphère v composée d'états de V tous à distance de L inférieure à η . Dès lors $p^{(\nu)}(F) > 0$ sur un ensemble v de mesure positive: (A) est réalisée.

D'autre part, si l'on suppose, en outre, que $P^{(\nu)}(E, F)$ est borné (ce qui aura lieu nécessairement si V est fini) alors $P^{(\nu)}(F)$ et par suite $P(F)$ sont bornés: (B) sera vérifiée d'elle-même.

On observera que ces deux simplifications subsistent si l'on suppose simplement que $P^{(\nu)}(E, F)$ représente pour les situations diverses de E , une famille de fonctions de F « également continues » sur V .

Conditions suffisantes pour le cas quasi-régulier. Considérons un état F tel que $P(F)$ soit fini et trois états arbitraires E, E_1, G . Pour n assez grand ($n - 1 \geq m$), $P^{(n-1)}(F)$ sera aussi fini et par suite aussi $P^{(n-1)}(G, F)$, $P^{(n)}(F)$, $P^{(n)}(E, F)$, $P^{(n)}(E_1, F)$. Etendons avec M. Hostinsky

(I, p. 44, I^{bis}, p. 21) au cas actuel la méthode de Markoff. Formons la différence

$$(6) \quad P^{(n)}(E, F) - P^{(n)}(E_1, F) = \int_V [\rho(E, G) - \rho(E_1, G)] P^{(n-1)}(G, F) dG.$$

Soient V' , V'' les deux parties de V sur lesquelles le crochet est respectivement ≥ 0 , < 0 . En posant

$$\begin{aligned} \theta &= \int_{V'} [\rho(E, G) - \rho(E_1, G)] dG = \int_{V'} K dG \\ \theta' &= \int_{V''} [\rho(E_1, G) - \rho(E, G)] dG = \int_{V''} K dG \end{aligned}$$

où K est la valeur absolue du crochet, on aura $\theta \geq 0$ et $\theta - \theta' = \int_V \rho(E, G) dG - \int_V \rho(E_1, G) dG = 1 - 1 = 0$, d'où

$$\theta' = \theta \geq 0.$$

Supposons d'abord $\theta > 0$. Alors on pourra écrire

$$\begin{aligned} P^{(n)}(E, F) - P^{(n)}(E_1, F) &= \\ \theta \left\{ \frac{\int_{V'} K P^{(n-1)}(G, F) dG}{\int_{V'} K dG} - \frac{\int_{V''} K P^{(n-1)}(G, F) dG}{\int_{V''} K dG} \right\} \\ &\leq \theta [P^{(n-1)}(F) - \rho^{(n-1)}(F)] \end{aligned}$$

L'inégalité

$$(7) \quad P^{(n)}(E, F) - P^{(n)}(E_1, F) \leq \theta [P^{(n-1)}(F) - \rho^{(n-1)}(F)]$$

qui vient d'être établie pour $\theta \neq 0$, reste d'ailleurs exacte pour $\theta = 0$, car, dans ce cas, K serait nul identiquement ou presque partout nul sur V et alors en vertu de l'égalité (6) le premier membre de (7) serait aussi nul.

Dans cette inégalité, θ est compris entre 0 et 1, car

$$0 \leq \theta \leq \int_{V'} \rho(E, G) dG \leq \int_V \rho(E, G) dG = 1.$$

Seulement, d'après sa définition, θ dépend de E et de E_1 . Mais on a

$$\begin{aligned}\theta &= \int_{V'} p(E, G) dG - \int_{V'} p(E_1, G) dG \\ &= \int_V p(E, G) dG - \int_{V''} p(E, G) dG - \int_{V'} p(E_1, G) dG.\end{aligned}$$

Or, sur V' comme sur V'' , $p(E, G) \geq p^{(1)}(G)$; $p(E_1, G) \geq p^{(1)}(G)$.

Donc
$$\theta \leq 1 - \int_{V'} p^{(1)}(G) dG = \theta_1.$$

Le nombre θ_1 est évidemment ≤ 0 , ≤ 1 et ne dépend ni de E , ni de E_1 . Comme on a

$$P^{(n)}(E, F) - P^{(n)}(E_1, F) \leq \theta_1 [P^{(n-1)}(F) - p^{(n-1)}(F)]$$

on voit qu'on aura pour $n > m$

$$(8) \quad P^{(n)}(F) - p^{(n)}(F) \leq \theta_1 [P^{(n-1)}(F) - p^{(n-1)}(F)]$$

où θ , est indépendant de n (pour $n > m$) et de F , avec $0 \leq \theta_1 \leq 1$. Seulement si $\theta_1 = 1$, on ne retrouverait qu'une inégalité déjà obtenue. Le cas intéressant est donc celui où $\theta_1 < 1$, c'est-à-dire où $p^{(1)}(G)$ n'est pas presque partout nul sur V . C'est celui où la condition (A) se trouve vérifiée quand on suppose, comme nous allons le faire d'abord, que dans son énoncé de la p. 201, l'entier ν est égal à l'unité. Dans ce cas, $\theta_1 < 1$ et, en vertu de (8), on aura

$$P^{(n)}(F) - p^{(n)}(F) \leq \theta_1^{n-m} [P^{(m)}(F) - p^{(m)}(F)].$$

Revenons maintenant à la condition (A), mais sous sa forme générale: supposons qu'il existe un rang ν pour lequel $p^{(\nu)}(G)$ ne soit pas presque partout nul sur V . On pourra recommencer le raisonnement précédent, mais en partant de l'égalité

$$\begin{aligned}P^{(m+s\nu)}(E, F) - P^{(m+s\nu)}(E_1, F) = \\ \int_V [P^{(n)}(E, G) - P^{(n)}(E_1, G)] P^{(m+(s-1)\nu)}(G, F) dG.\end{aligned}$$

On trouvera alors

$$P^{(m+sv)}(F) - p^{(m+sv)}(F) \leq (\theta_v)^s [P^{(m)}(F) - p^{(m)}(F)]$$

avec

$$0 \leq \theta_v = 1 - \int_V p^{(v)}(G) dG < 1.$$

Pour tout entier $n \geq m$, il y a un entier $s \geq 0$ tel que

$$m + sv \leq n < m + (s+1)v$$

d'où

$$P^{(n)}(F) - p^{(n)}(F) \leq P^{(m+sv)}(F) - p^{(m+sv)}(F)$$

et

$$(\theta_v)^s < \left[(\theta_v)^{\frac{1}{v}} \right]^{n-m-v}.$$

En posant $q_v = (\theta_v)^{\frac{1}{v}}$, on aura

$$(9) \quad |P^{(n)}(E, F) - p(F)| \leq P^{(n)}(F) - p^{(n)}(F) \leq (q_v)^n B(F)$$

avec

$$B(F) = \frac{P^{(m)}(F) - p^{(m)}(F)}{(q_v)^{m+v}}.$$

Comme q_v est indépendant de F et de n et inférieur à l'unité, on voit que $P^{(n)}(F) - p^{(n)}(F)$ tend vers zéro. Finalement, quand la condition (A) est réalisée, on a $P(F) = p(F)$ pour tout état F où $P(F)$ est fini. D'ailleurs $p(F) \geq p^{(v)}(F)$, donc, $P(F)$ ne sera pas, dans ce cas, presque partout nul.

Faisons maintenant entrer en ligne de compte la seconde condition nécessaire, la condition (B) de la p. 201; nous aurons alors un ensemble à la fois nécessaire et suffisant:

Pour que — à l'exception tout au plus d'un ensemble u de mesure nulle d'états F de V — $P^{(n)}(E, F)$ converge — uniformément pour F fixe quand E varie sur G — vers une limite finie indépendante de l'état initial E et non presque partout nulle sur V , il faut et il suffit: I qu'il existe au moins un rang v tel que $p^{(v)}(F)$ ne soit pas presque partout nul sur V , II que $P(F)$ soit fini presque partout sur V .

Observons aussi que si la condition (A) n'est pas réalisée, tous les $p^{(m)}(G)$ seront nuls presque partout. La fonction $p(G)$ sera donc nulle sauf, peut-être, sur un ensemble dénombrable α d'ensembles de mesure nulle; α étant alors de mesure nulle, $p(G)$ sera nul presque partout sur V . Nous pouvons donc dire. Pour qu'on soit dans le cas quasi-régulier :

Ou bien la condition (A) est réalisée et alors il faut et il suffit que $P(F)$ soit fini presque partout sur V .

Ou bien (A) n'étant pas réalisée et, par suite, $p(G)$ étant nul presque partout sur V , il faut et il suffit que $P(F)$ soit nul presque partout sur V . Dans ce dernier cas $\int_V P(F) dF = 0$. On sera certainement dans le cas quasi-régulier proprement dit, c'est-à-dire sans être dans le cas régulier⁵⁾.

Nature de la convergence. Nous avons déjà vu, p. 193 que, dans le cas quasi-régulier, la convergence est uniforme sur toute partie W bornée de V , après avoir retranché de W un ensemble convenable dont on peut supposer la mesure aussi petite que l'on veut.

L'inégalité (9) permet de préciser un peu plus la nature de la convergence. Elle montre d'abord qu'en tout état F où $P(F)$ est fini, c'est-à-dire presque partout sur V , la convergence de $P^{(n)}(E, F)$ vers $P(F)$ est, à partir d'un certain rang, au moins aussi rapide que celle des termes d'une *progression géométrique* dont les termes dépendent de F , mais dont la raison q_v ($0 < q_v < 1$) est indépendante de E et de F .

En limitant encore le champ de convergence, on peut même s'arranger pour que les termes de la progression soient eux-mêmes indépendants de F comme de E (ce qui, en même temps entraînera l'uniformité de la convergence sur ce nouveau champ).

C'est d'abord ce qui aura lieu sur chaque ensemble $V_r^{(m)}$ en appelant ainsi la partie de V où l'on a $P^{(m)}(F) < r$. Car on a sur $V_r^{(m)}$ $|P^{(n)}(E, F) - P(F)| \leq (q_v)^n \frac{r}{(q_v)^m \theta_v}$ pour $n \geq m$. C'est, par suite, aussi ce qui aura lieu sur l'ensemble V_s formé des ensembles $V_r^{(m)}$ en nombre fini, pour lesquels $r + m = s$. Car en appelant λ_s le plus grand des nombres $\frac{r}{(q_v)^m \theta_v}$ où $r + m = s$ on aura sur V_s , pour $n > s$

$$|P^{(n)}(E, F) - P(F)| < \lambda_s (q_v)^n.$$

⁵⁾ Il serait d'ailleurs intéressant de chercher si l'on peut effectivement former l'exemple d'un cas quasi-régulier où $0 < \int_V P(F) dF < 1$.

Or, si w est encore l'ensemble — peut-être vide, mais de mesure nulle — où $P(F)$ et $p(F)$ ne sont pas finis et égaux, il est clair que tout état de $V - w$ appartient à l'un des V_s et que v_s appartient à V_{s+1} .

Ainsi, on a formé une suite non décroissante d'ensembles V_s , remplissant V , à un ensemble de mesure nulle près et sur chacun, V_s , desquels, la suite $|P^{(n)}(E, F) - P(F)|$ est majorée à partir du rang s par une progression géométrique indépendante non seulement de E mais de F . Cette progression change avec V_s , mais sa raison reste la même.

Dans le cas où V est borné, l'égalité symbolique

$$V = w + (V_1 - V_2) + (V_2 - V_3) + \dots$$

peut aussi s'entendre en mesure, de sorte que $V - V_s$ est de mesure aussi petite que l'on veut avec $\frac{1}{s}$. Si V n'est pas borné, alors pour toute partie bornée W de V , c'est la partie de $V - V_s$ qui appartient à W qui est de mesure aussi petite que l'on veut, pour s assez grand.

Précisions. On peut même préciser dans des cas particulièrement importants dans les applications. Supposons d'abord qu'il existe un rang m tel que $P^{(m)}(F)$ soit borné sur V . La théorie des équations intégrales (II, p. 344, 357, 362) fournit des exemples de noyaux $p(E, F)$ discontinus, non bornés, mais tels que, pour n assez grand, $n = m$, $P^{(n)}(E, F)$ ait une borne supérieure finie μ quand E, F varient indépendamment sur V . Dans ce cas $P^{(n)}(F)$ sera borné sur V et même on aura $P^{(n)}(F) \leq \mu$ pour $n \geq m$. Alors $P(F)$ sera nécessairement borné aussi et la condition (B) sera remplie d'elle-même. De plus, si μ est la borne de $P^{(m)}(F)$ sur V , il résultera de la première condition, qu'on a d'après la formule (9)

$$0 \leq |P^{(n)}(E, F) - P(F)| \leq (q_v)^n \frac{\mu}{(q_v)^m \theta_v}.$$

Dès lors: si l'une au moins des probabilités itérées $P^{(n)}(E, F)$ a une borne supérieure finie quand E, F varient indépendamment sur V , la condition nécessaire et suffisante pour que $P^{(n)}(E, F)$ converge — quel que soit F sur V , et cela uniformément pour F fixe quand E varie sur V — vers une limite $P(F)$ indépendante de l'état initial E et qui ne soit pas presque partout nulle sur V est qu'il existe au moins un rang v pour lequel $p^{(v)}(F)$ ne soit pas presque partout nul sur V .

Et dans ce cas, non seulement la convergence est uniforme pour F fixe, mais: 1° elle est uniforme quand E, F varient indépendamment sur V , 2° la convergence de la suite des termes $|P^{(n)}(E, F) - P(F)|$ vers zéro est au moins aussi rapide que celle des termes d'une certaine progression géométrique convergente indépendante de E et de F .

D'autre part, si l'une au moins des probabilités itérées $P^{(n)}(E, F)$ est uniformément continue quand E, F varient sur V , la condition nécessaire et suffisante pour qu'on soit dans le cas quasi-régulier est

- ou bien que $P(F)$ soit nul presque partout sur V
ou bien que 1° $P(F)$ soit fini partout sur V
2° qu'il existe un rang ν et un état L tel que $p^{(\nu)}(L) \neq 0$.

Il faut que $P(F)$ soit fini presque partout sur V mais on a vu p. 184, qu'alors $P(F)$ est nécessairement fini partout sur V . Dans ces conditions $P(F)$ est uniformément continu sur V et $P^{(n)}(E, F)$ converge vers $P(F)$ quels que soient E et F sur V . La convergence est uniforme quand E varie sur V et F sur une partie bornée quelconque de V .

Ces énoncés se simplifient encore quand on les combine:

Si l'une au moins des probabilités itérées $P^{(n)}(E, F)$ est bornée et uniformément continue quand E, F varient sur V , la condition nécessaire et suffisante pour qu'on soit dans le cas quasi-régulier est ou bien que $P(F)$ soit nul presque partout sur V ou bien qu'il existe un rang ν et un état L tel que $p^{(\nu)}(L) \neq 0$.

Et dans ce cas $P^{(n)}(E, F)$ converge uniformément quand E, F varient sur V , vers la limite $P(F)$ (qui est continue). Dans les deux cas que nous allons envisager maintenant, le cas quasi-régulier ne se distingue pas du cas régulier.

Supposons que l'une au moins des $P^{(n)}(F)$ soit sommable sur V . Alors $P^{(n)}(F)$ sera fini presque partout sur V ; il en sera donc de même de $P(F)$ et ici encore la condition (B) sera vérifiée d'elle-même. De plus, nous avons vu p. 198 que dans ce cas $\int_V P(F) dF$ est fini et ≥ 1 .

Ainsi dans ce cas $P(F)$ ne peut être presque partout nul. Alors, pour qu'on soit dans le cas quasi-régulier, il faut que la condition (A) soit réalisée. Mais alors $1 \leq \int_V P(F) dF = \int_V p(F) dF \leq 1$ donc $\int_V P(F) dF = 1$; on sera dans le cas régulier.

En résumé: 1° le cas quasi-régulier ne peut se présenter sans être en même temps régulier que si aucune des fonctions $P^{(n)}(F)$ n'est sommable sur V et

2° lorsque l'une au moins des fonctions $P^{(n)}(E, F)$ est majorée par une fonction de F sommable sur V , la condition nécessaire et suffisante pour qu'on soit dans le cas régulier est que l'une au moins des $p^{(n)}(F)$ ne soit pas presque partout nulle. De plus, dans ce même cas, si cette condition est réalisée, nous avons vu (p. 193) que $P^{(n)}(E, F)$ converge uniformément vers $P(F)$ quand E variant arbitrairement sur V , F varie sur V , sauf, peut-être, sur un certain ensemble indépendant de E et dont on peut rendre la mesure aussi petite qu'on veut.

Lorsque V est borné ou de mesure finie, si l'un au moins des $P^{(n)}(F)$, par exemple $P^{(m)}(F)$, est borné sur V , $P^{(m)}(F)$ sera aussi sommable sur V et on pourra appliquer à la fois les propriétés de deux des cas qui viennent d'être examinés :

Si l'une au moins des fonctions $P^{(n)}(E, F)$ a une borne supérieure finie quand E, F varient sur V et si V est borné ou de mesure finie, la condition nécessaire et suffisante pour qu'on soit dans le cas régulier est que l'une au moins des fonctions $p^{(n)}(G)$ ne soit pas nulle sur V presque partout. De plus, dans ce cas, 1° $P^{(n)}(E, F)$ converge uniformément vers sa limite $P(F)$ lorsque E, F varient arbitrairement sur tout l'ensemble V ; 2° la série $\sum |P^{(n)}(E, F) - P(F)|$ est majorée par une certaine progression géométriqueⁿ convergente indépendante de E et de F .

Enfin, l'énoncé suivant s'obtient comme cas particulier de tous les précédents :

Si $p(E, F)$ ou si plus généralement l'une $P^{(m)}(E, F)$ des probabilités itérées est continue sur un domaine V borné alors la condition nécessaire et suffisante pour qu'on soit dans le cas régulier est que, pour au moins un état L de V et un rang $v \geq m$, $P^{(v)}(E, L)$ ne soit nul pour aucun état E de V .

Dans ce cas, les probabilités itérées $P^{(n)}(E, F)$ sont aussi continues à partir d'un certain rang, elles convergent uniformément quand E, F varient sur V vers une limite $P(F)$ continue sur V et la convergence de $|P^{(n)}(E, F) - P(F)|$ vers zéro est au moins aussi rapide que celle des termes d'une certaine progression géométrique indépendante de E et de F .

Dans ce cas, il y a correspondance absolue entre la condition trouvée ici et celle qu'on obtient dans le cas plus simple des suites discontinues d'états (IV, p. p. 2, 3).

Cas positivement régulier. Par analogie avec le cas d'un nombre fini d'états possibles, il serait, à première vue, naturel, d'appeler ainsi le cas régulier où $p(F)$ serait partout $\neq 0$ sur V . Toutefois $p(F)$ n'est pas

une probabilité, mais une densité de probabilité. Il semblerait donc plus indiqué d'imposer la condition que la probabilité $\int_v p(F) dF$ soit toujours différente de zéro. Comme cette intégrale est, quel que soit $p(F)$, nulle quand v est de mesure nulle, on est amené à une réserve, à supposer seulement que $\int_v p(F) dF > 0$ pour toute partie v de V de mesure positive. Pour cela, il faut et il suffit que $p(F)$ soit presque partout positif. Ceci sera d'ailleurs d'accord avec l'équivalence observée dans les questions précédentes entre les fonctions égales presque partout.

On pourra définir finalement cas *presque positivement* (quasi) *régulier*, tout cas (quasi) régulier où $p(F)$ est positif presque partout sur V . Commençons par le cas presque positivement quasi-régulier, c'est-à-dire celui où $p(F)$ et $P(F)$ sont presque partout à la fois finis, positifs et égaux. L'ensemble (vide ou non) β où $p(F) = 0$ doit être de mesure nulle. Or c'est évidemment l'ensemble commun aux ensembles β_n , β_n étant l'ensemble où $p^{(n)}(F) = 0$. I. Si donc V est de mesure finie la limite de la mesure de β_n tend vers zéro avec $\frac{1}{n}$. II. Sinon pour toute partie W de mesure finie de V , la mesure de la partie commune à W et β_n tend vers 0 avec $\frac{1}{n}$. Réciproquement; si (I) a lieu, β est de mesure nulle; si II a lieu, la partie commune à β et W est de mesure nulle quel que soit le choix de W . Par conséquent, β est encore de mesure nulle. De plus, la condition (A) est alors nécessairement remplie. Nous arrivons donc à la conclusion suivante.

Pour que $P^{(n)}(E, F)$ converge uniformément quand, F étant fixe, E varie sur V , vers une limite finie, indépendante de E et positive, sauf, peut-être, quand F appartient à un certain ensemble de mesure nulle, il faut et il suffit: 1° que pour tout ε positif, il existe un rang v tel que $p^{(v)}(F) > 0$ sauf peut-être sur un ensemble de mesure $< \varepsilon$; ou si V est de mesure infinie que ceci ait lieu pour toute partie de V de mesure finie; 2° que $P(F)$ soit fini presque partout.

Si l'une des fonctions $P^{(n)}(F)$ est sommable sur V , ce qui précède sera aussi la condition nécessaire et suffisante pour qu'on soit dans le cas presque positivement régulier.

On peut aussi donner la condition suffisante suivante:

Pour qu'on soit dans le cas positivement régulier, il suffit que, V étant borné ou de mesure finie, l'une au moins des fonctions $P^{(n)}(E, F)$ ait, lorsque E, F varient indépendamment sur V une borne inférieure positive et une borne supérieure finie. En effet, dans ce cas, $p^{(n)}(F)$ est, pour

au moins une valeur de n , partout positif donc la condition (A) est réalisée et $p(F)$ est partout positif. D'autre part, les fonctions $P^{(n)}(F)$ sont bornées à partir d'un certain rang et par suite sommables.

La condition que V soit de mesure finie ne peut être évitée ici : $\int_V P^{(n)}(E, F) dF$ ne pourrait être égale à l'unité si $P^{(n)}(E, F)$ ayant une borne inférieure positive, la mesure de V était infinie. On peut cependant, en modifiant la condition, la rendre applicable au cas de V non borné :

On sera certainement dans le cas positivement régulier, si l'une des fonctions $P^{(n)}(F)$ est sommable sur V et si, en outre, l'une des fonctions $P^{(n)}(E, F)$ a une borne inférieure positive lorsque E variant sur V , F varie sur toute partie bornée W de V . Car, dans ce cas les conditions (A) et (B) sont visiblement vérifiées et l'une des $P^{(n)}(F)$ étant sommable sur V , on est dans le cas régulier. En outre, l'une des $p^{(n)}(F)$ — et, par suite, $p(F)$ — a une borne inférieure > 0 sur toute partie bornée de V . Donc $P(F)$ est partout $\neq 0$ sur V .

On a un résultat plus simple quand l'une au moins des densités $P^{(n)}(E, F)$ est uniformément continue sur V . On a vu (p. 183) qu'alors $p(F)$ y est aussi uniformément continue.

Supposons, en outre, pour commencer que V soit borné. Alors, comme on sait, $p(F)$ y atteint sa borne inférieure ; si donc $p(F) > 0$ sur V , $p(F)$ a sur V un minimum positif α . Or, si l'on est dans le cas quasi-régulier, $p^{(n)}(F)$ converge uniformément vers $p(F)$ (p. 188). Donc, pour n assez

grand, $p^{(n)}(F) > \frac{\alpha}{2}$ quel que soit F sur V et par suite $P^{(n)}(E, F) > 0$ pour E, F arbitraires sur V . Cette condition est d'ailleurs suffisante, car si $P^{(m)}(E, F) \neq 0$ sur V , son minimum $p^{(m)}(F)$ quand E varie étant atteint pour un certain état E sera $\neq 0$. En résumé :

Si l'une au moins des densités itérées $P^{(n)}(E, F)$ est uniformément continue quand E, F varient sur V et si le domaine V est borné, la condition nécessaire et suffisante pour que l'on soit dans le cas positivement régulier est que l'une au moins de ces densités $P^{(n)}(E, F)$ soit partout positive quand E, F varient sur V .

Abandonnons l'hypothèse que le domaine V soit borné. $p(F)$ sera encore uniformément continu, mais ce n'est que sur chaque partie bornée W de V qu'il atteindra son minimum. Ce minimum β_W est alors positif. Or, sur W , il y a convergence uniforme de $p^{(n)}(F)$ vers $p(F)$. Donc, pour n assez grand, $p^{(n)}(F) > \frac{\beta_W}{2} > 0$. Réciproquement si nous spécifions que sur toute partie bornée W de V , $p^{(n)}(F)$ reste positif pour

un certain rang, éventuellement variable avec W , alors la condition (A) sera réalisée, et $P(F)$ sera partout positif.

Ainsi, lorsque l'une au moins des densités $P^{(n)}(E, F)$ est uniformément continue quand E, F varient sur le domaine V — borné ou non — la condition nécessaire et suffisante pour qu'on soit dans le cas positivement quasi-régulier est que: 1° sur toute partie bornée W de V , l'un au moins des $p^{(n)}(F)$ reste positif sur W , à partir d'un certain rang, variable en général avec W mais indépendant de la position de F sur W . 2° $P(F)$ soit fini presque partout.

En particulier, si l'une, au moins, des densités $P^{(n)}(E, F)$ est bornée et uniformément continue quand E, F varient sur V , borné ou non, pour qu'on soit dans le cas positivement quasi-régulier, il faut et il suffit que, sur toute partie bornée W de V , l'un au moins des $p^{(n)}(F)$ reste positif.

Dans ce cas, il y a convergence uniforme de $P^{(n)}(E, F)$ vers $P(F)$ quand E, F varient sur V . De même, si l'une au moins des densités $P^{(n)}(E, F)$ est quand E, F varient sur V , borné ou non, uniformément continue et majorée par une fonction $\Psi(F)$ sommable sur V , la condition nécessaire et suffisante pour qu'on soit dans le cas positivement régulier est que pour toute partie bornée, W de V , l'un au moins, des $p^{(n)}(F)$ reste positif.

Condition de M. Hostinsky. En étendant au cas actuel sa remarque relative au cas discontinu, M. Hostinsky a observé (I, § 30, p. 50) que pour se trouver dans le cas positivement régulier, il suffit, si V est borné et $p(E, F)$ continu, que $p(E, F)$ soit $\neq 0$ lorsque, E, F variant dans V , la distance EF (voir la note (2) de la p. 180), reste inférieure à un certain nombre positif ρ (si petit soit-il choisi). La démonstration revient à prouver qu'au bout d'un nombre n assez grand d'épreuves $P^{(n)}(E, F)$ — qui est continu — est partout positif.

Toutefois, il est bien clair que si le domaine V n'est pas d'un seul tenant, si, par exemple, on peut y distinguer deux parties V_1 et V_2 dont les points sont à des distances toutes supérieures à un nombre positif μ , alors on peut imaginer d'une infinité de façons une répartition de la probabilité telle que le passage de V_1 à V_2 soit impossible et que le passage de n'importe quel état de V_k à n'importe quel état de V_k soit possible en une seule épreuve pour $k = 1, 2$. Alors, en prenant ρ inférieur à la plus courte distance de V_1 à V_2 , on aura bien $p(E, F) > 0$ pour $EF < \rho$ et pourtant on n'aura $P^{(n)}(E, F) > 0$ quand E appartient à V_1 et F à V_2 pour aucune valeur de n .

Reste donc à examiner le cas où V est d'un seul tenant; nous continuerons à supposer, en outre, que la frontière de V n'est pas trop irrégulière, au sens précisé p. 180 et que V est borné.

Observons d'abord que, V étant borné, si $p(E, F) > 0$ pour $EF < \rho$, alors $p(E, F)$ a une borne inférieure positive ε quand $EF \leq \theta \rho$, θ étant un nombre positif fixe < 1 . Sans quoi, il existerait pour tout entier n un couple E_n, F_n de V tel que $p(E_n, F_n) < \frac{1}{n}$ et $E_n F_n \leq \theta \rho$.

On pourrait extraire de la suite des entiers n une suite d'indices $n_1, n_2 \dots$ tels que $E_{n_1}, E_{n_2} \dots, F_{n_1}, F_{n_2} \dots$ tendent vers E_0, F_0 de V . On aurait $E_0 F_0 \leq \theta \rho$ et, puisque p est supposé continu $p(E_0, F_0) = 0$ contrairement à l'hypothèse.

Ceci étant, utilisons l'hypothèse que V est borné et d'un seul tenant. Nous entendrons par là que pour tout nombre η positif, il existe un entier N tel que pour tout couple d'états E, F de V , il existe des états $H_1, H_2, \dots, H_s$ de V en nombre $< N$ tels que $E H_1 < \eta, H_1 H_2 < \eta, \dots, H_s F < \eta$. Rien n'empêche, d'ailleurs, de supposer $s = N - 1$ car au cas où $s < N - 1$ on poserait $H_{s+1} \equiv \dots \equiv H_{N-1} \equiv H_s$. D'après l'hypothèse (p. 180) sur la frontière de V , dans toute sphère de centre H_j et de rayon $\frac{\eta}{2}$ il existe une portion v_j de V de mesure positive. Et si G_j, G_{j+1} appartiennent à v_j, v_{j+1} , on a

$$G_j G_{j+1} \leq G_j H_j + H_j H_{j+1} + H_{j+1} G_{j+1} < 2\eta$$

et en particulier $E G_1 < 2\eta, G_{N-1} F < 2\eta$. Prenons $2\eta = \theta \rho$. On aura

$$P^{(2)}(E, G_2) \geq \int_{v_1} p(E, G_1) p(G_1, G_2) dG_1 \geq \varepsilon^2 [\text{mes. } v_1]$$

$$P^{(3)}(E, G_3) \geq \int_{v_2} P^{(2)}(E, G_2) p(G_2, G_3) dG_2 \geq \varepsilon^3 (\text{mes. } v_1) (\text{mes. } v_2)$$

$$P^{(N)}(E, F) \geq \int_{v_{N-1}} P^{(N-1)}(E, G_{N-1}) p(G_{N-1}, F) dG_{N-1} \geq \varepsilon^N (\text{mes. } v_1) \dots (\text{mes. } v_{N-1}).$$

Le second membre peut varier quand E, F varient; mais, en tout cas, il est $\neq 0$, de sorte qu'il existe un entier N , indépendant de E et de F sur G , tel que

$$P^{(N)}(E, F) > 0.$$

On est bien dans le cas positivement régulier. Plus généralement, supposons qu'il existe un entier ν et un nombre $\rho > 0$ tels que $P^{(\nu)}(E, F) > 0$

pour $EF < \rho$. Alors en raisonnant sur $P^{(v)}(E, F)$ comme sur $p(E, F)$ on voit qu'il existe un entier N tel que $P^{(Nv)}(E, F) \neq 0$ quels que soient E, F sur V . On est alors encore dans le cas positivement régulier. D'ailleurs, si l'on est dans le cas positivement régulier, la dernière condition est bien nécessaire puisqu'elle est même réalisée pour toute valeur positive de ρ . Donc :

Si V est borné et d'un seul tenant et si l'une au moins, $P^{(m)}(E, F)$, des densités $P^{(n)}(E, F)$ est continue sur V , la condition nécessaire et suffisante pour qu'on soit dans le cas positivement régulier est qu'il existe un nombre positif ρ et un entier $v > m$ tels que $P^{(v)}(E, F) > 0$ pour $EF < \rho$.

Dans le cas où $p(E, F)$ n'est pas continu, on peut recommencer la démonstration, mais en faisant entrer cette fois dans la condition de M. Hostinsky ce qui en était déduit précédemment comme conséquence :

quand $p(E, F)$, ni aucun des $P^{(n)}(E, F)$ n'est supposé continu, la condition de M. Hostinsky consistera en ce que pour l'une au moins des densités itérées $P^{(n)}(E, F)$ — soit $P^{(v)}(E, F)$ — il existe un nombre $\rho > 0$ tel que $P^{(v)}(E, F)$ a une borne inférieure positive ε quand E, F varient arbitrairement sur V de sorte que $EF < \rho$.

Alors, en supposant encore V borné et d'un seul tenant, on verra comme plus haut qu'il existe un nombre Nv tel que

$$P^{(Nv)}(E, F) \geq \varepsilon^N (\text{mes. } v_1) \dots (\text{mes. } v_{N-1}).$$

Il en résulte que, pour n assez grand, $P^{(n)}(E, F)$ reste $\neq 0$ quels que soient E, F sur V . Cela ne veut pas dire que la borne inférieure $p^{(n)}(F)$ soit $\neq 0$, car le second membre varie avec E, F . Mais, V étant borné et d'un seul tenant, on peut certainement choisir sur V des états $K_1, \dots, K_N$ assez nombreux pour que tout état de V soit à distance $< \eta$ de l'un des K_i et pour qu'on puisse choisir les états $H_1, \dots, H_s$ de la démonstration précédente parmi les $K_1 \dots K_N$. Cette fois, le produit $\omega = \varepsilon^N (\text{mes. } v_1) \dots (\text{mes. } v_{N-1})$ sera indépendant de E et de F sur V et on aura en particulier $p^{(Nv)}(F) \geq \omega$, d'où $p^{(n)}(F) \geq \omega > 0$ pour $n > Nv$, quel que soit F sur V . Il en résulte que la condition (A) sera vérifiée ; d'ailleurs $p(F) \geq p^{(n)}(F) > 0$. Si la condition (B) est vérifiée, on aura donc non seulement le cas quasi-régulier, mais le cas positivement quasi-régulier.

En résumé : si V est borné et d'un seul tenant et si — condition de M. Hostinsky généralisée — $p(E, F)$ étant ou non continue — il existe un rang v et une distance ρ tels que $P^{(n)}(E, F)$ ait une borne inférieure $\neq 0$ quand $EF \leq \rho$, alors la condition nécessaire et suffisante pour qu'on soit dans le cas positivement quasi-régulier est que $P(F)$ soit fini presque partout sur V .

On peut aussi particulariser les énoncés de la p. 208 :

Si l'un des $P^{(n)}(E, F)$ a, en outre une borne supérieure finie quand E et F varient indépendamment sur V , ou si plus généralement l'un des $P^{(n)}(E, F)$ est majoré constamment par une fonction $\Psi(F)$ sommable sur V , alors, si la condition de M. Hostinsky, modifiée comme il vient d'être dit, est vérifiée, on est sûrement dans le cas positivement régulier.

Convergence des moyennes arithmétiques. Revenons à l'inégalité

$$(9) \quad |P^{(n)}(E, F) - p(F)| \leq (q_v)^n B(F) \quad \text{pour } n > m.$$

Puisque, dans le cas quasi-régulier $P^{(n)}(E, F)$ tend vers $p(F)$ sur l'ensemble V' où $P(F)$ et $p(F)$ sont finis et égaux, alors, en vertu d'une propriété connue des séries convergentes, il en sera de même de la moyenne arithmétique

$$\Pi^{(n)}(E, F) = \frac{P^{(1)}(E, F) + \dots + P^{(n)}(E, F)}{n}.$$

Non seulement la différence $\eta_n(E, F) = \Pi^{(n)}(E, F) - p(F)$ est sur V' infiniment petite avec $\frac{1}{n}$, mais l'inégalité (9) va nous permettre d'obtenir un renseignement sur son ordre. En effet, la série

$$s(E, F) = \sum_{t=1}^{t=\infty} [P^{(t)}(E, F) - p(F)]$$

est d'après (9) convergente quand F est sur V' et cela uniformément quand E varie sur V . Or on a évidemment

$$\lim_{n \rightarrow \infty} n [\Pi^{(n)}(E, F) - p(F)] = s(E, F)$$

c'est-à-dire que : ou bien $s(E, F)$ est fini et $\neq 0$ et $\eta_n(E, F)$ est un infiniment petit du premier ordre en $\frac{1}{n}$, avec une partie principale égale à $s(E, F)$; ou bien $s(E, F) = 0$ et $\eta_n(E, F)$ est un infiniment petit d'ordre supérieur à $\frac{1}{n}$. Il faut cependant observer qu'il reste une troisième hypothèse, celle pour celle où $s(E, F)$ serait infini. Bien que restant convergente sur V' à partir d'un certain rang, il peut en effet arriver qu'avant ce rang, certains termes de $s(E, F)$ soient infinis. Cette difficulté disparaît quand au lieu des densités, on étudie les probabilités elles-mêmes.

Comportement des probabilités. Passons de l'étude des *densités* de probabilités à celle des probabilités. Soit v une portion quelconque de V , de mesure positive. La probabilité qu'on passe en n épreuves de l'état E à l'un quelconque F des états de v est:

$$\omega_v^{(n)}(E) = \int_v P^{(n)}(E, F) dF$$

$$\text{Soit} \quad A_n = \omega_v^{(n)}(E) - \int_v p(F) dF = I_n - \mathcal{J}_n$$

avec

$$\begin{aligned} 0 &\leq I_n = \int_v [P^{(n)}(E, F) - p^{(n)}(F)] dF \\ &\leq \int_v [P^{(n)}(E, F) - p^{(n)}(F)] dF = 1 - \int_v p^{(n)}(F) dF \end{aligned}$$

et de même

$$0 \leq \mathcal{J}_n = \int_v [p(F) - p^{(n)}(F)] dF \leq \int_v p(F) dF - \int_v p^{(n)}(F) dF.$$

Nous avons vu p. 198 que dans tous les cas ce dernier membre et par suite aussi $\mathcal{J}_n$ tendent vers zéro avec $\frac{1}{n}$. Le comportement de A_n quand n croît est donc le même que celui de I_n . Donc, *dans le cas régulier* A_n tend vers zéro, c'est-à-dire que la *probabilité cherchée* $\omega_v^{(n)}(E)$ *tend vers une limite indépendante de* E et qui est égale à $\int_v p(F) dF$.

Si l'on n'est pas dans le cas régulier, alors, par définition il n'en est pas ainsi déjà pour $v \equiv V$: $1 = \omega_V^{(n)}(E)$ ne tendra pas vers $\int_V p(F) dF$.

On vient de voir que dans le cas régulier $\omega_v^{(n)}(E)$ tend vers $\int_v p(F) dF$.

Il en est donc de même de la moyenne arithmétique $f_v^{(n)}(E) = \frac{1}{n} \sum_{t=1}^{t=n} \omega_v^{(t)}(E)$.

Non seulement la différence $\varepsilon_n(E) = f_v^{(n)}(E) - \int_v p(F) dF$ est infiniment petite avec $\frac{1}{n}$, mais nous pouvons avoir des renseignements sur son ordre. On a en effet

$$n \varepsilon_n(E) = \sum_{t=1}^{t=n} \left[\int_v P^{(t)}(E, F) dF - \int_v p(F) dF \right].$$

Tout d'abord, chacun des crochets est, en valeur absolue $\leq \int_V P^{(t)}(E, F) dF + \int_V p(F) dF \leq 2$. D'autre part si v appartient à V_s l'ensemble V_s où $P^{(m)}(F) < r$, avec $r + m = s$, on aura

$$\left| \int_V P^{(n)}(E, F) dF - \int_V p(F) dF \right| \leq (q_v)^n \frac{s v}{(q_v)^{s+v}} \quad \text{pour } n > m.$$

La série

$$(10) \quad \sigma_v(E) = \sum_{t=1}^{t=\infty} \left[\int_V P^{(t)}(E, F) dF - \int_V p(F) dF \right]$$

a donc ses termes tous finis quand E parcourt V et elle est majorée par une série convergente indépendante de E , quand v appartient à V_s . Ainsi quand v est contenu dans l'un des ensembles $V_1, V_2, \dots, V_s, \dots$ (dont la réunion forme V'), $\sigma_v(E)$ est une fonction bornée sur V et alors pour chaque état initial E , ou bien $\sigma_v(E) \neq 0$ et alors, non seulement ε_n est infiniment petit avec $\frac{1}{n}$, mais son ordre est égal à l'unité et sa partie principale est $\frac{1}{n} \sigma_v(E)$; ou bien $\sigma_v(E) = 0$ et ε_n est d'ordre supérieur à l'unité.

Remarque. Si l'on voulait calculer la partie principale du moyen de la formule (10), il faudrait auparavant calculer toutes les densités itérées $P^{(n)}(E, F)$. On verra plus loin, p. 227, qu'on peut obvier à cet inconvénient.

Valeur de la densité limite $P(F)$

Cas quasi-régulier. Nous avons vu que, dans ce cas, $P(F)$ et $p(F)$ sont finis et égaux partout sur V , sauf peut-être sur un certain ensemble w de mesure nulle. De plus $P^{(n)}(E, F)$ converge vers $P(F)$ quand n tend vers l'infini, si F est sur $V - w$. Pour essayer de déterminer $P(F)$, il paraît indiqué de passer à la limite dans les formules

$$(I_1) \quad P^{(n+1)}(E, F) = \int_{V-w} P^{(n)}(E, G) p(G, F) dG.$$

$$(T) \quad \int_{V-w} P^{(n)}(E, F) dF = 1.$$

Toutefois, nous avons déjà vu qu'on n'a pas le droit, sans précaution, de substituer dans (T) à $P^{(n)}(E, F)$, sa limite $P(F)$. Ainsi, dans l'exemple de la p. 189, $P(F) \equiv 0$. Et même nous avons réservé le nom de cas régulier au cas où cette substitution est légitime, c'est-à-dire où

$$\int_V P(F) dF = 1.$$

On a vu que, dans le cas quasi-régulier, on peut seulement écrire

$$\int_V P(F) dF \leq 1.$$

Considérons de même le passage à la limite dans (I₁) ou plus généralement dans la formule

$$(I) \quad P^{(m+n)}(E, F) = \int_V P^{(n)}(E, G) P^{(m)}(G, F) dG.$$

Si la substitution était légitime à la limite, l'expression

$$P(F) - \int_V P(G) P^{(m)}(G, F) dG$$

serait nulle.

Pour qu'elle ait d'abord un sens, il faut que $P(F)$ soit fini.

Si F appartient à l'ensemble $V' = V - w$, $P(F)$ est fini et, par suite, pour n assez grand, $P^{(m+n)}(E, F)$ est fini. Alors $\int_V P^{(n)}(G) P^{(m)}(G, F) dG \leq \int_V P^{(n)}(E, G) P^{(m)}(G, F) dG = P^{(m+n)}(E, F)$. Pour ϵ positif donné, on aura, pour $n > n'$, $P^{(m+n)}(E, F) < P(F) + \epsilon$. Donc $P^{(n)}(G) P^{(m)}(G, F)$ est une fonction (≥ 0) de G qui est sommable sur V' ; elle reste sur V' au plus égale à $P(G) P^{(m)}(G, F)$, tend sans décroître vers cette fonction quand n croît et son intégrale sur V' reste inférieure ou égale à $P(F) + \epsilon$. Dès lors $P(G) P^{(m)}(G, F)$ est sommable sur V' et son intégrale sur V' est $\leq P(F) + \epsilon$. Comme ϵ est arbitraire, on voit finalement que: *dans le cas quasi-régulier, pour tout état F de V' , $P(G) P^{(m)}(G, F)$ est sommable sur V' et l'on a, sur V'*

$$(II) \quad \int_V P(G) P^{(m)}(G, F) dG \leq P(F).$$

En particulier,

$$(12) \quad \int_V P(G) p(G, F) dG \leq P(F).$$

On peut même préciser la variation avec m de l'expression

$$\omega^{(m)}(F) = \int_V P(G) P^{(m)}(G, F) dG.$$

On a vu que $\omega^{(m)}(F) \leq P(F).$

Or
$$\begin{aligned} \omega^{(m+1)}(F) &= \int_V P(G) \left[\int_V P^{(m)}(G, H) p(H, F) dH \right] dG \\ &= \int_V \left[\int_V P(G) P^{(m)}(G, H) dG \right] p(H, F) dH = \int_V \omega^{(m)}(H) p(H, F) dH. \end{aligned}$$

Donc
$$\omega^{(2)}(F) = \int_V \omega^{(1)}(H) p(H, F) dH \leq \int_V P(H) p(H, F) dH = \omega^{(1)}(F).$$

et en général si $\omega^{(m)}(F) \leq \omega^{(m-1)}(F)$, alors

$$\omega^{(m+1)}(F) = \int_V \omega^{(m)}(H) p(H, F) dH \leq \int_V \omega^{(m-1)}(H) p(H, F) dH = \omega^{(m)}(F).$$

Dès lors, $\omega^{(m)}(F)$ est une fonction de m non croissante et ≥ 0 . Elle a donc une limite $\omega(F) \geq 0$ et l'on a sur V'

$$(13) \quad \omega(F) \leq \dots \leq \omega^{(n+1)}(F) \leq \omega^{(n)}(F) \leq \dots \leq \omega^{(1)}(F) \leq P(F).$$

La question se pose maintenant de savoir si la relation (11) qui a lieu partout où elle a un sens, c'est-à-dire quand F est sur V' peut devenir une égalité, une égalité permettant de concourir à la détermination de $P(F)$ partout où $P(F)$ est défini, c'est-à-dire une égalité ayant lieu partout sur V' .

Observons d'abord que, si l'on a pour une valeur particulière de m et quel que soit F sur V' :

$$(\mathcal{F}_m) \quad \int_{V'} P(G) P^{(m)}(G, F) dG = P(F)$$

alors, en multipliant par $P^{(m)}(F, H) dF$ et intégrant, on aura si H est sur V'

$$\int_{V'} P(G) P^{(2m)}(G, H) dG = P(H)$$

et de même, en général

$$\int_{V'} P(G) P^{(rm)}(G, H) dG = P(H)$$

quel que soit l'entier r . On a alors $\omega^{(n)}(H) = P(H)$ sur V' pour une infinité de valeurs rm de n et par suite à la limite $\omega(H) = P(H)$ sur V' . En raison des relations (13), on voit donc que $\omega^{(n)}(F) = P(H)$ sur V' quel que soit n . Ainsi, lorsque l'égalité $(\mathcal{I}_m)$ a lieu pour une valeur de m , elle a lieu pour toute valeur entière de m . Or, si l'on a l'égalité $(\mathcal{I}_m)$ en un état F de V' , comme d'après le raisonnement fait plus haut, on a

$$(14) \quad \lim_{n \rightarrow \infty} \int_{V'} p^{(n)}(G) P^{(m)}(G, F) dG = \int_{V'} P(G) P^{(m)}(G, F) dG$$

on aura

$$0 = \lim_{n \rightarrow \infty} [P^{(m+n)}(E, F) - p(F)] = \lim_{n \rightarrow \infty} \int_{V'} [P^{(n)}(E, G) - p^{(n)}(G)] P^{(m)}(G, F) dG.$$

La dernière intégrale est

$$\geq p^{(m)}(F) [1 - \int_{V'} p^{(n)}(G) dG].$$

On a donc, à la limite

$$0 \geq p^{(m)}(F) [1 - \int_{V'} P(G) dG] \geq 0.$$

Dès lors si l'on est dans le cas quasi-régulier *proprement dit*, celui où $\int_{V'} P(G) dG < 1$, on devra avoir $p^{(m)}(F) = 0$.

Or, si l'équation $(\mathcal{I}_m)$ a lieu presque partout sur V , il en est de même de $(\mathcal{I}_{rm})$ quel que soit l'entier r , donc $p^{(rm)}(F) = 0$ et par suite, à la

limite, $P(F) = 0$ presque partout sur V . En résumé, *dans le cas quasi-régulier proprement dit*, $P(F)$ vérifie la relation (11) quel que soit m sur V' , mais $P(F)$ n'est, pour aucune valeur de m , une solution non identiquement nulle de l'équation intégrale homogène de Fredholm ($\mathcal{F}_m$).

Ainsi, le passage à la limite qui fournit ($\mathcal{F}_m$) n'est jamais à la fois légitime et utile que dans le cas régulier.

Cas régulier. D'après la formule (14), on a

$$P(F) = \int_V P(G) P^{(m)}(G, F) dG = \lim_{n \rightarrow \infty} A_n(F)$$

avec, d'après (9)

$$\begin{aligned} 0 \leq A_n(F) &= \int_V [P^{(n)}(E, G) - p^{(n)}(G)] P^{(m)}(G, F) dG \\ &\leq P^{(m)}(F) \left[1 - \int_V p^{(n)}(G) dG \right] \end{aligned}$$

si $P^{(m)}(F)$ est fini. Si donc, on est dans le cas régulier, on aura

$$P(F) = \int_V P(G) P^{(m)}(G, F) dG$$

pour tout état F où $P^{(m)}(F)$ est fini.

Or, pour tout état F de V' , il y a un entier N tel que $P^{(m)}(F)$ soit fini pour $m > N$. On a donc $\omega^{(m)}(F) = P(F)$ pour $m > N$ et par suite $\omega(F) = P(F)$. Ainsi, pour F sur V'

$$P(F) = \omega(F) \leq \omega^{(n)}(F) \leq \omega^{(1)}(F) \leq P(F).$$

On a donc $\omega^{(n)}(F) = P(F)$ sur V' quel que soit n .

En résumé, *dans le cas régulier, la densité-limite $P(F)$ est une solution, vérifiant les conditions*

$$(15) \quad \int_{V'} P(F) dF = 1 \quad P(F) \geq 0,$$

de l'équation intégrale homogène de Fredholm

$$(\mathcal{F}) \quad P(F) = \int_{V'} P(G) p(G, F) dG$$

et de ses équations „itérées“

$$(\mathcal{F}_m) \quad P(F) = \int_{V'} P(G) P^{(m)}(G, F) dG.$$

Les conditions (15) montrent d'ailleurs que *c'est une solution „effective“ de ($\mathcal{F}$)*, c'est-à-dire une solution non identiquement nulle sur V' .

Cas de la densité-limite constante. On a vu que, dans le cas quasi-régulier, $P^{(n)}(E, F)$ converge presque partout sur V vers une limite indépendante de E . On peut se demander dans quel cas cette limite sera aussi, quand E varie presque partout sur V , indépendante de F . Alors, la limite étant indépendante de E et de F aura une valeur constante P . On a d'ailleurs vu qu'on a toujours dans le cas quasi-régulier

$$(16) \quad 0 \leq \int_V P(F) dF \leq 1.$$

Ceci montre aussitôt que *dans ce cas si la mesure de V est infinie, $P(F)$ ne peut être constant que s'il est nul.*

Si, au contraire, V est borné ou de mesure finie, on aura

$$P \leq \frac{1}{\text{mes. } V}.$$

Supposons $P > 0$ et, par suite, V de mesure finie. On a vu (p. 217, pour $m = 1$) que dans le cas quasi-singulier, si F est sur V' , $P(G) p(G, F)$ est une fonction de G sommable sur V . Donc, ici, l'intégrale $\int_V p(G, F) dG$ doit avoir une valeur finie. D'autre part, on a sur V'

$$(12) \quad P(F) \geq \int_V P(G) p(G, F) dG.$$

Donc ici

$$(17) \quad 1 \geq \int_V p(G, F) dG.$$

En multipliant par dF et intégrant

$$\begin{aligned} \text{mes. } V &\geq \int_V \left[\int_V p(G, F) dG \right] dF \\ &= \int_V \left[\int_V p(G, F) dF \right] dG = \text{mes. } V. \end{aligned}$$

L'avant-dernière ligne est donc une égalité et l'on a

$$\int_V \left\{ 1 - \int_V p(G, F) dG \right\} dF = 0.$$

L'accolade étant ≥ 0 , en vertu de (17), doit donc être nulle presque partout. Ainsi, on a la condition

$$(T_1') \quad \int_V p(G, F) dG = 1, \quad \text{presque partout sur } V.$$

Observons d'ailleurs que, s'il en est ainsi, la relation (17) devient une égalité sur V' . Or nous savons (p. 220) que pour $P > 0$, il ne peut en être ainsi dans le cas quasi-régulier proprement dit. Ainsi, pour que $P(F)$ soit presque partout égal à une constante positive, il faut :

- 1° qu'on soit dans le cas régulier,
- 2° que la condition (T_1') soit réalisée,
- 3° que V soit de mesure finie.

Et alors la limite constante est égale à $\frac{1}{V}$. Car on a $\int_V P(F) dF = 1$ d'après 1°.

Réciproquement, supposons d'abord simplement vérifiée la condition (T_1') et posons

$$(18) \quad \begin{aligned} L_n(F) &= \int_V P^{(n)}(G, F) dG, \text{ d'où} \\ L_{n+1}(F) &= \int_V \left[\int_V P^{(n)}(G, H) p(H, F) dH \right] dG \\ &= \int_V L_n(H) p(H, F) dH \end{aligned}$$

et $L_1(F) = 1$ presque partout. Si $L_m(F) = 1$ presque partout pour $m \leq n$, on aura, d'après (18)

$$L_{n+1}(F) = \int_V p(H, F) dH = 1 \text{ presque partout.}$$

Ainsi, si (T_1') a lieu, on a nécessairement aussi

$$(T') \quad \int_V P^{(n)}(G, F) dG = 1 \quad \text{presque partout sur } V,$$

quel que soit l'entier n .

Supposons, en outre, maintenant qu'on soit au moins dans le cas quasi-singulier, de sorte qu'il y ait d'abord presque partout une limite

$P(F)$ de $P^{(n)}(E, F)$ indépendante de E . En vertu de (T'), on a, presque partout sur V

$$1 \geq \int_V p^{(n)}(F) dG = p^{(n)}(F) \text{ mes. } V.$$

Dès lors, si la mesure de V était infinie, on déduirait de 2° que $p^{(n)}(F) = 0$ presque partout sur V et par suite que $p(F)$ est nul presque partout sur V .

Si donc la condition 2° est réalisée: ou bien la densité limite $P(F)$ est nulle presque partout sur V , ou bien V ne peut être de mesure infinie et la condition 3° est réalisée. Si $P(F)$ était nul presque partout sur V , on ne serait pas dans le cas régulier. Donc, si 1° et 2° sont réalisées, il en est de même de 3°.

Or, dans cette hypothèse, on déduit de la dernière formule que

$$P(F) \leq \frac{1}{\text{mes. } V}$$

et en multipliant par dF et intégrant

$$0 \leq \int_V \left[\frac{1}{\text{mes. } V} - P(F) \right] dF = 1 - \int_V P(F) dF = 0.$$

Le crochet est ≥ 0 et son intégrale sur V est nulle. Donc, il est nul presque partout, c'est-à-dire que $P(F)$ est presque partout égal à la constante $\frac{1}{V}$.

En particulier, s'il existe une fonction $\Psi(F)$ sommable sur V et qui majore, quels que soient E et F sur V , l'une des densités itérées $P^{(n)}(E, F)$, alors la condition nécessaire et suffisante pour que la suite des $P^{(n)}(E, F)$ converge vers une constante non nulle quand F varie presque partout sur V (et cela uniformément quand E varie sur V) est: 1° que la condition (T₁') soit vérifiée, 2° que pour une assez grande valeur de v , $p^{(v)}(F)$ ne soit pas nul presque partout sur V .

Plus particulièrement, si le domaine V est borné et si les densités itérées $P^{(n)}(E, F)$ sont continues sur V à partir d'un certain rang n , alors la condition nécessaire et suffisante pour que la suite des $P^{(n)}(E, F)$ converge vers une constante non nulle et cela uniformément quand E, F varient arbitrairement sur V , est que

1° $\int_V p(E, F) dE = 1$ quand F est arbitraire sur V .

2° Il existe au moins un état L de V tel que pour v assez grand, $P^{(v)}(E, L)$ ne s'annule jamais quand E parcourt V ou que $P^{(v)}(L, E)$ ne s'annule jamais quand E parcourt V .

Cas du domaine infini. Nous avons vu que si le domaine est de mesure infinie, on ne peut, dans le cas régulier, avoir pour densité limite une constante. Si le domaine étant infini, la condition (T') est cependant réalisée, on aura, presque partout sur V ,

$$1 = \int_V P^{(n)}(E, F) dE \geq \int_V p^{(n)}(F) dE \geq 0.$$

Si $p^{(n)}(F)$ n'était pas nul, la dernière intégrale serait infinie. Donc $p^{(n)}(F)$ est nul, presque partout sur V , quel que soit l'entier n et par suite $p(F) = 0$, presque partout sur V , qu'on soit ou non dans le cas quasi-régulier. Dès lors :

Quand, le domaine V étant de mesure infinie, la condition (T') est réalisée, $p(F)$ est nul presque partout sur V et l'on ne peut être dans le cas quasi-régulier que si la densité limite $P(F) = p(F)$ est nulle presque partout sur V .

Exemples de réalisation de T'. On peut signaler deux catégories simples de fonctions $p(E, F)$ pour lesquelles la condition (T') est réalisée nécessairement quand la condition (T) l'est déjà.

La première catégorie est formée des fonctions *symétriques*, c'est-à-dire telles que $p(E, F) \equiv p(F, E)$.

Une autre catégorie (qui n'a pas son analogue dans le cas d'un nombre fini d'états possibles) est celle où — le système aléatoire, ne dépendant que d'un paramètre numérique pouvant prendre toutes les valeurs possibles, et où par suite $p(E, F)$ se réduisant à une fonction $p(x, y)$ de deux variables numériques —, on suppose que $p(x, y)$ est une fonction de $y - x$, soit $f(y - x)$.

La condition

$$(T) \quad \int_{-\infty}^{+\infty} f(y - x) dy = 1$$

étant supposée vérifiée, il suffit d'observer, en posant $y - x = u$ que l'on a toujours

$$\int_{-\infty}^{+\infty} f(y - x) dy = \int_{-\infty}^{+\infty} f(u) du = \int_{-\infty}^{+\infty} f(y - x) dx$$

d'où la condition

$$(T') \quad \int_{-\infty}^{+\infty} f(y-x) dx = 1.$$

Dans le cas où les états du système sont définis par k paramètres:

$$p(E, F) = \Phi(q_1, \dots, q_k; u_1, \dots, u_k),$$

la condition (T) entraînera encore d'elle-même la condition (T') si Φ ne dépend de $q_1, \dots, u_k$ que par l'intermédiaire des différences correspondantes $(q_1 - u_1), \dots, (q_k - u_k)$.

Mais, comme nous l'avons vu plus haut, ce second cas ne peut conduire au cas régulier puisque V est ici illimité. Plus précisément, on aura $p(F) = 0$ et: ou bien on n'aura même pas le cas quasi-régulier c'est-à-dire que $p(F) < P(F)$ sur un ensemble de mesure positive, ou bien — comme dans l'exemple de la p. 189 — la limite $P(F) = p(F)$ de $P^{(n)}(E, F)$ sera nulle sur V' .

Expression de la densité-limite $P(F)$. Laissons maintenant de côté la condition (T'). On a vu que, dans le cas régulier, la limite $P(F)$ de $P^{(n)}(E, F)$ vérifie sur V' les conditions:

$$(\mathcal{J}) \quad P(F) = \int_{V'} P(G) p(G, F) dG$$

$$(15) \quad \int_{V'} P(G) dG = 1.$$

Ainsi, dans le cas régulier, l'équation intégrale

$$(H) \quad X(F) = \int_{V'} X(G) p(G, F) dG$$

a, au moins, une solution non identiquement nulle et même partout ≥ 0 : $X(F) = P(F)$ et l'une au moins des solutions vérifie la condition

$$\int_{V'} X(G) dG = 1.$$

Nous allons même démontrer que l'équation intégrale homogène de Fredholm (H) ne peut avoir, dans le cas régulier, qu'une solution

sommable, non identiquement nulle $X_0(F)$ à un facteur constant près. S'il en est ainsi, on aura nécessairement sur V' $P(F) = \alpha X_0(F)$ et on déterminera la constante α par la relation $1 = \alpha \int_{V'} X_0(F) dF$. Cette relation montre que $\int_{V'} X_0(F) dF \neq 0$ et par suite la résolution classique de l'équation de Fredholm permet de déterminer entièrement $P(F)$ par la formule

$$P(F) = \frac{X_0(F)}{\int_{V'} X_0(F) dF}.$$

Pour démontrer le point admis, observons que l'équation (H) ayant au moins une solution non identiquement nulle $X_0(F)$, alors en itérant (H), on voit que $X_0(F)$ vérifie aussi, quel que soit n , l'équation

$$(H_n) \quad X_0(F) = \int_{V'} P^{(n)}(G, F) X_0(G) dG.$$

Quand n croît indéfiniment, on peut passer à la limite sous le signe $\int$. Car

$$\begin{aligned} & \left| \int_{V'} P^{(n)}(G, F) X_0(G) dG - \int_{V'} P(F) X_0(G) dG \right| \\ & \leq \int_{V'} |P^{(n)}(G, F) - P(F)| |X_0(G)| dG \\ & \leq [P^{(n)}(F) - p^{(n)}(F)] \int_{V'} |X_0(G)| dG. \end{aligned}$$

On a donc $X_0(F) = P(F) \int_{V'} X_0(G) dG$ et il n'y a pas d'autre solution de (H) que les fonctions proportionnelles à $P(F)$.

Ainsi, dans le cas régulier, 1° l'équation de Fredholm (H) a une solution non identiquement nulle et une seule, $X_0(F)$ à un facteur constant près; 2° la recherche de la densité limite est ramenée à la résolution classique de cette équation de Fredholm et on a $P(F) = \frac{X_0(F)}{\int_{V'} X_0(F) dF}$.

Calcul de $s(E, F)$. Supposons, pour simplifier que $p(E, F)$ ait une borne supérieure finie quand E et F varient sur V . Alors la série

$$s(E, F) = \sum_{t=1}^{t=\infty} [P^{(t)}(E, F) - P(F)]$$

est majorée par une série convergente indépendante de E et de F variant sur tout V .

Or, dans le cas régulier :

$$P^{(t)}(E, F) - P(F) = \int_V [P^{(t-1)}(E, G) - P(G)] p(G, F) dG$$

d'où :

$$[P^{(2)}(E, F) - P(F)] + \dots + [P^{(n)}(E, F) - P(F)] = \int_V \{ [P^{(1)}(E, G) - P(G)] + \dots + [P^{(n-1)}(E, G) - P(G)] \} p(G, F) dG.$$

et à la limite

$$s(E, F) - P(E, F) + P(F) = \int_V s(E, G) p(G, F) dG.$$

Ainsi le calcul de $s(E, F)$ est ramené à la résolution d'une équation intégrale de Fredholm.

$$Y(E, F) = P(E, F) - P(F) + \int_V Y(E, G) p(G, F) dG$$

dont on vient de prouver qu'elle admet au moins une solution. D'ailleurs, une solution quelconque $Y_0(E, F)$ n'est pas nécessairement égale à $s(E, F)$; mais la différence de ces deux solutions vérifie évidemment l'équation (H), de sorte qu'elle est de la forme $Y_0(E, F) - s(E, F) = \beta(E) X_0(F)$.

Or, on a, avec convergence uniforme

$$\int_V s(E, F) dF = \sum_{t=1}^{t=\infty} \left[\int_V P^{(t)}(E, F) dF - \int_V P(F) dF \right] = \Sigma(1 - 1) = 0.$$

$$\text{D'où } \beta(E) = \frac{\int_V Y_0(E, F) dF}{\int_V X_0(F) dF},$$

$$\text{et finalement } s(E, F) = Y_0(E, F) - \frac{\int_V Y_0(E, F) dF}{\int_V X_0(F) dF} X_0(F).$$

Ceci fait, on calculera la partie principale $\frac{1}{n} \sigma_v(E)$ de $f_v^{(n)}(E) = \int_V P(F) dF$ (p. 215) par la formule $\sigma_v(E) = \int_V s(E, F) dF$.

Remarque I. Le raisonnement qui précède nous fournit, chemin faisant, une condition nécessaire pour qu'on soit dans le cas régulier sous une forme qui peut être commode :

Pour qu'on soit dans le cas régulier, il faut que l'équation intégrale homogène de Fredholm (H) admette une solution partout ≥ 0 , non identiquement nulle et qu'elle n'admette qu'une telle solution à un facteur constant près.

Remarque II. On a vu que dans le cas régulier l'équation intégrale homogène

$$(H_\lambda) \quad X(F) = \lambda \int_{V'} X(G) p(G, F) dG$$

a au moins une solution effective $P(F)$ pour $\lambda = 1$. C'est ce qu'on exprime en disant que l'unité est constante caractéristique pour l'équation (H). Il n'est pas nécessaire pour obtenir ce résultat d'avoir établi que la limite $P(F)$ vérifie l'équation (H). En effet, la condition (T_1) montre immédiatement que l'équation intégrale homogène associée à (H_λ) , soit

$$(H'_\lambda) \quad Y(F) = \lambda \int_{V'} Y(G) p(F, G) dG$$

admet pour $\lambda = 1$, la solution effective $Y(G) \equiv 1$. L'unité est donc constante caractéristique de (H'_λ) et ceci qu'on soit dans le cas régulier ou non. On sait qu'alors, si les théorèmes de Fredholm s'appliquent à (H_λ) , l'unité sera aussi constante caractéristique de (H_λ) .

Autrement dit, la théorie de Fredholm nous aurait permis d'établir directement l'existence d'au moins une solution effective de (H), aussi bien dans le cas singulier que dans le cas régulier. Mais elle ne suffisait pas pour établir que dans le cas régulier $P(F)$ est déterminé par les conditions (15) et $(\mathcal{F})$. Et de plus, on n'a pas le droit d'appliquer cette théorie à tout noyau et tout domaine. On sait en particulier que son extension au cas des domaines illimités et aux noyaux discontinus ne peut se faire sans de sérieuses restrictions.

Valeurs moyennes de variables et de fréquences aléatoires dépendant d'événements „en chaîne“. Estimation de leurs dispersions

Valeur moyenne d'une variable aléatoire. Supposons qu'à chaque état possible F corresponde une valeur déterminée $Y(F)$ d'une certaine

variable. C'est une variable aléatoire dont la valeur dépend de la réalisation de l'épreuve qui doit fournir un des états possibles.

Appelons $X^{(n)}(E)$ la valeur aléatoire prise par cette variable quand n épreuves ont eu lieu à partir de l'état E . Si, par exemple, ces n épreuves ont abouti à l'état F , on aura $X^{(n)}(E) = Y(F)$. Mais quand E , n sont donnés, F n'est pas déterminé et les deux membres de cette égalité ont une valeur aléatoire dont nous désignerons la valeur moyenne par

$$\mathfrak{M} X^{(n)}(E) = \overline{Y(F)}.$$

Avec les notations précédemment employées, cette valeur est égale à

$$\int_V Y(F) P^{(n)}(E, F) dF$$

et le problème à résoudre est d'abord de trouver comment se comporte cette quantité quand n croît.

Si nous nous plaçons dans le cas quasi-régulier alors $P^{(n)}(E, F)$ converge vers $P(F)$ sauf peut-être sur un ensemble w de mesure nulle. Or on a en posant $V' = V - w$

$$(19) \quad \mathfrak{M} X^{(n)}(E) = \int_{V'} Y(F) P^{(n)}(E, F) dF$$

et le problème est de savoir s'il est légitime de passer à la limite dans le second membre sous le signe d'intégration.

Mais, déjà dans le cas simple où $Y(F)$ est une constante non nulle, nous savons que si l'on n'est pas dans le cas régulier

$$\int_{V'} Y(F) [\lim P^{(n)}(E, F)] dF = \int_{V'} Y(F) P(F) dF = Y(F) \int_{V'} P(F) dF$$

$$\neq Y(F) = \lim \int_{V'} Y(F) P^{(n)}(E, F) dF.$$

Il en est encore de même dans le cas plus général où $Y(F)$ garde un signe constant et reste en valeur absolue supérieur à une borne positive. Car si $Y(F)$ reste positif et supérieur à ε

$$(20) \quad \int_{V'} Y(F) P^{(n)}(E, F) dF - \int_{V'} Y(F) P(F) dF = \\ \int_{V'} Y(F) [P^{(n)}(E, F) - p^{(n)}(F)] dF - \int_{V'} Y(F) [P(F) - p^{(n)}(F)] dF.$$

La question du passage à la limite sous l'intégrale ne peut se poser que si les intégrales du premier membre sont finies. Dans ce cas $Y(F) p^{(n)}(F)$ serait aussi sommable sur V' et au plus égale à une fonction $Y(F) P(F)$ sommable sur V' , vers laquelle elle tend sur V' sans décroître. Donc la dernière intégrale du second membre tend vers zéro. Or

$$\int_{V'} Y(F) [P^{(n)}(E, F) - p^{(n)}(F)] dF \geq \varepsilon \left[1 - \int_{V'} p^{(n)}(F) dF \right] \geq 0.$$

Le premier membre de (20) ne peut donc tendre vers zéro que si la limite $\int_{V'} p(F) dF$ de $\int_{V'} p^{(n)}(F) dF$ est égale à 1, c'est-à-dire dans le cas régulier.

Si $Y(F)$ peut changer de signe, il pourrait se produire des compensations de signe permettant le passage à la limite de (19) dans le cas non régulier, pour des fonctions $Y(F)$ particulières. Mais nous savons maintenant que cela n'aura pas lieu pour des fonctions choisies parmi les plus simples. Dans la suite, nous allons donc nous borner au cas régulier.

Si l'on suppose que $Y(F)$ est borné sur V ou au moins sur V' , si, par exemple, on a $|Y(F)| \leq A$ sur V' , alors en vertu de (20)

$$(21) \quad \left| \int_{V'} Y(F) P^{(n)}(E, F) dF - \int_{V'} Y(F) P(F) dF \right| \\ \leq 2A \left[1 - \int_{V'} p^{(n)}(F) dF \right].$$

Et par suite le premier membre tend vers zéro.

Donc : si la fonction $Y(F)$ est bornée sur V , alors dans le cas régulier $\mathfrak{M} X^{(n)}(E)$ tend vers une limite quand n croît indéfiniment et cette limite est la quantité indépendante de E

$$M = \int_V Y(F) P(F) dF.$$

Cette quantité limite de valeurs moyennes est, elle-même, une *valeur moyenne* de $Y(E)$ puisque $P(F) \geq 0$ et $\int P(F) dF = 1$ et c'est même la valeur moyenne de $Y(F)$ qui correspond à la densité-limite $P(F)$.

Enfin, dans ce même cas, la convergence de $\mathfrak{M} X^{(n)}(E) - M$ vers zéro est, d'après (21) jointe à la formule (19) de la p. 229, uniforme quand E varie.

Observons d'ailleurs que

$$\mathfrak{M} X^{(m+n)}(E) = \int_V P^{(m)}(E, G) [\mathfrak{M} X^{(n)}(G)] dG.$$

Si donc $Y(E)$ n'étant pas borné sur V , $\mathfrak{M} X^{(n)}(E)$ est borné sur V pour une valeur convenable m de n , les résultats précédents subsistent puisqu'en remplaçant $Y(E)$ par $\mathfrak{M} X^{(m)}(G)$ on retrouve à un décalage près la même suite de valeurs moyennes.

Or, si le passage à la limite est légitime c'est que $M = \int_V Y(F) P(F) dF$ est finie et si la convergence de $\mathfrak{M} X^{(n)}(E)$ vers M est uniforme quand E varie, alors, pour n assez grand $\mathfrak{M} X^{(n)}(E)$ est nécessairement bornée quand E varie. Donc :

Dans le cas régulier, la condition nécessaire et suffisante pour que la valeur moyenne $\mathfrak{M} X^{(n)}(E)$ converge, uniformément quand E varie sur V , vers une limite finie indépendante de E est que $Y(E)$ (ou plus généralement $\mathfrak{M} X^{(n)}(E)$ pour au moins une valeur de n) soit bornée quand E varie sur V .

Si $Y(E)$ n'est pas borné, un cas simple où la condition de convergence est remplie est celui où $Y(E)$ est sommable sur V et où, de surcroît, $P^{(n)}(E, F)$ est, pour n assez grand, borné quand E, F varient arbitrairement sur V . (C'est par exemple ce qui a lieu quand $Y(E)$ étant sommable sur un domaine V borné, l'un au moins des $P^{(n)}(E, F)$ est continu pour E, F arbitraires sur V).

En effet

$$|\mathfrak{M} X^{(n)}(E) - \int_V Y(F) P(F) dF| \leq \int_{V'} [P^{(n)}(F) - P(F)] |Y(F)| dF.$$

Or si $P^{(n)}(E, F)$ a une borne supérieure finie μ quand E, F varient sur V , on a vu (p. 207) que $P^{(n)}(F) - P(F)$ est majoré pour n assez grand, par le terme correspondant u_n d'une progression géométrique indépendante de F . Le second membre sera donc $\leq u_n \int_V |Y(F)| dF$ et par suite tendra vers zéro avec $\frac{1}{n}$. On a donc encore

$$(22) \quad \lim_{n \rightarrow \infty} \mathfrak{M} X^{(n)}(E) = \int_V Y(F) P(F) dF$$

et la convergence est non seulement uniforme quand E varie, mais au moins aussi rapide que celle des termes d'une progression géométrique indépendante de E . Il est vrai que dans ce cas $\mathfrak{M} X^{(n)}(E)$ étant alors borné pour n assez grand quand E varie, on rentre dans un cas déjà examiné; mais il peut être parfois plus commode de s'assurer que $P^{(n)}(E, F)$ est borné que de le vérifier pour $\mathfrak{M} X^{(n)}(E)$.

Même dans certains cas où la démonstration précédente ne s'applique pas, il peut arriver qu'on ait encore l'égalité (22). Par exemple, dans le cas envisagé p. 189 et avec les mêmes notations, on posera

$$\mathfrak{M} X^{(n)}(x) = \int_{-\infty}^{+\infty} Y(y) P^{(n)}(x, y) dy$$

$$\text{avec } P^{(n)}(x, y) = \frac{1}{\sqrt{n\pi}} e^{-\frac{(x-y)^2}{n}} < \frac{1}{\sqrt{n\pi}}.$$

D'où

$$|\mathfrak{M} X^{(n)}(x)| \leq \frac{1}{\sqrt{n\pi}} \int_{-\infty}^{+\infty} |X(y)| dy.$$

Si donc $Y(y)$ est sommable sur le domaine V actuel, c'est-à-dire sur l'intervalle $(-\infty, +\infty)$, $\mathfrak{M} X^{(n)}(x)$ tend vers une limite qui est zéro. Pourtant, on n'est pas dans le cas régulier, puisque $P^{(n)}(x, y)$ tend vers zéro. Et cependant, on a encore

$$\lim_{n \rightarrow \infty} \mathfrak{M} X^{(n)}(x) = \int_{-\infty}^{+\infty} Y(y) P(y) dy.$$

Remarque. Dans ce qui précède nous avons cherché à déterminer des cas assez généraux — au reste tous réguliers — où l'on peut passer à la limite sous le signe $\int$ dans $\int_V Y(F) P^{(n)}(E, F) dF$.

Il peut cependant se produire des cas où cette intégrale ait une limite déterminée, sans que cette limite soit égale à $\int_V Y(F) P(F) dF$. C'est ce que nous allons constater dans l'exemple suivant (qui relève du cas quasi-singulier proprement dit).

Exemple. Généralisons un exemple dû à Lord Rayleigh. Considérons une molécule qui effectue une suite de déplacements sur une droite et supposons que la longueur de chaque déplacement ait pour un sens donné une probabilité donnée. Ou, plus précisément que la probabilité pour que la molécule partant de x arrive au bout d'un seul déplacement à une position comprise entre y et $y + dy$ soit le produit de dy par une fonction de la grandeur $|y - x|$ du déplacement et de son sens, c'est-à-dire une fonction de $y - x$, $p(y - x)$. Dans le cas examiné par Lord Rayleigh, des déplacements égaux en valeur absolue et de sens contraire avaient la même probabilité, la généralisation la plus immédiate consisterait donc à supposer que $p(z)$ est une fonction paire de z . Mais nous allons considérer aussi le cas général.

Si nous prenons pour $Y(y)$ la fonction y on aura $\mathfrak{M} X^{(n)}(x) = \int_{-\infty}^{+\infty} y P^{(n)}(y - x) dy$.

Mais, sans rien changer à la question de l'existence d'une limite du premier membre, on a profité à en retrancher une certaine quantité indépendante de n , qui lui donnera une signification concrète plus simple et d'un plus grand intérêt physique. Posons en effet

$$(23) \quad V^{(n)} = \int_{-\infty}^{+\infty} (y - x) P^{(n)}(y - x) dy = \mathfrak{M} X^{(n)}(x) - x \int_{-\infty}^{+\infty} P^{(n)}(y - x) dy \\ = [\mathfrak{M} X^{(n)}(x)] - x$$

$V^{(n)}$ est la valeur moyenne du déplacement résultant $y - x$ après la $n^{\text{ème}}$ opération à partir de l'abscisse x . On a

$$V = V^{(1)} = \int_{-\infty}^{+\infty} (y - x) p(y - x) dy = \int_{-\infty}^{+\infty} u p(u) du; \text{ valeur indépendante de la}$$

position initiale x mais qui n'a pas nécessairement une valeur déterminée et finie. Supposons $p(u)$ tel que cette intégrale soit finie ainsi que celles qui interviennent par la suite. V est « le premier moment » de $p(u)$. De même

$$V^{(n)} = \int_{-\infty}^{+\infty} (y - x) P^{(n)}(x, y) dy.$$

Or:

$$P^{(2)}(x, y) = \int_{-\infty}^{+\infty} p(z - x) p(y - z) dz = \int_{-\infty}^{+\infty} p(u) p(y - x - u) du.$$

En posant :

$$p^{(2)}(t) = \int_{-\infty}^{+\infty} p(u) p(t-u) du$$

on aura

$$P^{(2)}(x, y) = p^{(2)}(y-x).$$

De même, en posant

$$p^{(n)}(t) = \int_{-\infty}^{+\infty} p^{(n-1)}(u) p(t-u) du,$$

si l'on a pour $m < n$, $P^{(m)}(x, y) = p^{(m)}(y-x)$, alors

$$P^{(n)}(x, y) = \int_{-\infty}^{+\infty} P^{(n-1)}(z-x) p(y-z) dz = \int_{-\infty}^{+\infty} p^{(n-1)}(u) p(y-x-u) du$$

et l'on a

$$P^{(n)}(x, y) = p^{(n)}(y-x).$$

Donc :

$$\begin{aligned} V^{(n)} &= \int_{-\infty}^{+\infty} (y-x) p^{(n)}(y-x) dy = \int_{-\infty}^{+\infty} u p^{(n)}(u) du \\ &= \int u \left[\int p^{(n-1)}(t) p(u-t) dt \right] du = \int p^{(n-1)}(t) \left[\int u p(u-t) du \right] dt \\ &= \int p^{(n-1)}(t) \left[\int (t+v) p(v) dv \right] dt = \int p^{(n-1)}(t) [t + V^{(1)}] dt = V^{(n-1)} + V^{(1)}. \end{aligned}$$

Ainsi $V^{(1)} = V^{(n)} - V^{(n-1)} = V^{(n-1)} - V^{(n-2)} = \dots = V^{(2)} - V^{(1)}.$

D'où en ajoutant :

$$(24) \quad V^{(n)} = nV.$$

Si $p(u)$ est une fonction paire, $V = \int_{-\infty}^{+\infty} u p(u) du$ sera nul. Dans le cas contraire, V sera en général $\neq 0$. Les deux cas seront très différents :

I. Si $V = 0$, $V^{(n)}$ sera nul quel que soit n et, naturellement, tendra vers zéro avec $\frac{1}{n}$.

II. Si $V \neq 0$, $V^{(n)}$ croîtra indéfiniment en valeur absolue et gardera le même signe.

On peut dire qu'en moyenne au cours d'une suite infinie d'épreuves : ou bien la molécule oscillera sur place — c'est le cas de Lord Rayleigh — ou bien elle s'éloignera indéfiniment.

D'ailleurs, on a: d'après (23) et (24)

$$\mathfrak{M} X^{(n)}(x) = x + nV.$$

De sorte que si $V = 0$, $\mathfrak{M} X^{(n)}(x)$ reste égal à x , quand n croît et si $V \neq 0$, $\mathfrak{M} X^{(n)}(x)$ tend vers l'infini.

Observons que, dans le cas actuel, la condition (T') étant vérifiée et le domaine des états possibles étant de mesure infinie, $P^{(n)}(x, y)$ ne peut tendre vers une limite indépendante de l'état initial que s'il tend vers zéro (p. 224). De sorte que, sauf le cas particulier où $V = 0$, on n'aura pas

$$\lim_{n \rightarrow \infty} \mathfrak{M} X^{(n)}(x) = \int_{-\infty}^{+\infty} Y(y) P(y) dy,$$

puisque le premier membre est infini.

Notons d'autre part ce résultat intéressant que nous avons pu obtenir l'expression de $V^{(n)}$ en fonction explicite de n sans avoir besoin de connaître la densité de probabilité pour une épreuve, c'est-à-dire indépendamment de la fonction $p(u)$.

Effet de la condition (T'). Dans le cas le plus régulier, celui où $P(F)$ est une constante sur V' , on a

$$P(F) = \frac{1}{\text{mes. } V}. \text{ La formule (22) devient } M = \lim_{n \rightarrow \infty} \mathfrak{M} X^{(n)}(E) = \frac{\int_V Y(F) dF}{\int_V dF}.$$

Ainsi, dans le cas le plus régulier, M est, comme dans le cas régulier, indépendant de E , mais, au lieu d'être simplement une valeur moyenne de $Y(F)$ au sens du Calcul des Probabilités, c'est-à-dire une moyenne pondérée, c'est une valeur moyenne de $Y(F)$ au sens du Calcul Intégral.

Nous savons d'ailleurs que si le cas le plus régulier ne peut se présenter que sur un domaine fini, la condition (T') peut avoir lieu sur un domaine illimité. Dans ce cas, plusieurs circonstances peuvent se produire dont nous avons rencontré des exemples plus haut. Par exemple, dans le cas de la p. 233, si $\int_{-\infty}^{+\infty} u p(u) du \neq 0$, $\mathfrak{M} X^{(n)}(x)$ tend vers l'infini,

pour $Y(y) \equiv y$. Si l'on prend $p(u) = \frac{1}{\sqrt{\pi}} e^{-u^2}$, alors, dans le cas, où $Y(y) \equiv y$,

$\mathfrak{M} X^{(n)}(x)$ tend vers une limite, mais cette limite n'est pas indépendante de l'état initial, tandis que si $Y(y)$ est sommable, non seulement $\mathfrak{M} X^{(n)}(x)$ tend vers une limite indépendante de l'état initial, mais cette limite étant nulle est indépendante de la fonction $Y(y)$ et l'on a

$$\lim_{n \rightarrow \infty} \int_{-\infty}^{+\infty} Y(y) P^{(n)}(x, y) dy = \int_{-\infty}^{+\infty} Y(y) [\lim_{n \rightarrow \infty} P^{(n)}(x, y)] dy$$

ce qui n'avait pas lieu dans les deux cas précédents.

M. Hostinsky a étudié également le résultat de la condition (T') dans l'effet d'un grand nombre de rotations aléatoires d'une roulette ou d'une sphère (I, p. 51).

Moyennes arithmétiques. Nous avons vu que, dans le cas régulier, si, par exemple, $Y(F)$ est bornée sur V , la quantité

$$\mathfrak{M} X^{(n)}(E) = \int_V P^{(n)}(E, F) Y(F) dF$$

tend vers la quantité

$$M = \int_V P(F) Y(F) dF.$$

On peut aussi donner une autre interprétation de M , en appelant $M^{(n)}(E)$ la valeur moyenne de la moyenne arithmétique

$$A^{(n)}(E) = \frac{X^{(1)}(E) + \dots + X^{(n)}(E)}{n}.$$

On voit en effet que $M^{(n)}(E)$ est la moyenne arithmétique de la somme des n premiers termes d'une suite de nombres $\mathfrak{M} X^{(i)}(E)$ qui tendent vers M . On a donc :

$$M = \lim_{n \rightarrow \infty} M^{(n)}(E)$$

et on voit que cette limite est indépendante de E .

L'existence de la limite est ici établie dans le cas régulier ; il est vraisemblable qu'elle subsiste dans des cas plus généraux.

Nous allons particulariser ce qui précède, au cas où $X^{(n)}(E)$ ne peut prendre que les valeurs 0 ou 1.

Fréquences moyennes. Appelons encore *répétition* d'un événement A au cours de n épreuves, le nombre R de fois que A se produit et fréquence f de cet événement le rapport $\frac{R}{n}$.

Appliquons au cas où l'événement A consiste en ce que le système matériel étudié se trouve dans l'un des états F appartenant à une por-

tion v du domaine V des états possibles, et ceci en partant de l'état initial E . Alors en employant des notations analogues à celles de la p. 215, on aura :

$$\mathfrak{M}_{R_v^{(n)}}(E) = \sum_{t=1}^{t=n} \omega_v^{(t)}(E) = \sum_{t=1}^{t=n} \int_v P^{(t)}(E, F) dF.$$

Et on aura pour la valeur moyenne de la fréquence $f_v^{(n)}(E)$ avec laquelle on passera au cours des n premières épreuves par un des états de v , à partir de l'état E

$$\mathfrak{M}_{f_v^{(n)}}(E) = \int_v \pi^{(n)}(E, F) dF$$

où $\pi^{(n)}(E, F) = \frac{1}{n} \sum_{t=1}^{t=n} P^{(t)}(E, F)$. Donc :

$$\mathfrak{M}_{f_v^{(n)}}(E) - \int_v p(F) dF = \frac{1}{n} \sum_{t=1}^{t=n} \int_v [P^{(t)}(E, F) - p(F)] dF = H_n - K_n,$$

où :

$$\begin{aligned} H_n &= \frac{1}{n} \sum_{t=1}^{t=n} \int_v [P^{(t)}(E, F) - p^{(t)}(F)] dF \\ &\leq \frac{1}{n} \sum_{t=1}^{t=n} \left[\int_v P^{(t)}(E, F) dF - \int_v p^{(t)}(F) dF \right] \\ &= \frac{1}{n} \sum_{t=1}^{t=n} \left[1 - \int_v p^{(t)}(F) dF \right] \end{aligned}$$

et

$$K_n = \frac{1}{n} \sum_{t=1}^{t=n} \int_v [p(F) - p^{(t)}(F)] dF \leq \frac{1}{n} \sum_{t=1}^{t=n} \left[\int_v p(F) dF - \int_v p^{(t)}(F) dF \right].$$

Dans tous les cas $\int_v p^{(t)}(F) dF$ tend vers $\int_v p(F) dF$ quand t croît, donc K_n tend vers 0 avec $\frac{1}{n}$. Dans le cas régulier, $\int_v p^{(t)}(F) dF$ tend vers 1 quand t croît, donc H_n tend aussi vers zéro.

Ainsi, dans le cas régulier, la fréquence moyenne $\mathfrak{M}_{f_v^{(n)}}(E)$ tend vers une limite quand n croît indéfiniment et cette limite est : $\int_v P(F) dF$ qui est indépendante de E . (Par analogie avec ce qui se passe dans le cas d'un nombre fini d'états possibles (IV, p. 18), on est porté à penser que la première partie de cet énoncé s'étend au cas singulier et que la seconde partie s'étend à certains cas singuliers — prévision confirmée dans la note V après rédaction du présent mémoire —.)

Dispersion

Nous jugerons des dispersions des variables aléatoires qui viennent d'être étudiées en calculant leurs écarts quadratiques moyens.

1° Dispersion d'une variable aléatoire. Pour mesurer la dispersion de $X^{(n)}(E)$, calculons d'abord son écart quadratique moyen $\mu^{(n)}(E)$. On a

$$(25) \quad [\mu^{(n)}(E)]^2 = \int_V P^{(n)}(E, F) [Y(F) - \mathfrak{M} X^{(n)}(E)]^2 dF \\ = \int_V Y^2(F) P^{(n)}(E, F) dF - \left[\int_V Y(F) P^{(n)}(E, F) dF \right]^2.$$

De la relation

$$[\mu^{(n)}(E)]^2 = \mathfrak{M} [X^{(n)}(E)]^2 - [\mathfrak{M} X^{(n)}(E)]^2$$

et du résultat établi plus haut p. 231, et appliqué à $Y(F)$ et $Y^2(F)$, on déduit, que: dans le cas régulier, la condition nécessaire et suffisante pour que $\mathfrak{M} X^{(n)}(E)$ et $\mu^{(n)}(E)$ convergent uniformément, quand E varie sur V , vers des limites indépendantes de l'état initial E , est que $Y(F)$ soit borné sur E , ou que, pour une valeur assez grande de n , $\mathfrak{M} X^{(n)}(E)$ et $\mathfrak{M} [X^{(n)}(E)]^2$ soient bornés quand E varie sur V .

Un cas où $Y(E)$ n'est pas borné et où l'on peut cependant, s'assurer, sans itération, que cette double condition est satisfaite est celui où pour au moins une valeur de n , $P^{(n)}(E, F)$ a une borne supérieure finie quand E, F varient sur V et où $Y^2(F)$ est sommable sur V . (On sait qu'alors $Y(F)$ est aussi sommable sur V).

Il est clair que dans ces deux cas, la limite μ de l'écart quadratique moyen $\mu^{(n)}(E)$ est elle-même un écart quadratique moyen, à savoir celui qu'aurait $Y(F)$ si $P^{(n)}(E, F)$ était remplacé par $P(F)$. Car on a

$$\mu^2 = \int_V P(F) [Y(F) - M]^2 dF = \int_V P(F) Y^2(F) dF - \left[\int_V P(F) Y(F) dF \right]^2.$$

Pour simplifier, nous allons pour la suite nous contenter de conditions suffisantes; nous supposons, qu'on est dans le cas régulier et que $Y(F)$ est borné sur V (ou bien que $Y^2(F)$ est sommable sur V et que $P^{(n)}(E, F)$ est pour n assez grand, borné supérieurement quand E, F varient sur V).

On déduit de $\mu^{(n)}(E)$ la valeur de l'écart quadratique moyen $\lambda^{(n)}(E)$ de $X^{(n)}(E)$ avec M . Car

$$[\lambda^{(n)}(E)]^2 = [\mu^{(n)}(E)]^2 + [\mathfrak{M} X^{(n)}(E) - M]^2.$$

Par suite, quand n croît indéfiniment, $\lambda^{(n)}(E)$ qui reste $\geq \mu^{(n)}(E)$, tend vers la même limite μ indépendante de E .

Dispersion de la moyenne arithmétique. De même, appelons $\varrho^{(n)}(E)$ et $\delta^{(n)}(E)$, les écarts quadratiques moyens de $A^{(n)}(E)$ avec sa moyenne $M^{(n)}(E)$ et avec la limite M de cette moyenne. On a⁶)

$$n [\delta^{(n)}(E)]^2 = \frac{1}{n} \sum_{u=1}^{u=n} [\lambda^{(u)}(E)]^2 + L^{(n)}.$$

Le premier terme du second membre tend vers μ^2 quand n croît indéfiniment. Reste à étudier

$$\begin{aligned} L^{(n)} &\equiv \frac{2}{n} \sum_{u < v \leq n} \mathfrak{M} [X^{(u)}(E) - M] [X^{(v)}(E) - M] \\ &= 2 \int_V [X(F) - M] \left[\int_V [X(G) - M] R^{(n)}(F, G, E) dG \right] dF \end{aligned}$$

$$\begin{aligned} \text{avec } R^{(n)}(F, G, E) &= \frac{1}{n} \sum_{u < v \leq n} P^{(u)}(E, F) P^{(v-n)}(F, G) \\ &= \frac{1}{n} \left\{ P^{(1)}(E, F) P^{(1)}(F, G) + [P^{(1)}(E, F) P^{(2)}(F, G) + P^{(2)}(E, F) P^{(1)}(F, G)] + \dots \right. \\ &\quad \left. \dots + [P^{(1)}(E, F) P^{(n-1)}(F, G) + \dots + P^{(n-1)}(E, F) P^{(1)}(F, G)] \right\} \\ &= [S^{(n)}(F, G, E) + T^{(n)}(F, G, E)] \end{aligned}$$

où $S^{(n)}$ et $T^{(n)}$ s'obtiennent en remplaçant dans $R^{(n)}$ les $P^{(k)}(F, G)$ respectivement par $P^{(k)}(F, G) - P(G)$ et par $P(G)$. On pourra dans $L^{(n)}$ remplacer $R^{(n)}$ par $S^{(n)}$, car le terme négligé dans $L^{(n)}$ sera

$$\begin{aligned} &\sum_{u < v \leq n} \frac{2}{n} \int_V [X(F) - M] \int_V [X(G) - M] P^{(u)}(E, F) P(G) dF dG \\ &= \sum_{u < v \leq n} \frac{2}{n} \int_V [X(F) - M] P^{(u)}(E, F) dF \cdot \int_V [X(G) - M] P(G) dG \end{aligned}$$

où la dernière intégrale est nulle.

⁶) Pour un calcul et des raisonnements analogues à ceux qui suivent, mais un peu plus détaillés, voir IV, p. 23.

Or $S^{(n)}$ est le produit par $\frac{n-1}{n}$ de la moyenne arithmétique des $n-1$ premiers termes de la série $\alpha_1 + \alpha_2 + \dots + \alpha_n + \dots$ où

$$\alpha_n = P^{(1)}(E, F) \omega^{(n)}(F, G) + \dots + P^{(n-1)}(E, F) \omega^{(1)}(F, G)$$

avec $\omega^{(n)}(F, G) \equiv P^{(n)}(F, G) - P(G)$. Si on démontre que α_n tend vers α , il sera prouvé que $S^{(n)}$ tend vers α .

Supposons encore, pour simplifier, que non seulement l'un des $P^{(n)}(E, F)$ mais $p(E, F)$ ait une borne supérieure finie quand E, F varient indépendamment sur V . Alors, dans le cas régulier, $P^{(n)}(E, F)$ tend uniformément vers $P(F)$ et la série $\sum_n |\omega^{(n)}(F, G)|$ est bornée et uniformément convergente. Dès lors α_n tend uniformément vers

$$S(F, G) \equiv P(F) \sum_{n=1}^{n=+\infty} \omega^{(n)}(F, G) = P(F) s(F, G).$$

Par suite

$$(26) \quad \lim_{n \rightarrow \infty} n [\delta^{(n)}(E)]^2 = \mu'^2$$

avec

$$\mu'^2 = \mu^2 + L$$

où

$$L =$$

$$2 \sum_{n=1}^{n=\infty} \int_V [X(F) - M] P(F) \left\{ \int [X(G) - M] [P^{(n)}(F, G) - P(G)] dG \right\} dF.$$

On peut aussi écrire

$$(27) \quad \mu'^2 = \int_V P(F) [X(F) - M]^2 dF + 2 \int_V [X(F) - M] \left\{ \int [X(G) - M] S(F, G) dG \right\} dF.$$

D'autre part, on a évidemment

$$(28) \quad n [\delta^{(n)}(E)]^2 = n [\rho^{(n)}(E)]^2 + n [M^{(n)}(E) - M]^2$$

Et

$$\begin{aligned} \sqrt{n} |M^{(n)}(E) - M| &= \sqrt{n} \left| \sum_{u=1}^{u=n} \frac{1}{n} \int [P^{(u)}(E, F) - P(F)] X(F) dF \right| \\ &\leq \frac{1}{\sqrt{n}} \sum_{u=1}^{u=n} A q^u \int |X(F)| dF \end{aligned}$$

puisque, dans le cas actuel, les $|P^{(u)}(E, F) - P(F)|$ sont inférieurs aux termes d'une série convergente $\sum_u A q^u$

Finalement on a :

$$(29) \quad \lim_{n \rightarrow \infty} n [\delta^{(n)}(E)]^2 = \mu'^2.$$

Ainsi, lorsque n croît indéfiniment, non seulement $\delta^{(n)}(E)$ tend vers zéro, mais nous connaissons maintenant sa partie principale $\frac{\mu'}{\sqrt{n}}$.

Retour à l'exemple. Il n'est pas impossible d'étudier le comportement de la dispersion en dehors du cas régulier. Reprenons l'exemple de Lord Rayleigh et les notations de la p. 233, pour calculer la dispersion du déplacement $y - x$. Soit $\theta_x^{(n)}$ l'écart quadratique moyen du déplacement à partir de l'abscisse x après n chocs. On aura

$$[\theta_x^{(n)}]^2 = \int (y - x - V^{(n)})^2 P^{(n)}(x, y) dy = \int_{-\infty}^{+\infty} [u - V^{(n)}]^2 p^{(n)}(u) du.$$

C'est une quantité, $[\theta^{(n)}]^2$, indépendante de la position initiale x .

On a :

$$[\theta^{(n)}]^2 = \int_{-\infty}^{+\infty} u^2 p^{(n)}(u) du - [V^{(n)}]^2 = \alpha_n - n^2 V^2.$$

Or

$$\begin{aligned} \int u^2 p^{(n)}(u) du &= \int u^2 \left[\int p^{(n-1)}(t) p(u-t) dt \right] du = \\ &= \int p^{(n-1)}(t) \left[\int u^2 p(u-t) du \right] dt \end{aligned}$$

et

$$\begin{aligned} \int u^2 p(u-t) du &= \int (t+v)^2 p(v) dv = t^2 + 2t \int v p(v) dv + \int v^2 p(v) dv \\ &= t^2 + 2tV + W \quad \text{avec } W = \int_{-\infty}^{+\infty} v^2 p(v) dv. \end{aligned}$$

$\theta^{(n)}$ sera fini si $\int_{-\infty}^{+\infty} v p(v) dv$ et $\int_{-\infty}^{+\infty} v^2 p(v) dv$ sont finis ce que nous allons supposer.

Alors

$$\alpha_n = \int p^{(n-1)}(t) [t^2 + 2 t V + W] dt = \alpha_{n-1} + 2 (n-1) V^2 + W$$

d'où

$$[\theta^{(n)}]^2 - [\theta^{(n-1)}]^2 = (\alpha_n - \alpha_{n-1}) - [n^2 - (n-1)^2] V^2 = W - V^2, \text{ avec } [\theta^{(1)}]^2 = W - V^2. \text{ Donc}$$

$$[\theta^{(n)}]^2 = n [\theta^{(1)}]^2,$$

$$\theta^{(n)} = \theta^{(1)} \sqrt{n} \quad \text{et} \quad \theta^{(1)} = \sqrt{W - V^2}.$$

Sous la condition que le premier et le deuxième moments, V et W , de $p(u)$ soient finis, on a généralisé le résultat établi par Lord Rayleigh dans le cas de déplacement finis, constants et égaux, à savoir que $\theta^{(n)}$ est proportionnel à $\sqrt{n}$.

Observons que si l'on prend pour $Y(y)$, la fonction y , on a $y - \mathfrak{M} X^{(n)}(x) = y - x - V^{(n)}$ et par suite l'écart quadratique moyen $\mu^{(n)}$ de $X^{(n)}(x)$ est aussi égal à $\theta^{(n)}$. On a donc ici un exemple d'un cas quasi-régulier mais non régulier. Et ici $\mu^{(n)}$ au lieu de tendre vers une limite, croît indéfiniment.

II° Dispersion des fréquences. Il y a plusieurs fréquences à distinguer. La fréquence $F_v^{(n,N)}(E)$ avec laquelle dans N groupes de n épreuves, on aboutit à la $n^{\text{ème}}$ épreuve à un état appartenant au domaine v après être parti initialement de l'état E , a pour valeur moyenne la probabilité $\omega_v^{(n)}(E) = \int_v P^{(n)}(E, F) dF$ de cet événement. Et son écart quadratique

moyen est, comme on sait, la quantité $\mathcal{F}_v^{(n,N)}(E) = \sqrt{\frac{\omega_v^{(n)}(E)[1 - \omega_v^{(n)}(E)]}{N}}$.

On voit qu'on aura dans le cas régulier

$$\lim_{n \rightarrow \infty} \mathfrak{M} F_v^{(n,N)}(E) = \int_v P(F) dF$$

$$\lim_{n \rightarrow \infty} \mathcal{F}_v^{(n,N)}(E) = \sqrt{\frac{[\int_v P(F) dF] [\int_{V-v} P(F) dF]}{N}}.$$

La première limite ne dépend ni de n , ni de N , ni de l'état initial E ; la seconde ne dépend ni de n , ni de E .

Passons maintenant à l'écart quadratique moyen $\mathcal{E}_v^{(n)}(E)$ de la fréquence $f_v^{(n)}(E)$. Les limites des valeurs moyennes de $F_v^{(n,N)}(E)$ et de $f_v^{(n)}(E)$ (p. 237) sont les mêmes, quel que soit N , lorsque n croît indéfiniment; il n'en est pas de même de leur dispersion.

On aura avantage à calculer d'abord l'écart quadratique moyen $\mathcal{G}_v^{(n)}(E)$ de $f_v^{(n)}(E)$ avec la limite $p_v = \int_v P(F) dF$ de la valeur moyenne de $f_v^{(n)}(E)$. On aura évidemment

$$(30) \quad [\mathcal{G}_v^{(n)}(E)]^2 = [\mathcal{E}_v^{(n)}(E)]^2 + [\mathfrak{M} f_v^{(n)}(E) - p_v]^2.$$

On connaît déjà le dernier terme

$$\left[\frac{1}{n} \sum_{t=1}^{t=n} \int_v P^{(t)}(E, F) dF - p_v \right]^2 = \frac{1}{n^2} \left\{ \sum_{t=1}^{t=n} \int_v [P^{(t)}(E, F) - P(F)] dF \right\}^2$$

et on sait que si, pour simplifier, on suppose que $p(E, F)$ a une borne supérieure finie quand E, F varient sur V , l'accolade tend, quand n croît, vers une limite finie $\int_v s(E, F) dF$.

En ce qui concerne $\mathcal{E}_v^{(n)}$ on pourrait le calculer directement. Mais, il est clair que si $X(F) = 1$ ou 0 suivant que F se trouve ou non sur v , $A_E^{(n)}$, $M^{(n)}(E)$, M , $\delta^{(n)}(E)$ se réduiront à $f_v^{(n)}(E)$, $\mathfrak{M} f_v^{(n)}(E)$, p_v , $\mathcal{G}_v^{(n)}$. Alors d'après les formules (26), (27), on a

$$\lim_{n \rightarrow \infty} n [\mathcal{G}_v^{(n)}(E)]^2 = K^2 + H$$

avec :

$$K^2 = \int_v P(F) dF - \left[\int_v P(F) dF \right]^2, \quad \text{ou} \quad K^2 = p_v (1 - p_v)$$

et

$$\begin{aligned} H &= 2 \int_v (1 - p_v) \int_v (1 - p_v) S(F, G) dG dF + 2 \int_{V-v} (-p_v) \int_v (1 - p_v) S(F, G) dG dF \\ &\quad + 2 \int_v (1 - p_v) \int_{V-v} (-p_v) S(F, G) dG dF + 2 \int_{V-v} (-p_v) \int_{V-v} (-p_v) S(F, G) dG dF \\ &= 2 (1 - p_v)^2 \int_v \int_v S(F, G) dG dF + 2 (p_v)^2 \int_{V-v} \int_{V-v} S(F, G) dG dF \\ &\quad - 2 p_v (1 - p_v) \left[\int_v \int_{V-v} S(F, G) dG dF + \int_{V-v} \int_v S(F, G) dG dF \right]. \end{aligned}$$

D'ailleurs, on a, avec convergence uniforme

$$\begin{aligned} \int_V S(F, G) dG &= \sum_{t=1}^{t=+\infty} \int_V P(F) [P^{(t)}(F, G) - P(G)] dG \\ &= \sum_{t=1}^{t=+\infty} P(F) \left[\int_V P^{(t)}(F, G) dG - \int_V P(G) dG \right] = \sum_{t=1}^{t=+\infty} P(F) [1-1] = 0, \end{aligned}$$

et d'après $(\mathcal{F}_m)$

$$\begin{aligned} \int_V S(F, G) dF &= \sum_{t=1}^{t=\infty} \left[\int_V P(F) P^{(t)}(F, G) dF - P(G) \int_V P(F) dF \right] \\ &= \sum_{t=1}^{t=\infty} [P(G) - P(G)] = 0. \end{aligned}$$

$$\text{Dès lors } \int_{V-v} S(F, G) dG = - \int_v S(F, G) dG, \quad \int_{V-v} S(F, G) dF = - \int_v S(F, G) dF$$

et H se réduit à :

$$\begin{aligned} &2(1-p_v)^2 \int_v \left[\int_v S(F, G) dG \right] dF - 2(p_v)^2 \int_{V-v} \left[\int_v S(F, G) dG \right] dF \\ &- 2p_v(1-p_v) \left\{ - \int_v \left[\int_v S(F, G) dG \right] dF + \int_{V-v} \left[\int_v S(F, G) dG \right] dF \right\} \\ &= 2 \left\{ (1-p_v) \int_v \left[\int_v S(F, G) dG \right] dF - p_v \int_{V-v} \left[\int_v S(F, G) dG \right] dF \right\} \\ &= 2 \left\{ \int_v \int_v S(F, G) dG dF - p_v \int_v \left[\int_v S(F, G) dF + \int_{V-v} S(F, G) dF \right] dG \right\} \\ &= 2 \int_v \left[\int_v S(F, G) dF - p_v \int_V S(F, G) dF \right] dG \\ &= 2 \int_v \int_v S(F, G) dG dF. \end{aligned}$$

$$\text{Car } \int_v \left[\int_V S(F, G) dF \right] dG = \int_v \left[\int_v S(F, G) dG \right] dF = 0.$$

$$\text{Ainsi } H = 2 \int_v \int_v S(F, G) dG dF.$$

Il résulte alors de la formule (26), qu'on aura, non seulement $\mathcal{G}_v^{(n)}(E) \supseteq \mathcal{E}_v^{(n)}(E)$, mais encore

$$\begin{aligned} \lim_{n \rightarrow \infty} \sqrt{n} \mathcal{G}_v^{(n)}(E) &= \lim_{n \rightarrow \infty} \sqrt{n} \mathcal{E}_v^{(n)}(E) = \sqrt{K^2 + H} \\ &= \sqrt{p_v(1-p_v) + 2 \int_v \int_v P(F) s(F, G) dG dF}. \end{aligned}$$

(Reçu le 23 avril 1932)

Références bibliographiques

- I° *B. Hostinsky*: Méthodes générales du calcul des probabilités.
66 pages, Gauthier-Villars, 1931. Voir aussi:
Ibis le texte des conférences de M. Hostinsky en cours d'impression dans les Annales de l'Institut Henri Poincaré, 1932, vol. III, pp. 1—74.
- II° *E. Goursat*: Cours d'analyse.
T. III, Gauthier-Villars, 2^{me} éd., 1915.
- III° *Ch. J. de la Vallée-Poussin*: Cours d'analyse infinitésimale.
T. II, Gauthier-Villars, 2^{me} éd., 1912.
- IV° *M. Fréchet*: Compléments à la théorie des probabilités discontinues «en chaîne». Annali R. Scuole N. S. Pisa, 1932, pp. 1—30.
- V° *M. Fréchet*: Sur le comportement de certains noyaux de Fredholm itérés et sur les probabilités «en chaîne».
C. R., t. 195, 1932, p. 590—592.
- VI° *M. Fréchet*: On the behaviour when n becomes infinite of the n -th iterate of a Fredholm's kernel. Proceedings of the Amer. Nat. Acad., Washington, 1932.
- VII° *M. Fréchet*: Sur la solution continue la plus générale d'une équation fonctionnelle de la théorie des probabilités «en chaîne».
C. R., t. 195, 1932, p. 639—641.
M. Fréchet: Solution continue la plus générale d'une équation fonctionnelle de la théorie des probabilités.
Bull. Soc. Math. France, t. LX, 1932.

Grundlagen der elliptischen Geometrie

Von ANNA FISCHER, Bern.

In seiner Arbeit über die Grundlagen der Geometrie nimmt Hilbert¹⁾ die folgende Definition der Ebene an:

„Die Ebene ist ein System von Dingen, die Punkte heißen und die sich umkehrbar eindeutig auf die im Endlichen gelegenen Punkte der Zahlenebene oder auf ein gewisses Teilsystem derselben abbilden lassen; diese Punkte der Zahlenebene (d. h. die Bildpunkte) werden auch zugleich zur Bezeichnung der Punkte unserer Ebene selbst verwandt.

Zu jedem Punkte A unserer Ebene gibt es in der Zahlenebene Jordansche Gebiete, in denen der Bildpunkt von A liegt und deren sämtliche Punkte ebenfalls Punkte unserer Ebene darstellen. Diese Jordanschen Gebiete heißen Umgebungen des Punktes A .

Jedes in einer Umgebung von A enthaltene Jordansche Gebiet, innerhalb dessen der Punkt A (Bildpunkt von A) liegt, ist wiederum eine Umgebung von A .

Ist B irgendein Punkt in einer Umgebung von A , so ist diese Umgebung auch zugleich eine Umgebung von B .

Wenn A und B irgend zwei Punkte unserer Ebene sind, so gibt es stets eine Umgebung von A , die zugleich den Punkt B enthält.“

Die Bewegung wird wie folgt definiert:

„Eine Bewegung ist eine umkehrbar eindeutige stetige Transformation der Bildpunkte der Zahlenebene in sich von der Art, daß dabei der Umlaufssinn einer geschlossenen Jordanschen Kurve stets derselbe bleibt. Die Umkehrung der zu einer Bewegung gehörenden Transformation ist wieder eine Bewegung. Eine Bewegung, bei welcher ein Punkt M ungeändert bleibt, heißt eine Drehung um den Punkt M .“

Um nun die Euklidische oder hyperbolische Geometrie aufzubauen, genügen drei Axiome:

Axiom I: Die Bewegungen bilden eine Gruppe.

Bezeichnen wir als wahren Kreis die Gesamtheit der verschiedenen Lagen, die ein Punkt A bei den sämtlichen Drehungen um einen von ihm verschiedenen Punkt M annehmen kann, so lautet das

Axiom II: Jeder wahre Kreis besteht aus unendlich vielen Punkten.

Für die Gruppe der Bewegungen muß noch die Stetigkeit gefordert werden; dies geschieht in

¹⁾ Hilbert, Über die Grundlagen der Geometrie. Math. Ann. 56, 1902.

Axiom III: Die Bewegungen bilden ein abgeschlossenes System.

Auf Grund dieser Axiome gelingt es Hilbert, den Kreis und die Gerade mit ihren notwendigen Eigenschaften zu gewinnen.

Zu den Untersuchungen Hilbert's bemerkt Gonseth²⁾, daß es vorteilhafter wäre, anstatt in der Zahlenebene direkt in der *topologisch* definierten euklidischen oder hyperbolischen Ebene zu operieren. In der vorliegenden Arbeit wird dieser Weg eingeschlagen, und zwar zur Begründung der elliptischen Geometrie.

Als *Ebene* wählen wir die *projektive Ebene* in ihrer topologischen Bedeutung, d. h. die geschlossene einseitige Fläche vom Geschlecht eins.

Eine *Drehung* um den Punkt M ist eine eindeutige stetige Transformation der projektiven Ebene in sich, wobei der Punkt M ungeändert bleibt; der Umlaufssinn jeder geschlossenen Jordanschen Kurve, die M im Inneren enthält, soll dabei invariant sein.

Die drei Axiome nehmen wir in der folgenden Fassung an:

Axiom I: Die Drehungen um einen beliebigen Punkt M bilden eine Gruppe.

Axiom II: Jeder Kreis besteht aus unendlich vielen Punkten.

Axiom III: Die Drehungen um einen beliebigen Punkt M bilden ein abgeschlossenes System.

§ 1. Zunächst grenzen wir in der projektiven Ebene durch eine zweifurige doppelpunktfreie geschlossene Kurve C ein Gebiet ab. Darin gelingt es, genau nach dem Verfahren von Hilbert, den Kreis K um M zu konstruieren und zu beweisen, daß er eine geschlossene Jordansche Kurve ist. Die Einführung der Koordinaten x, y für die Punkte auf K , die Hilbert benützt, kann umgangen werden, da es sich beim Beweis nur um die Aufeinanderfolge der Punkte auf K handelt. Anstatt der Zahlengeraden und der Zahlenkreise betrachten wir Jordansche Bögen und geschlossene Jordansche Kurven.

Die Gruppe der Drehungen von K um M in sich erweist sich als holomorph mit der Gruppe der gewöhnlichen Drehungen des gewöhnlichen Kreises in sich. •

Erst von hier an ist mein Aufbauverfahren von demjenigen von Hilbert wesentlich verschieden.

§ 2. Es muß nun die Gruppe der Transformationen aller Punkte bei den Drehungen der projektiven Ebene um M untersucht werden.

²⁾ *Gonseth*, Les Fondements des Mathématiques. Paris, A. Blanchard. 1926, p. 66 ff.

Wir beweisen zunächst:

Ist μ ein Kreis um M , von dem bekannt ist, daß er eine geschlossene Jordansche Kurve ist, die M im Inneren enthält, so gibt es außer der Identität keine Drehung um M , die jeden Punkt von μ in Ruhe läßt.

Beweis. Die Drehung um M , die jeden Punkt auf μ festläßt, sei D . Wir nehmen im Gegensatz zu unserer Behauptung *erstens* an, es gäbe in beliebiger Nähe eines Punktes A auf μ Punkte, die durch D transformiert werden. Dann schlagen wir um A einen genügend kleinen Kreis, der durch einen gegenüber D veränderlichen Punkt geht. Dieser Kreis α trifft μ sicher in einem Punkt B . Die Rotation D um M ist für den Kreis α eine Transformation, bei der der Mittelpunkt A und ein Punkt B auf α invariant bleiben; es ist also eine Transformation des Kreises α in sich. Dabei kann entweder die Orientierung von α unverändert bleiben und wir haben es mit einer Drehung von α in sich zu tun; oder geht die Orientierung von α in die entgegengesetzte über und unsere Transformation ist eine Umlegung des Kreises α in sich.

Betrachten wir zunächst den letzteren Fall. Da B der Schnittpunkt von α mit μ ist, gibt es auf α eine unendliche Reihe von Punkten C_i , die gegen B konvergieren und außerhalb μ liegen, und eine unendliche Reihe von Punkten D_i , die gegen B konvergieren und innerhalb μ liegen. Bei der Drehung D um M bleibt der Punkt B fest und α soll eine Umlegung erleiden. Es müßte dabei die Reihe der Punkte C_i in die Reihe der Punkte D_i übergehen. Angenommen, dies wäre nicht der Fall und es würde die gegen B konvergierende Punktfolge C_i in eine gegen einen Punkt E auf α konvergierende Folge E_i übergehen. Wir können auf μ eine unendliche Folge von Punkten F_i wählen, die auf μ gegen den Punkt B konvergiert. Bei der Drehung D um M würde, nach Annahme, die Folge F_i in Ruhe bleiben, während die C_i in die E_i übergehen. Es müßte daher, nach Axiom III, eine Drehung um M geben, bei welcher der Punkt B fest bleiben und gleichzeitig in den Punkt E übergehen würde. Dies ist unmöglich und folglich muß bei der Umlegung des Kreises α die Folge C_i in die Folge D_i übergehen. Die Transformation der C_i in die D_i ist aber ihrerseits auch nicht möglich, da bei einer Drehung um M Punkte außerhalb μ nicht ins Innere von μ gelangen können. Unsere Annahme, daß die Transformation des Kreises α eine Umlegung sei, ist also falsch; die Drehung D um M ist für den Kreis α ebenfalls eine Drehung um seinen Mittelpunkt A . Dabei bleibt, nach Annahme, der Punkt B von α fest. Es müssen daher alle Punkte von α fest bleiben und es kann in beliebiger Nähe von A keine Punkte geben, die durch die Drehung D transformiert würden.

Wir konstruieren nun um M ein System von geschlossenen Kurven ohne Doppelpunkt, so daß durch jeden Punkt der projektiven Ebene eine und nur eine Kurve des Systems hindurchgeht.

Zunächst beweisen wir, daß sich in diesem System eine einufrige Kurve befinden muß. Angenommen, es seien alle Kurven des Systems zweiufrige. Da die projektive Ebene eine geschlossene einseitige Fläche ist, gibt es durch jeden Punkt mindestens eine einufrige geschlossene Kurve. Wir ziehen eine solche, γ durch M . Durch jeden Punkt von γ geht eine Kurve unseres Systems. Für jede zweiufrige Kurve des Systems wird ein Teil der Punkte von γ innerhalb, ein Teil auf und eventuell ein Teil außerhalb der Kurve sein. Geht man im System immer weiter, so muß man schließlich zu einer Kurve K kommen, die so ist, daß alle Punkte von γ innerhalb und auf K liegen. Es wäre somit möglich, eine zweiufrige Kurve K im System zu finden, die γ vollständig enthält. Dies ist aber unmöglich. Also gibt es in unserem System mindestens eine einufrige geschlossene Kurve. Es gibt eine einzige, da zwei einufrige Kurven sich in der projektiven Ebene (Geschlecht = 1) notwendig treffen müßten. Diese einufrige Kurve liegt außerhalb aller zweiufrigen des Systems um M .

Es sei nun λ eine zweiufrige Kurve des Systems, innerhalb oder außerhalb μ . Dann nehmen wir *zweitens* an, daß alle Punkte im Ring zwischen μ und λ bei der Drehung D fest bleiben, während in beliebiger Nähe von λ bewegliche Punkte sind. Wir legen um den Punkt B auf λ einen genügend kleinen Kreis α , so daß er durch das Gebiet zwischen μ und λ hindurchgeht. Dieser Kreis geht auch durch einen beweglichen Punkt A in der Nähe von . Die Drehung D ist für α eine Drehung in sich. Dabei bleiben unendlich viele Punkte von α fest; somit müssen alle Punkte von α fest bleiben.

Ist λ die einufrige Kurve des Systems, so können wir *drittens* annehmen, daß alle Punkte innerhalb und außerhalb μ bei der Drehung D fest bleiben, mit Ausnahme der Punkte auf λ . Dann ziehen wir um einen der festen Punkte einen genügend kleinen Kreis durch einen beweglichen Punkt auf λ und schließen gleich wie früher.

Damit ist unser Beweis erbracht: *Bei den Drehungen D um M , die alle Punkte von μ fest lassen, bleiben alle Punkte der Ebene fest.*

§ 3. Wir beweisen nun folgende Behauptung: *Jeder Kreis ist eine geschlossene Jordansche Kurve. Das System aller Kreise um M ist so beschaffen, daß durch jeden Punkt der projektiven Ebene ein und nur ein Kreis des Systems hindurchgeht. Im System befindet sich stets eine einzige einufrige Kurve, außerhalb aller Kreise. Diese einufrige Kurve*

nennen wir Gerade. Man erhält jeden Punkt des Kreises durch den Punkt P je einmal, wenn man den Drehungswinkel ω die Werte von 0 bis 2π durchlaufen läßt. Jeder Punkt der Geraden wird erhalten, wenn ω die Werte von 0 bis π durchläuft. Für die Zusammensetzung zweier Drehungen um die Winkel ω_1, ω_2 gilt die Formel:

$$\Delta(\omega_1) \Delta(\omega_2) = \Delta(\omega_1 + \omega_2).$$

Wir betrachten die Drehungen des Kreises K in sich um die Winkel von 0 bis 2π . Es ist zu beweisen, daß dabei jeder Punkt der projektiven Ebene ebenfalls einen Kreis (oder die Gerade) um M beschreibt. Mit jedem Punkt der projektiven Ebene betrachten wir eine kleine Umgebung desselben, die mit einer bestimmten Indikatrix versehen ist. Bei einer Drehung um M kann ein Punkt in sich selber transformiert werden, entweder unter Erhaltung oder unter Umkehrung der zugehörigen Indikatrix. Diese zwei Fälle sind verschieden. Unter dieser Voraussetzung gilt³⁾:

1. Für jeden Punkt der projektiven Ebene hängt die Lage des transformierten Punktes vom Drehwinkel ω *eindeutig* ab.

Jeder Drehung um den Winkel ω zwischen 0 und 2π entspricht eine Drehung von K in sich. Da bei einer Drehung um 2π , bei welcher die Punkte von K in Ruhe bleiben, auch *alle* Punkte der projektiven Ebene in Ruhe bleiben, entspricht jeder Drehung von K in sich eine einzige Drehung der Ebene um M . Nach einer Drehung um 2π erreicht jeder Punkt seine ursprüngliche Lage mit der ursprünglichen Indikatrix.

2. Die Lage des transformierten Punktes hängt *stetig* vom Drehwinkel ω ab.

Wir wählen auf K einen festen Punkt O . Es sei $\omega_1, \omega_2, \dots$ eine unendliche Reihe von Winkelwerten, die gegen den Wert ω konvergieren, $T_1, T_2, \dots$ eine unendliche Reihe von Punkten der projektiven Ebene, die gegen den Punkt T konvergieren. Bei den Drehungen um $\omega_1, \omega_2, \dots$ gehe der Punkt O auf K bezüglich in die Punkte $S_1, S_2, \dots$ über. Aus den Punkten $T_1, T_2, \dots$ entstehen bei den Drehungen um $\omega_1, \omega_2, \dots$ die Punkte $Z_1, Z_2, \dots$. Die Drehung um ω transformiert O in S und T in Z . Wir haben zu zeigen, daß die Z_i gegen Z konvergieren.

Angenommen, der Häufungspunkt der Z_i sei Z^* . Da die S_i auf K gegen S konvergieren, gibt es nach Axiom III eine Drehung um M , die O in S und T in Z^* transformiert. Bei der Drehung, die O nach S bringt, geht aber, nach Annahme, T in Z über. Es ist also $Z^* = Z$, der Häufungspunkt der Z_i .

³⁾ Der Beweis für die Kreise ist demjenigen von Hilbert nachgebildet.

Die Punkte der projektiven Ebene werden bei den Drehungen um M eindeutig und stetig transformiert. Die geschlossenen, doppelpunktfreien zweifurigen Kurven, die dabei entstehen, sind Kreise um M . Wir haben noch zu zeigen, daß alle Punkte eines Kreises erst erhalten werden, wenn der Winkel ω alle Werte von 0 bis 2π durchläuft. Angenommen, ein Kreis würde vollständig erhalten bei allen Drehungen von 0 bis ω^* ($\omega^* < 2\pi$). Führen wir die Drehung um ω^* aus, so blieben alle Punkte dieses Kreises in Ruhe; es müßten somit alle Punkte der Ebene in Ruhe bleiben. Die Punkte von K bleiben aber nur bei einer Drehung um 2π in Ruhe. Es muß also $\omega^* = 2\pi$ sein.

Eine der entstandenen geschlossenen, stetigen doppelpunktfreien Kurven ist sicher einufig; es ist die Gerade I' „um M “. (M soll auch „Mittelpunkt der Geraden I' “ heißen). Wir haben zu beweisen, daß jeder Punkt von I' erhalten wird, wenn der Drehwinkel ω alle Werte von 0 bis π durchläuft.

Wir betrachten einen Punkt P von I' und einen genügend kleinen orientierten Kreis α um ihn. Alle Punkte von α liegen entweder auf den Kreisen um M oder auf I' . Da I' einufig ist, muß es eine Transformation geben, die P in sich überführt, wobei α mit entgegengesetzter Indikatrix mit sich selbst zusammenfällt. Diese Transformation ist eine Drehung um M um einen Winkel $\omega' < 2\pi$. Dabei bleiben außer P noch zwei Punkte, A und B auf α fest und es bleibt eine ganze Achse von Punkten zwischen A , P und B fest. Diese Fixpunkte können nicht auf den Kreisen um M liegen; sie liegen auf I' . Bei der Drehung um ω' bleiben unendlich viele Punkte von I' in Ruhe; es bleibt also die ganze Gerade I' in Ruhe. α erleidet eine Umlegung. Da die Aufeinanderfolge von zwei Umlegungen die Identität ergibt, ist $\omega' = \pi$.

Damit ist unsere Behauptung in allen Teilen bewiesen.

Aus unserer Betrachtung folgt weiter:

Bleibt bei einer Drehung um M ein Punkt der Geraden I' in Ruhe, so bleiben alle Punkte von I' in Ruhe.

Definition: Jedes Punktsystem, das durch beliebige Drehung der Geraden I' um irgendeinen ihrer Punkte entsteht, ist wiederum eine Gerade.

§ 4. *Zwei Gerade in der projektiven Ebene haben stets einen und nur einen Schnittpunkt.*

Die Existenz von mindestens einem Schnittpunkt folgt aus dem Geschlecht der projektiven Ebene. Es ist zu zeigen, daß es nur einen Schnittpunkt gibt. Angenommen, zwei Gerade, I'_1 und I'_2 hätten zwei

Schnittpunkte A und B . Die „Mittelpunkte“ von I_1 und I_2 seien bezüglich M_1 und M_2 . Es sind nun zwei Fälle möglich:

a) Um A läßt sich ein orientierter Kreis durch B legen. Bei einer Drehung um π um M_1 erleidet dieser Kreis eine Umlegung an der Achse I_1 . Bei einer weiteren Drehung um π um M_2 erleidet er eine weitere Umlegung an der Achse I_2 . Nach Ausführung der beiden Drehungen ist der Kreis mit ursprünglicher Orientierung in sich übergegangen. Dabei blieb der Punkt B fest. Die Transformation des Kreises in sich ist somit die Identität. Es müßten alle Punkte der Ebene in Ruhe bleiben, was bei unserer Annahme nicht der Fall sein kann.

b) Um A läßt sich nur eine Gerade g durch B legen. Führen wir nun eine Drehung um π um M_1 und M_2 aus, so bleiben A und B fest *). Diese Drehung ist also für g eine Drehung in sich, bei welcher ein Punkt von g in Ruhe bleibt. Es müßte also die ganze Gerade g in Ruhe bleiben, was unmöglich ist.

Daraus folgt: *Zwei Punkte der projektiven Ebene lassen sich stets durch eine einzige Gerade verbinden.*

Man legt durch den einen Punkt eine Gerade und dreht sie um diesen Punkt, bis sie durch den zweiten Punkt geht.

Jeder Punkt der projektiven Ebene läßt sich in jeden andern Punkt durch eine Drehung um einen Winkel $< \pi$ transformieren.

Die beiden Punkte seien A und B . Wir legen durch sie eine Gerade, die einen „Mittelpunkt“ M besitzt. Dann gibt es sicher eine Drehung um M um einen Winkel $< \pi$, die A in B überführt.

§ 5. Wir sind nun in der Lage, den Begriff der Strecke zu definieren.

Definition: Die durch zwei Punkte A und B bestimmte Strecke AB ist der Bogen der Geraden zwischen A und B , der einem Drehwinkel $\leq \frac{\pi}{2}$ um den „Mittelpunkt“ dieser Geraden entspricht. Die Strecke wird durch diesen Drehwinkel gemessen.

Zwei Strecken AB und $A'B'$ sind *kongruent*, wenn die zugehörigen Drehwinkel $\left(\leq \frac{\pi}{2}\right)$ um die Mittelpunkte M und M' der Geraden durch A, B bzw. A', B' einander gleich sind.

Zwei kongruente Strecken AB und $A'B'$ lassen sich stets durch zwei Drehungen (durch eine *Bewegung*) zur Deckung bringen. Man lege zuerst durch A und A' die Gerade g und bringe A durch Drehung um den „Mittelpunkt“ von g nach A' . Darauf läßt sich der Punkt B durch Drehung um A' in B' transformieren.

*) Mit fester Indikatrix.

Zwei Kreise sind kongruent, wenn es eine Drehung gibt, die ihre Mittelpunkte und gleichzeitig sie selbst zur Deckung bringt.

§ 6. Jeder um einen Punkt auf einer Geraden gelegte Kreis trifft diese Gerade in zwei Punkten.

Bei der Drehung um π um den „Mittelpunkt“ der Geraden erleidet der Kreis eine Umlegung. Dabei bleiben auf ihm zwei und nur zwei Punkte fest.

§ 7. Wir müssen noch feststellen, daß in unserer Geometrie die Kongruenzaxiome gelten. Um den Schnittpunkt S von zwei Geraden Γ_1 und Γ_2 legen wir einen Kreis K . Jede der beiden Geraden trifft K in zwei Punkten, A und A' bzw. B und B' , und die Gerade Γ um S in je einem Punkt, P_1 bzw. P_2 . Denkt man sich aus den Geraden Γ_1 und Γ_2 den Punkt P_1 bzw. P_2 entfernt, so bilden die Strecken SAP_1 , $SA'P_1$, SBP_2 , $SB'P_2$ vier „Strahlen“. Der Winkel ASB z. B. wird nun wie folgt definiert: S heißt *Scheitel*, die Strahlen SAP_1 und SBP_2 *Schenkel* des Winkels. Der Winkel selbst wird gemessen durch den Drehungswinkel ω um S , der den Punkt A in B überführt.

Dann gilt in unserer Geometrie der erste Kongruenzsatz für Dreiecke in der folgenden Fassung:

Wenn für zwei Dreiecke ABC und $A'B'C'$ die Kongruenzen

$$AB \equiv A'B', \quad AC \equiv A'C', \quad \sphericalangle BAC \equiv \sphericalangle B'A'C'$$

gelten, so gelten stets auch die Kongruenzen

$$\sphericalangle ABC \equiv \sphericalangle A'B'C', \quad \sphericalangle ACB \equiv \sphericalangle A'C'B', \quad BC \equiv B'C'.$$

Der Beweis dieses Satzes ergibt sich mit Leichtigkeit aus unseren Definitionen der Drehung, der Bewegung und der Kongruenz.

(Eingegangen den 4. April 1932)

Note sur un type d'équations différentielles du premier ordre

par W. RIVIER, Lausanne

Soit $F(u, v, w)$ une fonction de trois variables indépendantes. Je suppose que cette fonction possède des dérivées partielles du premier ordre et que ces dernières satisfont à la condition suivante :

$$(1) \quad au \frac{\partial F}{\partial u} + bv \frac{\partial F}{\partial v} + cw \frac{\partial F}{\partial w} = k F,$$

a, b, c et k désignant des constantes données. Je considère alors l'équation différentielle

$$(2) \quad F\left(x, y, \frac{dy}{dx}\right) = 0.$$

Cette équation présente quelques particularités simples que je me propose d'indiquer ici.

1^o) Opérons dans (2) les substitutions $x = x'^{\mu}$, $y = y'^{\nu}$, où μ et ν désignent des constantes arbitraires différentes de zéro. Soit alors

$$F'\left(x', y', \frac{dy'}{dx'}\right) = 0$$

l'équation transformée de (2). Envisageons la fonction $F'(u, v, w)$. Cette fonction satisfera à la condition

$$a'u \frac{\partial F'}{\partial u} + b'v \frac{\partial F'}{\partial v} + c'w \frac{\partial F'}{\partial w} = k F'$$

pour $a' = \frac{a}{\mu}$, $b' = \frac{b}{\nu}$ et c' déterminé par la relation

$$a' - b' + c' = a - b + c.$$

2°) Supposons a et b différents de zéro. Choisissons μ et ν de manière que l'on ait

$$(3) \quad \frac{a}{\mu} - \frac{b}{\nu} = a - b + c.$$

$\frac{dy'}{dx'}$ sera alors fonction du produit $x'^{-\frac{b\mu}{a\nu}} y'$ seulement. Si l'on a $a - b + c \neq 0$, l'exposant qui affecte x' dans ce produit pourra prendre n'importe quelle valeur donnée à l'avance, la valeur -1 exceptée. Si l'on a $a - b + c = 0$, cet exposant, au contraire, prendra la valeur -1 quelles que soient μ et ν satisfaisant à (3).

Dans le premier cas, il n'existe pas de méthode générale connue pour intégrer (2).

Dans le second cas, l'équation transformée de (2) sera une équation homogène. Il en résulte une méthode d'intégration qui fera dépendre l'intégrale générale de (2) d'une fonction implicite d'une variable. Plus exactement, cette intégrale générale sera définie par les relations

$$\begin{cases} x = \left(C e^{\int \frac{dt}{z-t}} \right)^{\frac{a\nu}{b}} \\ y = \left(C t e^{\int \frac{dt}{z-t}} \right)^{\nu}, \end{cases}$$

où C désigne la constante arbitraire, z et t des variables liées par l'équation

$$(4) \quad F\left(1, t^{\nu}, \frac{b}{a} t^{\nu-1} z\right) = 0,$$

et ν une quantité dont on pourra disposer arbitrairement. Supposons, par exemple, que $F(u, v, w)$ soit un trinôme. Si l'on prend alors, dans l'équation (4), t comme fonction et le produit $t^{\lambda} z$ comme variable, l'équation qui liera t à cette dernière pourra, par un choix convenable des quantités λ et ν , être rendue linéaire par rapport à t .

3°) Les remarques que nous venons de faire, concernant le cas où l'on a $a - b + c = 0$, s'appliquent naturellement aux équations différentielles homogènes du premier ordre, puisque ces équations différentielles, comme on s'en sera aperçu tout de suite, appartiennent au type envisagé dans cette note et sont caractérisées par le fait que, dans la condition (1) à laquelle elles satisfont, on a $a = b$ et $c = 0$.

(Reçu le 30 juillet 1932)

Laplace'sche Integraltransformation und Integration partieller Differentialgleichungen vom hyperbolischen und parabolischen Typus

(Ein Beitrag zum Heaviside'schen Operatorenkalkül)

Von W. MÄCHLER, Zürich

Im Jahre 1917 hat *T. J. I. A. Bromwich*⁴⁾ eine Arbeit veröffentlicht betitelt: „Normal coordinates in dynamical systems“. Bromwich gibt hier eine interessante Methode an zur Lösung von Anfangswert-Randwertproblemen für partielle Differentialgleichungen vom hyperbolischen und parabolischen Typus, für die aber bisher kein Beweis gegeben wurde. Die vorliegende Abhandlung hat nun den Zweck, im Anschluß an eine Arbeit von *M. Plancherel*⁵⁾, die Methode von Bromwich einer genauen Analyse zu unterziehen.

Die Bromwich'sche Methode hängt mit dem Operatorenkalkül von *O. Heaviside*⁷⁾ eng zusammen, so daß die von uns erhaltenen Resultate den ersten strengen Beweis der Anwendbarkeit dieses Kalküls für eine große Klasse von Problemen mit zwei unabhängigen Variablen und nicht konstanten Koeffizienten^{5) 6)} liefern *).

§ 1. Formulierung des Problems, der Voraussetzungen und Angabe der Resultate

Wir betrachten die folgenden zwei Probleme:

Problem I:

$$I_1) \quad a(x) \frac{\partial^2 u}{\partial t^2} + b(x) \frac{\partial u}{\partial t} - L(u) = f(x, t), \quad 0 \leq x \leq 1, \quad 0 < t < \infty,$$

$$I_2) \quad \begin{cases} u(x, 0) = u_0(x), \\ \left. \frac{\partial u(x, t)}{\partial t} \right|_{t=0} = u_1(x), \end{cases}$$

$$I_3) \quad \begin{cases} L_1(u) \equiv a_{11} u(0, t) + a_{12} u'(0, t) + b_{11} u(1, t) + b_{12} u'(1, t) = 0, \\ L_2(u) \equiv a_{21} u(0, t) + a_{22} u'(0, t) + b_{21} u(1, t) + b_{22} u'(1, t) = 0. \end{cases}$$

*) *G. Doetsch*^{5) 6)} untersucht spezielle, einfachere Fälle des vorliegenden Problems mit konstanten Koeffizienten und andern Randbedingungen.

Die Größen a_{ik} , b_{ik} ($i, k = 1, 2$) sind Konstanten und die Striche bedeuten Ableitungen nach x . Ferner ist:

$$(1) \quad L(u) \equiv p(x) \frac{\partial^2 u}{\partial x^2} + q(x) \frac{\partial u}{\partial x} + r(x) u.$$

Problem II:

$$\begin{aligned} \Pi_1) \quad & (a(x)\lambda^2 + b(x)\lambda) \mathcal{U} - L(\mathcal{U}) = (a(x)\lambda + b(x))u_0(x) + a(x)u_1(x) \\ & + g(x, \lambda), \\ \Pi_2) \quad & \begin{cases} L_1(\mathcal{U}) = 0, \\ L_2(\mathcal{U}) = 0. \end{cases} \end{aligned}$$

Dabei ist λ ein komplexer Parameter und $g(x, \lambda)$, $f(x, t)$ sind durch die Formeln (Laplace'sche Integraltransformation)

$$(2) \quad g(x, \lambda) = \int_0^\infty e^{-\lambda t} f(x, t) dt,$$

$$(3) \quad f(x, t) = \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} e^{\lambda t} g(x, \lambda) d\lambda, *$$

miteinander verknüpft.

Voraussetzungen: Es sei $0 \leq x \leq 1$; dann fordern wir:

$$(4) \quad \begin{cases} a(x), p(x) \text{ zweimal stetig differentiierbar,} \\ b(x), q(x) \text{ einmal stetig differentiierbar,} \\ r(x) \text{ stetig,} \end{cases}$$

und

$$(5) \quad \begin{cases} a(x) > 0, ** \\ p(x) > 0. \end{cases}$$

*) Ueber die Definition dieses Integrals verweisen wir auf die Bemerkungen in der Einleitung des § 8.

**) Wäre $a(x) \equiv 0$, $b(x) > 0$ so ließe sich dieser Fall analog behandeln, indem man im Problem II $\lambda = \mu^2$ setzen würde.

Ferner,

$$(6) \quad \begin{cases} u_0(x) \text{ dreimal stetig differenzierbar,} \\ u_1(x) \text{ zweimal stetig differenzierbar,} \\ \text{und } u_0(x), u_1(x) \text{ sollen die Randbedingungen } I_3 \text{ erfüllen.} \end{cases}$$

Von den Funktionen

$$(7) \quad f(x, t), \quad \frac{\partial f}{\partial t}, \quad \frac{\partial^2 f}{\partial t^2}, \quad \frac{\partial f}{\partial x},$$

fordern wir die Stetigkeit in x, t für $0 \leq x \leq 1, t \geq 0$ und die Integrale

$$(8) \quad \int_0^\infty |f(x, t)| dt, \quad \int_0^\infty \left| \frac{\partial f}{\partial t} \right| dt, \quad \int_0^\infty \left| \frac{\partial^2 f}{\partial t^2} \right| dt$$

sollen gleichmäßig konvergent in x sein für $0 \leq x \leq 1^*$.

Für die Koeffizienten in den Randbedingungen soll gelten:

$$(9) \quad \begin{vmatrix} a_{12} & b_{12} \\ a_{22} & b_{22} \end{vmatrix} \neq 0.$$

Wie weit man sich von dieser Bedingung befreien kann ist in der vorliegenden Arbeit gezeigt (§ 5). Die Rechnungen werden aber nur durchgeführt wenn (9) besteht.

Unter diesen Voraussetzungen läßt sich beweisen, daß das Problem I lösbar ist und das Problem II, mit Ausnahme von abzählbar unendlich vielen Werten von λ , ebenfalls lösbar ist, wobei zwischen den Lösungen dieser Probleme die Beziehungen bestehen

$$(10) \quad u(x, t) = \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} e^{\lambda t} \mathcal{V}(x, \lambda) d\lambda, \quad \mathcal{V}(x, \lambda) = \int_0^\infty e^{-\lambda t} u(x, t) dt.$$

Es ist $\sigma > 0$ und größer als die obere Schranke der Realteile der Eigenwerte des Problems II. (§ 8). Die Funktion $u(x, t)$ ist einmal stetig nach x und t differenzierbar und erfüllt bis auf eine zweidimensionale Punktmenge vom Maß Null die Differentialgleichung I_1). Mit Hilfe des Residuensatzes läßt sich für das Funktionenpaar $u_0(x), u_1(x)$ ein Entwicklungssatz herleiten. (Satz VIII, § 14).

*) Sind die Bedingungen (8) nicht erfüllt, so läßt sich die Lösung $u(x, t)$ in jedem Intervall $0 \leq t \leq T$ noch durch (10) ausdrücken, wenn man für $t > T, f(x, t) = 0$ setzt und $g(x, \lambda), \mathcal{V}(x, \lambda)$ entsprechend berechnet.

**) Man beachte die Bemerkungen zu Beginn des § 8.

§ 2. Lösung der homogenen Differentialgleichung II₁)

Unter der homogenen Differentialgleichung II₁) verstehen wir die folgende Gleichung:

$$(a(x)\lambda^2 + b(x)\lambda) v - L(v) = 0.$$

Ueber diese gilt ein grundlegender Satz, der von *G. D. Birkhoff*¹⁾ und *J. Tamarkin*^{12) 13)} herrührt. Eine Nachprüfung des Birkhoff'schen Beweises gestattet es, diesen Satz in folgender Form auszusprechen:

Satz I: Sind die Bedingungen 4), 5) erfüllt und setzt man

$$(11) \quad \varphi_1(x) = -\varphi_2(x) = \sqrt{\frac{a(x)}{p(x)}},$$

bedeuten ferner l, L endliche positive Zahlen, wobei L genügend groß ist, so gibt es zwei linear unabhängige Lösungen $v_1(x, \lambda)$, $v_2(x, \lambda)$ der homogenen Differentialgleichungen II₁), die transzendente Funktionen von λ sind und die ferner im Gebiet*) $|\lambda| > L$, $\Re(\lambda) \leq l$ für $0 \leq x \leq 1$ folgende asymptotische Darstellung haben:

$$(12) \quad v_k(x, \lambda) = e^{\lambda \int_0^x \varphi_k(x) dx} \left\{ \eta_k(x) + \frac{E_{k0}(x, \lambda)}{\lambda} \right\} \quad (k = 1, 2)$$

$$\frac{\partial v_k(x, \lambda)}{\partial x} = e^{\lambda \int_0^x \varphi_k(x) dx} \lambda \varphi_k(x) \left\{ \eta_k(x) + \frac{E_{k1}(x, \lambda)}{\lambda} \right\};$$

dabei ist $\eta_k(x)$ zweimal stetig differentiierbar in $0 \leq x \leq 1$ und es ist

$$(13) \quad \eta_k(x) = \frac{1}{\sqrt{|\varphi_k(x)|}} e^{-\frac{1}{2} \int_0^x \frac{q(x)\varphi_k(x) - b(x)}{p(x)\varphi_k(x)} dx} \quad (k = 1, 2).$$

*) $\Re(\lambda)$ bedeutet den Realteil von λ .

Ebenso existiert ein zweites derartiges Fundamentalsystem $v_1^*(x, \lambda)$, $v_2^*(x, \lambda)$ mit folgender asymptotischen Darstellung im Gebiet $|\lambda| > L$, $\Re(\lambda) \geq -l$ und für $0 \leq x \leq 1$

$$(14) \quad v_k^*(x, \lambda) = e^{\lambda \int_0^x \varphi_k(x) dx} \left\{ \eta_k(x) + \frac{E_{ko}^*(x, \lambda)}{\lambda} \right\} \quad (k = 1, 2)$$

$$\frac{\partial v_k^*(x, \lambda)}{\partial x} = e^{\lambda \int_0^x \varphi_k(x) dx} \lambda \varphi_k(x) \left\{ \eta_k(x) + \frac{E_{k1}^*(x, \lambda)}{\lambda} \right\}.$$

Für die Funktionen $E_{k\nu}(x, \lambda)$ bzw. $E_{k\nu}^*(x, \lambda)$, ($k = 1, 2$), ($\nu = 0, 1$), gelten in $0 \leq x \leq 1$ die Ungleichungen:

$$(15) \quad \begin{cases} |E_{k\nu}(x, \lambda)| < M \text{ im Gebiet } |\lambda| > L, \Re(\lambda) \leq l, \\ |E_{k\nu}^*(x, \lambda)| < M \text{ im Gebiet } |\lambda| > L, \Re(\lambda) \leq -l, \end{cases}$$

wobei M eine von x unabhängige Konstante ist.

§ 3. Die Green'sche Funktion des Randwertproblems II ²⁾ ³⁾ ¹²⁾ ¹⁴⁾

Bilden die Funktionen $v_k(x, \lambda)$, ($k = 1, 2$) irgend ein Fundamentalsystem der homogenen Differentialgleichung II₁) und setzen wir

$$(16) \quad \delta(s, \lambda) = \begin{vmatrix} v_1'(s, \lambda) & v_2'(s, \lambda) \\ v_1(s, \lambda) & v_2(s, \lambda) \end{vmatrix} = \delta(0, \lambda) e^{-\int_0^s \frac{q(s)}{p(s)} ds}$$

so ist

$$(17) \quad g(x, s, \lambda) = \pm \frac{1}{2p(s)\delta(s, \lambda)} \begin{vmatrix} v_1(x, \lambda) & v_2(x, \lambda) \\ v_1(s, \lambda) & v_2(s, \lambda) \end{vmatrix} \quad \begin{array}{l} + \text{ für } x \geq s \\ - \text{ für } x \leq s \end{array}$$

eine Gröndlösung der homogenen Differentialgleichung II₁).

Wir benützen ferner die Abkürzungen

$$(18) \quad \Delta(x, s; \lambda) = \begin{vmatrix} v_1(x, \lambda) & v_2(x, \lambda) & g(x, s; \lambda) \\ L_1(v_1) & L_1(v_2) & L_1(g)_x \\ L_2(v_1) & L_2(v_2) & L_2(g)_x \end{vmatrix},$$

wobei

$$\begin{aligned} L_k(g)_x = & a_{k1} g(0, s; \lambda) + a_{k2} \frac{\partial g(x, s; \lambda)}{\partial x} \Big|_{x=0} \\ & + b_{k1} g(1, s; \lambda) + b_{k2} \frac{\partial g(x, s; \lambda)}{\partial x} \Big|_{x=1} \quad (k = 1, 2) \end{aligned}$$

ist, und

$$(19) \quad \Delta(\lambda) = \begin{vmatrix} L_1(v_1) & L_1(v_2) \\ L_2(v_1) & L_2(v_2) \end{vmatrix}.$$

Damit ist die Green'sche Funktion des Randwertproblems II gegeben durch

$$(20) \quad G(x, s; \lambda) = \frac{\Delta(x, s; \lambda)}{\Delta(\lambda)}$$

Da wir $v_k(x, \lambda)$, ($k = 1, 2$) als ganze transzendente Funktionen von λ voraussetzen können, so folgt, daß $\delta(s, \lambda)$, $\Delta(\lambda)$ ganze transzendente Funktionen sind und $g(x, s; \lambda)$, $\Delta(x, s; \lambda)$, $G(x, s; \lambda)$, im allgemeinen meromorphe Funktionen von λ .

Die Funktionen $g(x, s; \lambda)$, $G(x, s; \lambda)$ sind zweimal stetig nach x bzw. s differenzierbar für $0 \leq x, s \leq 1$ sobald $x \neq s$ ist. Bei $x = s$ hat die erste Ableitung beider Funktionen eine Unstetigkeit, gegeben durch

$$(21_1) \quad \frac{\partial g(x, s; \lambda)}{\partial x} \Big|_{x=s+0} - \frac{\partial g(x, s; \lambda)}{\partial x} \Big|_{x=s-0} = \frac{1}{p(s)}$$

und

$$(21_2) \quad \frac{\partial G(x, s; \lambda)}{\partial x} \Big|_{x=s+0} - \frac{\partial G(x, s; \lambda)}{\partial x} \Big|_{x=s-0} = \frac{1}{p(s)}.$$

§ 4. Existenz von Nullstellen von $\Delta(\lambda)$ für grosse λ

Wegen der Invarianz der Green'schen Funktion bezüglich der linearen Transformationen der Lösungen eines Fundamentalsystems, wobei die Koeffizienten unabhängig von x sind und eine nicht verschwindende Determinante haben, können wir, je nach Bedarf, das Fundamentalsystem 12) oder 14) benützen.

Setzt man

$$(22) \quad \int_0^1 \varphi_k(x) dx = \gamma_k \quad (k = 1, 2)$$

so folgt unter Verwendung von 12)

$$(23) \quad L_i(v_k) = \lambda \left\{ a_{i2} \varphi_k(0) \eta_k(0) + \frac{a_{i1} \left(\eta_k(0) + \frac{E_{k0}(0, \lambda)}{\lambda} \right) + a_{i2} \varphi_k(0) E_{k1}(0, \lambda)}{\lambda} \right. \\ \left. + e^{\lambda \gamma_k} \left(b_{i2} \varphi_k(1) \eta_k(1) + \frac{b_{i1} \left(\eta_k(1) + \frac{E_{k0}(1, \lambda)}{\lambda} \right) + b_{i2} \varphi_k(1) E_{k1}(1, \lambda)}{\lambda} \right) \right\}$$

Benützen wir die Abkürzungen

$$(24) \quad A_{ik} = a_{i2} \varphi_k(0) \eta_k(0), \quad B_{ik} = b_{i2} \varphi_k(1) \eta_k(1)$$

und verstehen wir, nach einer Bezeichnung von Birkhoff¹⁾, unter $[\alpha]$ den Ausdruck

$$(25) \quad [\alpha] = \alpha + \frac{E}{\lambda},$$

wo E eine von λ und eventuell auch anderen Variablen abhängige GröÙe bedeutet, die für große $|\lambda|$ im Gebiet $\Re(\lambda) \leq l$ oder im Gebiet $\Re(\lambda) \geq -l$ gleichmäßig beschränkt ist in diesen Variablen, so ergibt sich, sobald $\Re(\lambda) \leq l$

$$(23_1) \quad L_i(v_k) = \lambda \{ [A_{ik}] + e^{\lambda \gamma_k} [B_{ik}] \}$$

und der Ausdruck 19) wird

$$\Delta(\lambda) = \lambda^2 \begin{vmatrix} [A_{11}] + e^{\lambda \gamma_1} [B_{11}], [A_{12}] + e^{\lambda \gamma_2} [B_{12}] \\ [A_{21}] + e^{\lambda \gamma_1} [B_{21}], [A_{22}] + e^{\lambda \gamma_2} [B_{22}] \end{vmatrix}.$$

Für zweireihige Determinanten verwenden wir die Abkürzung

$$\begin{vmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{vmatrix} = |m_{ik}|_2$$

woraus folgt:

$$(26) \quad \Delta(\lambda) = \lambda^2 | [A_{ik}] + e^{\lambda \gamma_k} [B_{ik}] |_2.$$

Setzt man

$$(27) \quad \Delta_1(\lambda) = | [A_{ik}] + e^{\lambda \gamma_k} [B_{ik}] |_2$$

so ergibt sich, wegen $\gamma_1 = -\gamma_2$ und 24)

$$(27_1) \quad \Delta_1(\lambda) = e^{\lambda \gamma_2} \{ [A_0] + [A_1] e^{\lambda \gamma_1} + [A_2] e^{2\lambda \gamma_1} \}$$

wobei A_0, A_1, A_2 durch

$$(28) \quad \begin{cases} A_0 = \begin{vmatrix} A_{11} & B_{12} \\ A_{21} & B_{22} \end{vmatrix} = \varphi_1(0) \eta_1(0) \varphi_2(1) \eta_2(1) \begin{vmatrix} a_{12} & b_{12} \\ a_{22} & b_{22} \end{vmatrix} \\ A_1 = 0 \\ A_2 = \begin{vmatrix} B_{11} & A_{12} \\ B_{21} & A_{22} \end{vmatrix} = -\varphi_1(1) \eta_1(1) \varphi_2(0) \eta_2(0) \begin{vmatrix} a_{12} & b_{12} \\ a_{22} & b_{22} \end{vmatrix} \end{cases}$$

definiert sind. Nach Voraussetzung (9) und wegen $\varphi_k(x) \neq 0, \eta_k(x) \neq 0$ in $0 \leq x \leq 1$, sind A_0 und A_2 von entgegengesetztem Zeichen. Aus

$$[A_\nu] = A_\nu + \frac{E_\nu(\lambda)}{\lambda} \quad (\nu = 1, 2)$$

und mit folgender Definition von $A(\lambda)$

$$(29) \quad A(\lambda) = A_0 + A_2 e^{2\lambda \gamma_1}$$

ergibt sich

$$e^{\lambda \gamma_1} \Delta_1(\lambda) = A(\lambda) + \frac{E_0(\lambda) + E_1(\lambda) e^{\lambda \gamma_1} + E_2(\lambda) e^{2\lambda \gamma_1}}{\lambda}$$

eine Beziehung, die für große $|\lambda|$ und $\Re(\lambda) \leq l$ besteht.

Die Funktion $A(\lambda)$ ist periodisch mit der Periode $\frac{i\pi}{\gamma_1}$ und hat die einfachen Wurzeln

$$(30) \quad \lambda_k = \frac{1}{2\gamma_1} \operatorname{Lg} \left(-\frac{A_0}{A_2} \right) + \frac{i k \pi}{\gamma_1} \quad (k = 0, \pm 1, \pm 2, \dots)$$

wo Lg den Hauptwert des Logarithmus bedeutet*). Wir finden nach 28) und 13)

$$(31) \quad -\frac{A_0}{A_2} = e^{-\int_0^1 \frac{b(x)}{\sqrt{a(x)p(x)}} dx};$$

also insbesondere wenn in $0 \leq x \leq 1$

$b(x) \equiv 0$ ist, wird $\Re(\lambda_k) = 0$

$b(x) > 0$ ist, wird $\Re(\lambda_k) < 0$

$b(x) < 0$ ist, wird $\Re(\lambda_k) > 0$.

Wir untersuchen jetzt das Verhalten von $A(\lambda)$ im Streifen**)

$$(32) \quad -\frac{\pi}{2\gamma_1} \leq \Im(\lambda) \leq \frac{\pi}{2\gamma_1}.$$

Setzen wir $\lambda = \xi + i\eta$, so ist

$$(33) \quad |A(\lambda)| > \frac{|A_0|}{2} \text{ sobald } |\xi| > h \text{ und}$$

passend gewähltes $h > 0$. In dem durch die Ungleichungen definierten Rechteck

$$(34) \quad \begin{aligned} -h &\leq \Re(\lambda) \leq h \\ -\frac{\pi}{2\gamma_1} &\leq \Im(\lambda) \leq \frac{\pi}{2\gamma_1} \end{aligned}$$

liegt nur eine Nullstelle von $A(\lambda)$, wobei h genügend groß gewählt sei. Umgibt man diese mit einem genügend kleinen Kreis vom Radius ϱ und bezeichnet m das Maximum von $|A(\lambda)|$ auf der Kreisperipherie, so ist

*) Das Argument sei zwischen $-\pi, \pi$ gewählt.

**) $\Im(\lambda)$ bedeutet den Imaginärteil von λ .

im Rechteck (34) und wegen (33) auch im ganzen durch (32) definierten Streifen

$$(35) \quad |A(\lambda)| \geq m.$$

Aus der Periodizität von $A(\lambda)$ folgt also, daß man die Nullstellen von $A(\lambda)$ mit Kreisen von festem genügend kleinem Radius ρ so umgeben kann, daß in der ganzen λ -Ebene, ausschließlich dieser Kreise, die Ungleichung (35) gilt, wobei m nur eine Funktion von ρ ist.

Als „gelochte λ -Ebene“ wollen wir die λ -Ebene ausschließlich dieser Kreise um die Nullstellen von $A(\lambda)$ bezeichnen.

In der gelochten λ -Ebene betrachten wir die Funktion

$$\frac{e^{\lambda\gamma_1} A(\lambda)}{\lambda^2} = e^{\lambda\gamma_1} \cdot A_1(\lambda) = A(\lambda) + \frac{E_0(\lambda) + E_1(\lambda) e^{\lambda\gamma_1} + E_2(\lambda) e^{2\lambda\gamma_1}}{\lambda}$$

für große $|\lambda|$. Nach (30) können wir eine Zahl $l_1 > 0$ so wählen, daß $|\Re(\lambda_k)| < l_1$ ist. Da $|E_v(\lambda)| < M$ für $|\lambda| > L$ und $\Re(\lambda) \leq l_1$, wo M eine Konstante ist, so folgt:

$$\left| \frac{E_0(\lambda) + E_1(\lambda) e^{\lambda\gamma_1} + E_2(\lambda) e^{2\lambda\gamma_1}}{\lambda A(\lambda)} \right| < \frac{M(1 + e^{l_1\gamma_1} + e^{2l_1\gamma_1})}{|\lambda| \cdot m}.$$

Dieser Ausdruck wird kleiner als Eins, sobald $|\lambda|$ genügend groß ist. Wendet man also den Satz von *Rouché* auf die Funktionen

$$A(\lambda) \quad \text{und} \quad \frac{e^{\lambda\gamma_1} \cdot A(\lambda)}{\lambda^2}$$

an, so sieht man, daß für $\Re(\lambda) \leq l_1$ und $|\lambda|$ genügend groß, die Funktion $A(\lambda)$ unendlich viele einfache Nullstellen besitzt die in den oben betrachteten Kreisen liegen.

Analog ergibt sich, wenn man (14) gebraucht

$$\frac{e^{\lambda\gamma_2} A(\lambda)}{\lambda^2} \cdot A_0 e^{2\lambda\gamma_2} + A_2 + \frac{E_0^*(\lambda) + E_1^*(\lambda) e^{\lambda\gamma_1} + E_2^*(\lambda) e^{2\lambda\gamma_1}}{\lambda}$$

mit $|E_v^*(\lambda)| < M^*$ für $|\lambda| > L$ und $\Re(\lambda) \geq -l_1$. Da die Funktionen $A(\lambda)$ und $e^{2\lambda\gamma_2} A(\lambda) = A_0 e^{2\lambda\gamma_2} + A_2$ die gleichen Nullstellen haben, so ergibt sich der

Satz II: Unter der Voraussetzung (9) hat die Funktion $A(\lambda)$ unendlich viele Nullstellen die, abgesehen von endlich vielen, einfach sind und in den oben betrachteten Kreisen mit dem festen Radius ρ liegen.

§ 5. Bemerkungen zum § 4

Wir haben bei der Herleitung des Satzes II vorausgesetzt, daß die Bedingung (9) erfüllt sei. Wenn diese Determinante (9) gleich Null ist, so können wir den Satz II im allgemeinen doch aufrecht erhalten, eventuell mit dem Zusatz, daß die Determinante $\Delta(\lambda)$ für große $|\lambda|$ auch Doppelwurzeln haben kann. Im besonderen werden wir erkennen, daß der Satz II, unter Berücksichtigung der vorhin gemachten Bemerkung, für die meisten in der mathematischen Physik wichtigen Randbedingungen noch gilt.

Ist der Ausdruck (9) gleich Null, also

$$a_{12} b_{22} - a_{22} b_{12} = 0,$$

so sind folgende Fälle möglich:

- a) $a_{12} = a_{22} = b_{12} = b_{22} = 0$
- b)
$$\begin{cases} a_{12} = a_{22} = 0, b_{12} = 0, b_{22} \neq 0 \\ a_{12} = a_{22} = 0, b_{12} \neq 0, b_{22} = 0 \\ a_{12} = a_{22} = 0, b_{12} = b_{22} \neq 0^* \end{cases}$$
- c)
$$\begin{cases} a_{12} \neq 0, a_{22} = b_{12} = b_{22} = 0 \\ a_{12} = 0, a_{22} \neq 0, b_{12} = b_{22} = 0 \\ a_{12} = a_{22} \neq 0^*, b_{12} = b_{22} = 0 \end{cases}$$
- d)
$$\begin{cases} a_{12} = 0, a_{22} \neq 0, b_{12} = 0, b_{22} \neq 0 \\ a_{12} \neq 0, a_{22} = 0, b_{12} \neq 0, b_{22} = 0 \\ a_{12} = a_{22} \neq 0, b_{12} = b_{22} \neq 0 \end{cases}$$

Bezeichnen wir die eventuell von Null verschiedenen Koeffizienten in der Koeffizientenmatrix der Randbedingungen durch Punkte, dann können wir die oben genannten Fälle auf folgende Schemata zurückführen (eventuell durch Umnummerierung von L_1 und L_2 oder passende Linearkombinationen von L_1 und L_2):

$$\begin{array}{ll} \text{a)} & \begin{pmatrix} \cdot & 0 & \cdot & 0 \\ \cdot & 0 & \cdot & 0 \end{pmatrix} & \text{b)} & \begin{pmatrix} \cdot & 0 & \cdot & \cdot \\ \cdot & 0 & \cdot & 0 \end{pmatrix} \\ \\ \text{c)} & \begin{pmatrix} \cdot & \cdot & \cdot & 0 \\ \cdot & 0 & \cdot & 0 \end{pmatrix} & \text{d)} & \begin{pmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & 0 & \cdot & 0 \end{pmatrix} \end{array}$$

*) Wenn zwei entsprechende Koeffizienten in den Randbedingungen von Null verschieden sind, so kann man diese durch Multiplikation mit einem konstanten Faktor stets als gleich voraussetzen.

Wie früher wollen wir die Elemente in jedem Schema wieder mit a_{ik} bzw. b_{ik} ($i, k = 1, 2$) bezeichnen und untersuchen unter Benützung von (23) die Determinante $A(\lambda)$.

Zu Fall a). Es wird:

$$A(\lambda) = e^{\lambda\gamma_2} \{ [A_0] + [A_1] e^{\lambda\gamma_1} + [A_2] e^{2\lambda\gamma_1} \}$$

mit

$$A_0 = \eta_1(0) \eta_2(1) \begin{vmatrix} a_{11} & b_{11} \\ a_{21} & b_{21} \end{vmatrix}$$

$$A_1 = 0$$

$$A_2 = -\eta_1(1) \eta_2(0) \begin{vmatrix} a_{11} & b_{11} \\ a_{21} & b_{21} \end{vmatrix}$$

Es ist also der Satz II in § 4 wieder gültig, wenn man voraussetzt, daß

$$\begin{vmatrix} a_{11} & b_{11} \\ a_{21} & b_{21} \end{vmatrix} \neq 0 \text{ ist.}^*)$$

Das ist auch zugleich die Bedingung dafür, daß die Randbedingungen voneinander unabhängig sind.

Zu Fall b). Hier darf man $b_{12} \neq 0$ voraussetzen, da sonst Fall a) vorliegt. Man erhält:

$$A(\lambda) = \lambda e^{\lambda\gamma_2} \{ [A_0] + [A_1] e^{\lambda\gamma_1} + [A_2] e^{2\lambda\gamma_1} \}$$

wo

$$A_0 = a_{21} b_{12} \eta_1(0) \eta_2(1) \varphi_1(1)$$

$$A_1 = 2 b_{12} b_{21} \eta_1(1) \eta_2(1) \varphi_1(1)$$

$$A_2 = a_{21} b_{12} \eta_1(1) \eta_2(0) \varphi_1(1).$$

*) Als Beispiel erwähnen wir

$$\left. \begin{aligned} v'' - \pi^2 \lambda^2 v &= 0 \\ v(0) &= 0 \\ v(1) &= 0 \end{aligned} \right\} \text{ mit den Eigenwerten } \lambda = \pm i, \pm 2i, \pm 3i, \dots$$

Es muß also, da $b_{12} \neq 0$ ist, $a_{21} \neq 0$ vorausgesetzt werden, sonst versagt hier Satz II. Aus der Gleichung

$$A_0 + A_1 e^{\lambda \gamma_1} + A_2 e^{2\lambda \gamma_1} = 0$$

folgt dann:

$$\lambda_k^{(1)} = \frac{1}{\lambda_1} \operatorname{Lg} \left(\frac{-A_1 + \sqrt{A_1^2 - 4 A_0 A_2}}{2 A_2} \right) + \frac{2 k \pi i}{\gamma_1} \quad (k = 0, \pm 1, \pm 2, \dots)$$

$$\lambda_k^{(2)} = \frac{1}{\gamma_1} \operatorname{Lg} \left(\frac{-A_1 - \sqrt{A_1^2 - 4 A_0 A_2}}{2 A_2} \right) + \frac{2 k \pi i}{\gamma_1},$$

wobei Lg den Hauptwert des Logarithmus bedeutet. Wir erhalten also, solange $A_1^2 - 4 A_0 A_2 \neq 0$ ist, zwei verschiedene Reihen von einfachen Wurzeln und eine der früheren analoge Ueberlegung zeigt die Gültigkeit von Satz II. Ist aber $A_1^2 - 4 A_0 A_2 = 0$, so sind die zwei Wurzelreihen identisch, wir haben Doppelwurzeln. In diesem Fall kann also die Gleichung $\Delta(\lambda)$ für große $|\lambda|$ auch Doppelwurzeln haben.*)

Nehmen wir jetzt an, daß im Schema b) $a_{21} = 0$ sei. Es läßt sich dann zurückführen auf das folgende:

$$\begin{pmatrix} \cdot & 0 & 0 & \cdot \\ 0 & 0 & \cdot & 0 \end{pmatrix}$$

mit $b_{21} \neq 0$, sollen zwei unabhängige Randbedingungen vorliegen. Wendet man auf dieses Schema (23) an, so ergibt sich:

$$\Delta(\lambda) = b_{21} e^{\lambda \gamma_2} \{ [a_{11} \eta_1(0) \eta_2(1)] + 2 \lambda b_{12} [\eta_1(1) \eta_2(1) \varphi_1(1)] e^{\lambda \gamma_1} - [a_{11} \eta_1(1) \eta_2(0)] e^{2\lambda \gamma_1} \}.$$

Man muß also $b_{12} = 0$ annehmen, sollen die Betrachtungen von § 4 auch hier gelten. Damit hat man aber den Fall a).

*) Betrachten wir z. B. das Problem

$$\begin{cases} \lambda^2 v - v'' = \lambda u_0 + u_1 \\ L_1(v) \equiv v'(1, \lambda) = 0 \\ L_2(v) \equiv v(0, \lambda) + v(1, \lambda) = 0 \end{cases}$$

so ergibt sich $\Delta(\lambda) = \lambda e^{-\lambda} (1 + e^\lambda)^2$ und die λ -Werte $\lambda_k = i(1 + 2k)\pi$, wo $k = 0, \pm 1, \pm 2, \dots$, sind Doppelwurzeln. Für die Green'sche Funktion sind diese λ -Werte Pole zweiter Ordnung, wie man leicht nachrechnet.

Zu Fall c). Es ist

$$\Delta(\lambda) = \lambda e^{\lambda \gamma_2} \{ [A_0] + [A_1] e^{\lambda \gamma_1} + [A_2] e^{2\lambda \gamma_1} \}$$

wo

$$A_0 = a_{12} b_{21} \eta_1(0) \eta_2(1) \varphi_1(0)$$

$$A_1 = 2 a_{12} a_{21} \eta_1(0) \eta_2(0) \varphi_1(0)$$

$$A_2 = a_{12} b_{21} \eta_1(1) \eta_2(0) \varphi_1(0)$$

woraus sich ergibt, daß $a_{12} \cdot b_{21} \neq 0$ sein muß, damit wir Satz II beweisen können. Ist $b_{21} = 0$, dann muß man das Schema untersuchen

$$\begin{pmatrix} 0 & \cdot & 0 \\ \cdot & 0 & 0 \end{pmatrix}.$$

Wir erhalten:

$$\Delta(\lambda) = a_{21} e^{\lambda \gamma_2} \{ - [b_{11} \eta_1(0) \eta_2(1)] + 2 \lambda a_{12} [\eta_1(0) \eta_2(0) \varphi_1(0)] e^{\lambda \gamma_1} + [b_{11} \eta_1(1) \eta_2(0)] e^{2\lambda \gamma_1} \}.$$

Es muß also gelten $a_{12} = 0$ und $a_{21} b_{11} \neq 0^*$ d. h. es liegt dann Fall a) vor.

Zu Fall d). Hier darf man $a_{12} \neq 0$, $b_{12} \neq 0$ voraussetzen, da sonst einer der früher genannten Fälle vorliegt. Es wird:

$$\Delta(\lambda) = \lambda e^{\lambda \gamma_2} \{ [A_0] + [A_1] e^{\lambda \gamma_1} + [A_2] e^{2\lambda \gamma_1} \}$$

mit

$$A_0 = (a_{12} b_{21} \varphi_1(0) + a_{21} b_{12} \varphi_1(1)) \eta_1(0) \eta_2(1)$$

$$A_1 = 2 (a_{12} a_{21} \varphi_1(0) \eta_1(0) \eta_2(0) + b_{12} b_{21} \varphi_1(1) \eta_1(1) \eta_2(1))$$

$$A_2 = (a_{12} b_{21} \varphi_1(0) + a_{21} b_{12} \varphi_1(1)) \eta_1(1) \eta_2(0).$$

Wir müssen also fordern

$$a_{12} b_{21} \varphi_1(0) + a_{21} b_{12} \varphi_1(1) \neq 0^{**}).$$

*) z. B. hat das Randwertproblem

$$\left. \begin{aligned} v'' - \lambda^2 v &= 0 \\ v'(0) &= 0 \\ v(0) &= 0 \end{aligned} \right\}$$

nur die triviale Lösung $v=0$ und es existiert kein Eigenwert.

**) Betrachtet man das Randwertproblem

$$\left. \begin{aligned} \frac{1}{4} v'' - \lambda^2 v &= 0 \\ v'(0) + v'(1) &= 0 \\ v(0) - v(1) &= 0 \end{aligned} \right\}$$

wo also $a_{12} b_{21} \varphi_1(0) + a_{21} b_{12} \varphi_1(1) = 0$ ist, so ist jeder Wert λ ein Eigenwert.

Zusammenfassung: Satz II in § 4 bleibt, eventuell mit dem Zusatz, daß auch Doppelwurzeln auftreten können für große $|\lambda|$, richtig bei folgenden Randbedingungen:

- a) $\begin{cases} a_{11} u(0, t) + b_{11} u(1, t) = 0 \\ a_{21} u(0, t) + b_{21} u(1, t) = 0 \end{cases} \quad \text{mit} \quad \begin{vmatrix} a_{11} & b_{11} \\ a_{21} & b_{21} \end{vmatrix} \neq 0$
- b) $\begin{cases} a_{11} u(0, t) + b_{11} u(1, t) + u'(1, t) = 0 \\ u(0, t) + b_{21} u(1, t) = 0 \end{cases}$
- c) $\begin{cases} a_{11} u(0, t) + u'(0, t) + b_{11} u(1, t) = 0 \\ a_{21} u(0, t) + u(1, t) = 0 \end{cases}$
- d) $\begin{cases} a_{11} u(0, t) + a_{12} u'(0, t) + b_{11} u(1, t) + b_{12} u'(1, t) = 0 \\ a_{21} u(0, t) + b_{21} u(1, t) = 0 \end{cases}$

wobei im Fall d)

$$a_{12} b_{21} \varphi_1(0) + a_{21} b_{12} \varphi_1(1) \neq 0$$

sein muß.

Man beachte, daß der Fall d) die Periodizitätsbedingung enthält; man braucht nur

$$\begin{aligned} a_{11} &= b_{11} = 0 \\ a_{12} &= a_{21} = 1 \\ b_{12} &= b_{21} = -1 \end{aligned}$$

zu setzen. Es ist dann

$$a_{12} b_{21} \varphi_1(0) + a_{21} b_{12} \varphi_1(1) = -(\varphi_1(0) + \varphi_1(1)) \neq 0.$$

§ 6. Das asymptotische Verhalten der Green'schen Funktion¹²⁾

Satz III: Für große Werte von $|\lambda|$ gilt in der gelochten λ -Ebene die Ungleichung

$$(36) \quad |G(x, s; \lambda)| < \frac{G_0}{|\lambda|} \quad 0 \leq x, s \leq 1$$

wo G_0 nur eine Funktion von ϱ ist (§ 4).

Zur Begründung dieses Satzes beschränken wir uns zuerst auf den Fall $\Re(\lambda) \leq l$ und verwenden deshalb das Fundamentalsystem (12). Man hat:

$$p(s) \delta(s, \lambda) = p(s) \delta(0, \lambda) e^{-\int_0^s \frac{q(s)}{p(s)} ds}$$

mit

$$\delta(0, \lambda) = \begin{vmatrix} \lambda \varphi_1(0) [\eta_1(0)], & \lambda \varphi_2(0) [\eta_2(0)] \\ [\eta_1(0)], & [\eta_2(0)] \end{vmatrix} = \lambda [2 \eta_1(0) \eta_2(0) \varphi_1(0)].$$

Setzen wir

$$k(s) = 2 p(s) e^{-\int_0^s \frac{q(s)}{p(s)} ds} \cdot \eta_1(0) \eta_2(0) \varphi_1(0),$$

so ist

$$p(s) \delta(s, \lambda) = \lambda [k(s)]$$

wobei $k(s) \neq 0$ in $0 \leq x \leq 1$. Damit wird:

$$g(x, s; \lambda) = \pm \frac{1}{2 [k(s)]} \begin{cases} v_1(x, \lambda) v_2(s, \lambda) - v_2(x, \lambda) v_1(s, \lambda) \end{cases} \begin{cases} + & \text{für } x \geq s \\ - & \text{,, } x \leq s \end{cases}$$

und

$$L_i(g)_x = \frac{1}{2\lambda [k(s)]} \left\{ \begin{aligned} & a_{i1} (-v_1(0, \lambda) v_2(s, \lambda) + v_2(0, \lambda) v_1(s, \lambda)) \\ & + a_{i2} (-v_1'(0, \lambda) v_2(s, \lambda) + v_2'(0, \lambda) v_1(s, \lambda)) \\ & + b_{i1} (v_1(1, \lambda) v_2(s, \lambda) - v_2(1, \lambda) v_1(s, \lambda)) \\ & + b_{i2} (v_1'(1, \lambda) v_2(s, \lambda) - v_2'(1, \lambda) v_1(s, \lambda)) \end{aligned} \right\}.$$

Indem wir in $\Delta(x, s; \lambda)$ die

$$\begin{array}{l} 1. \\ 2. \end{array} \left\{ \begin{array}{l} \text{Kolonne mit} \end{array} \right\} \begin{cases} \frac{1}{2} \frac{v_2(s, \lambda)}{p(s) \delta(s, \lambda)} \\ \frac{1}{2} \frac{v_1(s, \lambda)}{p(s) \delta(s, \lambda)} \end{cases}$$

multiplizieren und zur letzten Kolonne addieren, erhalten wir

$$A(x, s; \lambda) = \begin{vmatrix} v_1(x, \lambda) & v_2(x, \lambda) & g_0(x, s; \lambda) \\ L_1(v_1) & L_1(v_2) & L_1(g)_x \\ L_2(v_1) & L_2(v_2) & L_2(g)_x \end{vmatrix}$$

wo

$$g_i(s, \lambda) = L_i(g)_x + \frac{1}{2} \frac{v_2(s, \lambda)}{p(s) \delta(s, \lambda)} L_i(v_1) + \frac{1}{2} \frac{v_1(s, \lambda)}{p(s) \delta(s, \lambda)} L_i(v_2) \\ (i = 1, 2)$$

und

$$g_0(x, s; \lambda) = \begin{cases} \frac{1}{p(s) \delta(s, \lambda)} v_1(x, \lambda) v_2(s, \lambda) & x \geq s \\ \frac{1}{p(s) \delta(s, \lambda)} v_2(x, \lambda) v_1(s, \lambda). & x \leq s \end{cases}$$

Wir erhalten weiter

$$g_i(s, \lambda) = \frac{v_2(s, \lambda)}{\lambda [k(s)]} \left\{ b_{i1} v_1(1, \lambda) + b_{i2} v_1'(1, \lambda) \right\} \\ + \frac{v_1(s, \lambda)}{\lambda [k(s)]} \left\{ a_{i1} v_2(0, \lambda) + a_{i2} v_2'(0, \lambda) \right\}.$$

Indem wir die Abkürzungen

$$(37) \quad \int_0^x \varphi_k(x) dx = X_k, \quad \int_0^s \varphi_k(s) ds = S_k$$

einführen, erhalten wir mit (24)

$$g_0(x, s; \lambda) = \begin{cases} \lambda^{-1} e^{\lambda(X_1 - S_1)} \left[\frac{\eta_1(x) \eta_2(s)}{k(s)} \right] & x \geq s \\ \lambda^{-1} e^{\lambda(X_2 - S_2)} \left[\frac{\eta_2(x) \eta_1(s)}{k(s)} \right] & x \leq s \end{cases}$$

und

$$g_i(s, \lambda) = e^{\lambda S_1} \left[A_{i2} \frac{\eta_1(s)}{k(s)} \right] + e^{\lambda(X_1 - S_1)} \left[B_{i1} \frac{\eta_2(s)}{k(s)} \right].$$

Die Ausdrücke (23₁) schreiben wir folgendermaßen:

$$\begin{aligned} L_i(v_1) &= \lambda e^{\lambda\gamma_2} \left\{ e^{\lambda\gamma_1} [A_{i1}] + e^{2\lambda\gamma_1} [B_{i1}] \right\} \\ L_i(v_2) &= \lambda e^{2\lambda\gamma_2} \left\{ e^{2\lambda\gamma_1} [A_{i2}] + e^{\lambda\gamma_1} [B_{i2}] \right\} \end{aligned} \quad (i = 1, 2)$$

und setzen zur Abkürzung

$$\begin{aligned} E_{i1}(\lambda) &= e^{\lambda\gamma_1} [A_{i1}] + e^{2\lambda\gamma_1} [B_{i1}] \\ E_{i2}(\lambda) &= e^{2\lambda\gamma_1} [A_{i2}] + e^{\lambda\gamma_1} [B_{i2}]; \end{aligned} \quad (i = 1, 2)$$

dann wird

$$\begin{aligned} L_i(v_1) &= \lambda e^{\lambda\gamma_2} E_{i1}(\lambda) \\ L_i(v_2) &= \lambda e^{2\lambda\gamma_2} E_{i2}(\lambda). \end{aligned} \quad (i = 1, 2)$$

Wenn man nun setzt

$$E(\lambda) = e^{\lambda\gamma_2} |E_{ik}(\lambda)|_2$$

und

$$G(x, s; \lambda) = g_o(x, s; \lambda) + G^{(o)}(x, s; \lambda),$$

so liefert eine einfache Umformung

$$g_o(x, s; \lambda) = \begin{cases} \lambda^{-1} e^{\lambda(X_1 - S_1)} \left[\frac{\eta_1(x) \eta_2(s)}{k(s)} \right] & x \geq s \\ \lambda^{-1} e^{\lambda(X_2 - S_2)} \left[\frac{\eta_2(x) \eta_1(s)}{k(s)} \right] & x \leq s \end{cases}$$

und

$$\begin{aligned} G^{(o)}(x, s; \lambda) &= \lambda^{-1} \left\{ \left(\left[B_{21} \frac{\eta_1(x) \eta_2(s)}{k(s)} \right] \frac{E_{12}}{E} - \left[B_{11} \frac{\eta_1(x) \eta_2(s)}{k(s)} \right] \frac{E_{22}}{E} \right) e^{\lambda X_1} \cdot e^{\lambda(\gamma_1 - S_1)} \right. \\ (43) \quad &+ \left(\left[A_{22} \frac{\eta_1(x) \eta_1(s)}{k(s)} \right] \frac{E_{12}}{E} - \left[A_{12} \frac{\eta_1(x) \eta_1(s)}{k(s)} \right] \frac{E_{22}}{E} \right) e^{\lambda X_1} \cdot e^{\lambda S_1} \\ &+ \left(\left[B_{11} \frac{\eta_2(x) \eta_2(s)}{k(s)} \right] \frac{E_{21}}{E} - \left[B_{21} \frac{\eta_2(x) \eta_2(s)}{k(s)} \right] \frac{E_{11}}{E} \right) e^{\lambda(\gamma_1 - X_1)} \cdot e^{\lambda(\gamma_1 - S_1)} \\ &\left. + \left(\left[A_{12} \frac{\eta_2(x) \eta_1(s)}{k(s)} \right] \frac{E_{21}}{E} - \left[A_{22} \frac{\eta_2(x) \eta_1(s)}{k(s)} \right] \frac{E_{11}}{E} \right) e^{\lambda(\gamma_1 - X_1)} \cdot e^{\lambda S_1} \right\}. \end{aligned}$$

Die Funktionen $\frac{E_{ik}(\lambda)}{E(\lambda)}$ ($i, k = 1, 2$) sind für große $|\lambda|$ und $\Re(\lambda) \leq l$ in der gelochten λ -Ebene beschränkt. Wir zeigen dies z. B. für $\frac{E_{i1}(\lambda)}{E(\lambda)}$. Es ist

$$E = e^{\lambda\gamma_2} |E_{ik}(\lambda)|_2 = e^{2\lambda\gamma_1} \cdot A_1(\lambda)$$

wo $A_1(\lambda)$ durch (27₁) definiert ist. Mit (27) erhält man:

$$E(\lambda) = e^{\lambda\gamma_1} \{ [A_0] + [A_1] e^{\lambda\gamma_1} + [A_2] e^{2\lambda\gamma_1} \}$$

also folgt mit (38)

$$(44) \quad \frac{E_{i1}(\lambda)}{E(\lambda)} = \frac{[A_{i1}] + [B_{i1}] e^{\lambda\gamma_1}}{A(\lambda) + \frac{E_0(\lambda) + E_1(\lambda) e^{\lambda\gamma_1} + E_2(\lambda) e^{2\lambda\gamma_1}}{\lambda}}$$

und $A(\lambda)$ ist durch (29) erklärt. Nach (35) ist in der ganzen gelochten λ -Ebene $|A(\lambda)| \geq m$, mit $m > 0$ und weil

$$\left| \frac{E_0(\lambda) + E_1(\lambda) e^{\lambda\gamma_1} + E_2(\lambda) e^{2\lambda\gamma_1}}{\lambda} \right| \leq \frac{M(1 + e^{l\gamma_1} + e^{2l\gamma_1})}{|\lambda|}$$

sobald $\Re(\lambda) \leq l$, so wird der Nenner im Ausdruck (44) größer als

$$m - \frac{M(1 + e^{l\gamma_1} + e^{2l\gamma_1})}{|\lambda|}.$$

Für genügend große $|\lambda|$ ist er also größer als $\frac{m}{2}$. Der Zähler im Ausdruck (44) ist beschränkt für $\Re(\lambda) \leq l$, wo l irgend eine endliche positive Zahl bedeutet. Damit ist also (36) bewiesen, sobald $\Re(\lambda) \leq l$; denn die in (42) und (43) auftretenden Exponentialfunktionen sind für $\Re(\lambda) \leq l$ beschränkt.

Unter Benützung von (14) und passende Umformung von $A(x, s; \lambda)$ gelingt für $\Re(\lambda) \geq -l$ eine analoge Darstellung der Green'schen Funktion. Wir schreiben

$$(45) \quad G(x, s; \lambda) = g^*(x, s; \lambda) + G^*(x, s; \lambda)$$

und die hier auftretenden Exponentialfunktionen sind für $\Re(\lambda) \geq -l$ gleichmäßig beschränkt. Damit ist (36) als richtig erkannt.*)

*) Für die in § 5 angegebenen Randbedingungen erhält man eine analoge asymptotische Darstellung der Green'schen Funktion.

§ 7. Lösung des inhomogenen Problems II, wobei wir $g(x, \lambda) \equiv 0$ voraussetzen

Jeder Pol der Green'schen Funktion ist ein Eigenwert unseres homogenen Problems II¹³⁾, also auch Nullstelle von $\Delta(\lambda)$. Wenn nun λ von einem Eigenwert verschieden ist, so ist das inhomogene Problem lösbar und man hat:

$$(46) \quad v(x, \lambda) = - \int_0^1 G(x, s; \lambda) \{ (a(s)\lambda + b(s)) u_0(s) + a(s) u_1(s) \} ds,$$

wobei also $v(x, \lambda)$ eine meromorphe Funktion von λ ist.

Nach Voraussetzung (6) ist $u_i(x)$ ($i = 1, 2$) Lösung von

$$(47) \quad \begin{cases} -(a(x)\lambda^2 + b(x)\lambda)v + L(v) = -(a(x)\lambda^2 + b(x)\lambda)u_i(x) + L(u_i(x)) \\ L_k(v) = 0 \end{cases} \quad (i, k = 1, 2)$$

d. h.

$$u_i(x) = \int_0^1 G(x, s; \lambda) \{ L(u_i(s)) - (a(s)\lambda^2 + b(s)\lambda) u_i(s) \} ds$$

woraus folgt:

$$\begin{aligned} - \int_0^1 G(x, s; \lambda) \{ a(s)\lambda + b(s) \} u_0(s) ds \\ = \frac{u_0(x)}{\lambda} - \frac{1}{\lambda} \int_0^1 G(x, s; \lambda) L(u_0(s)) ds \end{aligned}$$

und

$$\begin{aligned} - \int_0^1 G(x, s; \lambda) a(s) u_1(s) ds = \frac{u_1(x)}{\lambda^2} + \frac{1}{\lambda} \int_0^1 G(x, s; \lambda) b(s) u_1(s) ds \\ - \frac{1}{\lambda^2} \int_0^1 G(x, s; \lambda) L(u_1(s)) ds \end{aligned}$$

also

$$(48) \quad v(x, \lambda) = \frac{u_0(x)}{\lambda} + \frac{u_1(x)}{\lambda^2} - \frac{1}{\lambda} \int_0^1 G(x, s; \lambda) \left\{ L(u_0(s)) - b(s) u_1(s) + \frac{1}{\lambda} L(u_1(s)) \right\} ds.$$

§ 8. Untersuchung des Integrals

$$u(x, t) = \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} e^{\lambda t} v(x, \lambda) d\lambda.$$

Wir schicken zunächst einige Bemerkungen voraus über die im Verlaufe unserer Ausführungen auftretenden Integrale

$$(49) \quad \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} * d\lambda \quad \text{bzw.} \quad (50) \quad \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} * d\lambda.$$

Unter dem Integral (49) soll der Grenzwert eines Kurvenintegrals verstanden werden, was durch den dem Integralzeichen beigegeführten Kreis bezeichnet sei, wobei die auftretenden Integrationswege durch die unten näher beschriebenen Fig. 1 und 2 festgelegt sind. Im Gegensatz dazu soll das Integral (50) durch folgende Definition festgelegt sein:

$$\frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} * d\lambda = \lim_{R_1, R_2 \rightarrow \infty} \int_{\sigma-iR_1}^{\sigma+iR_2} * d\lambda.$$

Liefern (49) und (50) das gleiche Resultat, so gebrauchen wir die Darstellung (50).

$\sigma > 0$ bedeutet eine Zahl, die größer ist als die obere Schranke der Realteile der Eigenwerte des Problems II. Da nach § 4 alle Eigenwerte innerhalb eines Streifens endlicher Breite liegen, welcher parallel zur imaginären Achse ist, so existiert eine solche Zahl σ .

Wir untersuchen nun den Grenzwert des Kurvenintegrals

$$(51) \quad \frac{1}{2\pi i} \int_{\Gamma_n} e^{\lambda t} v(x, \lambda) d\lambda$$

wobei der Integrationsweg Γ_n durch die Figur 1 gegeben ist und $\lim_{n \rightarrow \infty} R_n = \infty$.

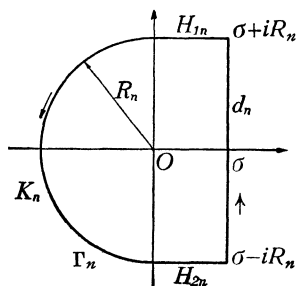


Fig. 1.

Γ_n treffe keinen der in § 4 betrachteten Kreise mit den Radien ϱ .
Setzen wir zur Abkürzung

$$(52) \quad v^*(x, \lambda) = \frac{1}{\lambda} \int_0^1 G(x, s; \lambda) \{L(u_0(s)) - b(s)u_1(s) + \frac{1}{\lambda} L(u_1(s))\} ds,$$

so können wir unter Benutzung von (48) schreiben

$$\frac{1}{2\pi i} \int_{\Gamma_n} e^{\lambda t} v(x, \lambda) d\lambda = u_0(x) + u_1(x)t - \frac{1}{2\pi i} \int_{\Gamma_n} e^{\lambda t} v^*(x, \lambda) d\lambda.$$

Das Integral rechts ist aber gleich

$$\int_{K_n} + \int_{H_{1n}} + \int_{H_{2n}} + \int_{d_n}.$$

Nun hat man für große $|\lambda|$, $|G(x, s; \lambda)| < \frac{G_0}{|\lambda|}$, gleichmäßig in x und s , also hat der Integrand die Ordnung $\frac{1}{|\lambda|^2}$, woraus sich ergibt, daß die

ersten drei Integrale mit wachsendem n gegen Null streben, während das vierte dem endlichen Grenzwert

$$(53) \quad \int_{\sigma-i\infty}^{\sigma+i\infty} e^{\lambda t} v^*(x, \lambda) d\lambda$$

zustrebt. Die Konvergenz ist übrigens gleichmäßig in x für $0 \leq x \leq 1$ und in t für $0 \leq t < \infty$.

Wegen der gleichmäßigen Konvergenz von (53) bezüglich t für $0 \leq t < \infty$ hat man

$$\lim_{t \rightarrow 0} \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} e^{\lambda t} v^*(x, \lambda) d\lambda = \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} v^*(x, \lambda) d\lambda.$$

Dieses letztere Integral ist aber gleich Null. Denn die einzig möglichen singulären Stellen der meromorphen Funktion v^* in der λ -Ebene sind die Pole von $G(x, s; \lambda)$ und $\lambda = 0$; folglich ist v^* regulär im Gebiet $\Re(\lambda) \geq \sigma$. Es verschwindet also das Integral

$$\int_{\Gamma_n^*} v^*(x, \lambda) d\lambda$$

wenn Γ_n^* der Integrationsweg der Fig. 2 ist.

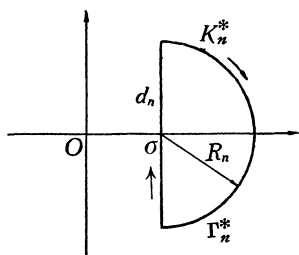


Fig. 2

Da der Integrand die Ordnung $\frac{1}{|\lambda|^2}$ hat, gilt

$$\lim_{R_n \rightarrow \infty} \int_{K_n^*} v^*(x, \lambda) d\lambda = 0,$$

und daher auch

$$(54) \quad \int_{\sigma-i\infty}^{\sigma+i\infty} v^*(x, \lambda) d\lambda = 0.$$

Definiert man also $u(x, t)$ durch

(55)

$$u(x, t) = \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} e^{\lambda t} v(x, \lambda) d\lambda = u_0(x) + u_1(x)t - \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} e^{\lambda t} v^*(x, \lambda) d\lambda,$$

so ist

$$\lim_{t \rightarrow 0} u(x, t) = u(x, 0) = u_0(x),$$

gleichmäßig in x für $0 \leq x \leq 1$.

Wenn man weiter beachtet, daß nach einer leichten Abschätzung

$$\lim_{R \rightarrow \infty} \frac{1}{2\pi i} \int_{K_R} \frac{e^{\lambda t}}{\lambda} d\lambda = 0$$

sich ergibt, gleichmäßig in t für $0 < \varepsilon \leq t \leq T < \infty$, und daß

$$\frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} \frac{e^{\lambda t}}{\lambda} d\lambda$$

für $0 < \varepsilon \leq t \leq T < \infty$ gleichmäßig konvergent ist, wie der zweite Mittelwertsatz zeigt, so kann man auch schreiben

$$u(x, t) = \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} e^{\lambda t} v(x, \lambda) d\lambda$$

für $t > 0$ und dieses Integral ist für $0 \leq x \leq 1$, $0 < \varepsilon \leq t \leq T < \infty$ gleichmäßig konvergent. Es gilt zudem*)

$$\lim_{t \rightarrow 0} \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} e^{\lambda t} v(x, \lambda) d\lambda = u_0(x)$$

*) Das Integral $\frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} v(x, \lambda) d\lambda$ ist nach den zu Beginn des § 8 gegebenen Erklärungen

natürlich sinnlos; es existiert aber der Grenzwert $\lim_{R \rightarrow \infty} \frac{1}{2\pi i} \int_{\sigma-iR}^{\sigma+iR} v(x, \lambda) d\lambda = \frac{u_0(x)}{2}$. Wir machen von dieser Tatsache im folgenden keinen Gebrauch.

§ 9. Untersuchung des Integrals

$$\frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} \lambda e^{\lambda t} v(x, \lambda) d\lambda.$$

Es werde der Grenzwert des folgenden Integrals untersucht für $\lim_{n \rightarrow \infty} R_n = \infty$

$$(56) \quad \frac{1}{2\pi i} \int_{\Gamma_n} \lambda e^{\lambda t} v(x, \lambda) d\lambda.$$

Dabei bedeutet Γ_n wieder den Integrationsweg der Figur 1. Wir benutzen dazu die asymptotische Darstellung der Green'schen Funktion.

Es ist

$$G(x, s; \lambda) = g_0(x, s; \lambda) + G^{(0)}(x, s; \lambda)$$

und die rechts auftretenden Funktionen sind durch (42), (43) gegeben. Die Einsetzung liefert:

$$\begin{aligned} \int_{K_n} \lambda e^{\lambda t} v^*(x, \lambda) d\lambda &= \int_{K_n} e^{\lambda t} d\lambda \int_0^1 g_0(x, s; \lambda) \left\{ L(u_0(s)) - b(s) u_1(s) + \right. \\ &\quad \left. \frac{1}{\lambda} L(u_1(s)) \right\} ds + \int_{K_n} e^{\lambda t} d\lambda \int_0^1 G^{(0)}(x, s; \lambda) \left\{ L(u_0(s)) - \right. \\ &\quad \left. b(s) u_1(s) + \frac{1}{\lambda} L(u_1(s)) \right\} ds. \end{aligned}$$

Nach § 8 genügt es, die folgenden zwei Integrale zu betrachten.

$$(57) \quad \mathcal{F}_1 = \int_{K_n} e^{\lambda t} d\lambda \int_0^1 g_0(x, s; \lambda) \left\{ L(u_0(s)) - b(s) u_1(s) \right\} ds$$

und

$$(58) \quad \mathcal{F}_2 = \int_{K_n} e^{\lambda t} d\lambda \int_0^1 G^{(0)}(x, s; \lambda) \left\{ L(u_0(s)) - b(s) u_1(s) \right\} ds.$$

Setzen wir

$$(59) \quad \Phi_v(s) = \frac{\eta_v(s)}{k(s)} \left\{ L(u_0(s)) - b(s) u_1(s) \right\}$$

und beschäftigen wir uns zunächst mit (57). Nach (42) ist

$$\mathcal{F} = \mathcal{F}_1^* + \mathcal{F}_2^*$$

wo

$$(57_1) \quad \mathcal{F}_1^* = \eta_1(x) \int_{K_n} \frac{e^{\lambda t}}{\lambda} d\lambda \int_0^x \Phi_2(s) e^{\lambda(X_1 - S_1)} ds$$

und

$$(57_2) \quad \mathcal{F}_2^* = \eta_2(x) \int_{K_n} \frac{e^{\lambda t}}{\lambda} d\lambda \int_x^1 \Phi_1(s) e^{\lambda(X_2 - S_2)} ds.$$

Untersuchen wir (57₁). Nach Voraussetzung (6) und wegen

$$\frac{dS_1(s)}{ds} = \varphi_1(s) \neq 0$$

liefert eine partielle Integration nach s

$$\int_0^x \Phi_2(s) e^{\lambda(X_1 - S_1)} ds = -\frac{1}{\lambda} \left\{ \frac{\Phi_2(s)}{\varphi_1(s)} e^{\lambda(X_1 - S_1)} \right\}_0^x + \frac{1}{\lambda} \int_0^x \left(\frac{\Phi_2(s)}{\varphi_1(s)} \right)' e^{\lambda(X_1 - S_1)} ds$$

also hat der Integrand die Ordnung $\frac{1}{|\lambda|^2}$, d. h.

$$\lim_{n \rightarrow \infty} \mathcal{F}_1^* = 0$$

gleichmäßig in x für $0 \leq x \leq 1$ und in t für $0 \leq t \leq T < \infty$. Analog für (57₂).

Betrachten wir nun (58). Nach (43) genügt es, ein Integral der Form zu untersuchen

$$\int_{K_n} \frac{e^{\lambda t}}{\lambda} d\lambda \int_0^1 \Phi_v(s) \mathcal{C}(\lambda) e^{\lambda f(x)} \cdot e^{\lambda h(s)} ds$$

wo $\mathcal{C}(\lambda)$ für große $|\lambda|$ beschränkt ist und $\Phi_v(s)$ die Bedeutung (59) hat. Beachtet man, daß $h'(s) \neq 0$ ist in $0 \leq s \leq 1$, so liefert eine partielle Integration nach s für den Integranden die Ordnung $\frac{1}{|\lambda|^2}$, woraus folgt, daß $\mathcal{Y}_2$ mit wachsendem n gleichmäßig in x für $0 \leq x \leq 1$ und in t für $0 \leq t \leq T < \infty$ gegen Null strebt.

Die Integrale über die horizontalen Strecken gehen mit $\frac{1}{R_n}$ gegen Null, wieder gleichmäßig in x, t für die vorhin genannten Intervalle. Wir schließen also daraus, daß das Integral

$$\int_{\sigma-i\infty}^{\sigma+i\infty} \lambda e^{\lambda t} v^*(x, \lambda) d\lambda$$

gleichmäßig in x und t konvergiert für $0 \leq x \leq 1, 0 \leq t \leq T < \infty$.

Man hat nun

$$\frac{\partial u(x, t)}{\partial t} = u_1(x) - \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} \lambda e^{\lambda t} v^*(x, \lambda) d\lambda = \frac{1}{2\pi i} \oint_{\sigma-i\infty}^{\sigma+i\infty} \lambda e^{\lambda t} v(x, \lambda) d\lambda.$$

Wegen der gleichmäßigen Konvergenz bezüglich t gilt:

$$\lim_{t \rightarrow 0} \frac{\partial u(x, t)}{\partial t} = \left. \frac{\partial u(x, t)}{\partial t} \right|_{t=0} = u_1(x) - \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} \lambda v^*(x, \lambda) d\lambda = \frac{1}{2\pi i} \oint_{\sigma-i\infty}^{\sigma+i\infty} \lambda v(x, \lambda) d\lambda.$$

Das vorletzte Integral ist aber gleich Null, wie eine analoge Betrachtung zu der des § 8 zeigt. Man braucht sich dazu nur der Beziehung (45) und der dort gemachten Bemerkung zu erinnern.

Wir haben damit den

Satz IV: Unter den oben gemachten Voraussetzungen darf man in der Formel (55) unter dem Integralzeichen einmal partiell nach t derivieren. $u(x, t), \frac{\partial u(x, t)}{\partial t}$ sind stetige Funktionen in x, t und es gilt

$$u(x, 0) = \lim_{t \rightarrow 0} u(x, t) = u_0(x) \quad *)$$

$$\left. \frac{\partial u(x, t)}{\partial t} \right|_{t=0} = \lim_{t \rightarrow 0} \frac{\partial u(x, t)}{\partial t} = u_1(x)$$

*) Man beachte die Definition von $u(x, t)$ durch (55), § 8.

§ 10. Die Funktion $u(x, t)$ erfüllt die homogenen Randbedingungen

Wir weisen nach, daß man in der Formel (55) die Integrale rechts einmal unter dem Integralzeichen nach x ableiten darf. Aus (12) und den Herleitungen des § 6 ergibt sich die asymptotische Darstellung von $\frac{\partial G(x, s; \lambda)}{\partial x}$ wie folgt:

$$\frac{\partial G}{\partial x} = \frac{\partial g_0}{\partial x} + \frac{\partial G^{(0)}}{\partial x}$$

mit

$$\frac{\partial g_0}{\partial x} = \begin{cases} e^{\lambda(X_1 - S_1)} \left[\frac{\varphi_1(x) \eta_1(x) \eta_2(s)}{k(s)} \right] \\ e^{\lambda(X_2 - S_2)} \left[\frac{\varphi_2(x) \eta_2(x) \eta_1(s)}{k(s)} \right] \end{cases}$$

und

$$\begin{aligned} \frac{\partial G^{(0)}}{\partial x} = & \left(\left[B_{21} \frac{\varphi_1(x) \eta_1(x) \eta_2(s)}{k(s)} \right] \frac{E_{12}}{E} \right. \\ & \left. - \left[B_{11} \frac{\varphi_1(x) \eta_1(x) \eta_2(s)}{k(s)} \right] \frac{E_{22}}{E} \right) e^{\lambda X_1} \cdot e^{\lambda(\gamma_1 - S_1)} \\ & + \left(\left[A_{22} \frac{\varphi_1(x) \eta_1(x) \eta_1(s)}{k(s)} \right] \frac{E_{12}}{E} - \left[A_{12} \frac{\varphi_1(x) \eta_1(x) \eta_1(s)}{k(s)} \right] \frac{E_{22}}{E} \right) e^{\lambda X_1} \cdot e^{\lambda S_1} \\ & + \left(\left[B_{11} \frac{\varphi_2(x) \eta_2(x) \eta_2(s)}{k(s)} \right] \frac{E_{21}}{E} - \left[B_{21} \frac{\varphi_2(x) \eta_2(x) \eta_2(s)}{k(s)} \right] \frac{E_{11}}{E} \right) e^{\lambda(\gamma_1 - X_1)} \cdot e^{\lambda(\gamma_1 - S_1)} \\ & + \left(\left[A_{12} \frac{\varphi_2(x) \eta_2(x) \eta_1(s)}{k(s)} \right] \frac{E_{21}}{E} - \left[A_{22} \frac{\varphi_2(x) \eta_2(x) \eta_1(s)}{k(s)} \right] \frac{E_{11}}{E} \right) e^{\lambda(\gamma_1 - X_1)} \cdot e^{\lambda S_1}. \end{aligned}$$

Wir betrachten jetzt das Integral

$$\begin{aligned} \int_{\sigma - i\infty}^{\sigma + i\infty} e^{\lambda t} \frac{\partial v^*(x, \lambda)}{\partial x} d\lambda = & \int_{\sigma - i\infty}^{\sigma + i\infty} \frac{e^{\lambda t}}{\lambda} d\lambda \int_0^1 \frac{\partial G}{\partial x} \left\{ L(u_0(s)) \right. \\ & \left. - b(s) u_1(s) + \frac{1}{\lambda} L(u_1(s)) \right\} ds. \end{aligned}$$

Die asymptotische Darstellung von $\frac{\partial G}{\partial x}$ zeigt aber, daß die Ordnung von $\frac{\partial G}{\partial x}$ für große $|\lambda|$ durch $O(1)$ gegeben ist, gleichmäßig in x und s für $0 \leq x, s \leq 1$. Die gleiche partielle Integration, wie sie in § 9 benützt wurde, liefert für $\frac{\partial v^*}{\partial x}$ die Ordnung $\frac{1}{|\lambda|^2}$ gleichmäßig in x, s für $0 \leq x, s \leq 1$; also konvergiert das obige Integral gleichmäßig bezüglich x und t für $0 \leq x \leq 1, 0 \leq t \leq T < \infty$. Man schließt daraus

$$\frac{\partial u(x, t)}{\partial x} = u_0'(x) + u_1'(x)t - \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} e^{\lambda t} \frac{\partial v^*}{\partial x} d\lambda = \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} e^{\lambda t} \frac{\partial v(x, \lambda)}{\partial x} d\lambda,$$

und es gilt zudem:

$$\lim_{t \rightarrow 0} \frac{\partial u(x, t)}{\partial x} = \frac{\partial u(x, t)}{\partial x} \Big|_{t=0} = \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} \frac{\partial v(x, \lambda)}{\partial x} d\lambda = u_0'(x) - \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} \frac{\partial v^*}{\partial x} d\lambda.$$

Wie früher zeigt man auch hier, daß das letztere Integral Null ist. Also gilt:

$$\frac{\partial u(x, t)}{\partial x} \Big|_{t=0} = u_0'(x).$$

Nun erfüllt $v(x, \lambda)$ und wegen (6) und (52) auch $v^*(x, \lambda)$ die Randbedingungen, also gilt das gleiche für $u(x, t)$.

§ 11. Die Funktion $u(x, t)$ erfüllt fast überall in x und t die homogene Differentialgleichung I_1)

Wir zeigen jetzt, daß die nach x und t integrierte homogene Differentialgleichung I_1)

$$a(x) \frac{\partial^2 u}{\partial t^2} + b(x) \frac{\partial u}{\partial t} - L(u) = 0$$

durch (55) erfüllt wird.

Die integrierte Gleichung lautet:

$$\begin{aligned}
 (60) \quad & \int_0^x a(x) \left\{ \frac{\partial u}{\partial t} - u_1(x) \right\} dx + \int_0^x b(x) \left\{ u - u_0(x) \right\} dx \\
 & - \int_0^t \left\{ p(x) \frac{\partial u}{\partial x} - p(0) \frac{\partial u}{\partial x} \Big|_{x=0} - \int_0^x p'(x) \frac{\partial u}{\partial x} dx \right\} dt \\
 & - \int_0^t \left\{ q(x) u - q(0) u(0, t) - \int_0^x q'(x) u dx \right\} dt \\
 & - \int_0^x r(x) dx \int_0^t u dt = 0.
 \end{aligned}$$

Wegen der gleichmäßigen Konvergenz des in (55) auftretenden Integrals

$$\frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} e^{\lambda t} v^*(x, \lambda) d\lambda$$

hat man

$$\begin{aligned}
 \int_0^t u(x, t) dt = u_0(x) t + u_1(x) \frac{t^2}{2} - \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} \frac{e^{\lambda t}}{\lambda} v^*(x, \lambda) d\lambda \\
 + \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} \frac{v^*(x, \lambda)}{\lambda} d\lambda.
 \end{aligned}$$

Das zweite Integral rechts ist aber Null, wie die Ausführungen des § 8 zeigen. Es bleibt:

$$(61) \quad \int_0^t u(x, t) dt = u_0(x) t + u_1(x) \frac{t^2}{2} - \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} \frac{e^{\lambda t}}{\lambda} v^*(x, \lambda) d\lambda$$

Substituieren wir (55) und (61) in die linke Seite von (60), so wird diese linke Seite:

$$\begin{aligned}
(62) \quad & t \left\{ \int_0^x b(x) u_1(x) dx - \left(p(x) u_0'(x) - p(0) u_0'(0) \right) + \int_0^x p'(x) u_0'(x) dx \right. \\
& - \left(q(x) u_0(x) - q(0) u_0(0) \right) + \int_0^x q'(x) u_0(x) dx - \int_0^x r(x) u_0(x) dx \left. \right\} \\
& - \frac{t^2}{2} \left\{ \left(p(x) u_1'(x) - p(0) u_1'(0) \right) - \int_0^x p'(x) u_1'(x) dx \right. \\
& + \left(q(x) u_1(x) - q(0) u_1(0) \right) - \int_0^x q'(x) u_1(x) dx + \int_0^x r(x) u_1(x) dx \left. \right\} \\
& - \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} \frac{e^{\lambda t}}{\lambda} \left\{ \int_0^x (a(x)\lambda^2 + b(x)\lambda) v^* dx - \left(p(x) \frac{\partial v^*}{\partial x} - p(0) \frac{\partial v^*}{\partial x} \Big|_{x=0} \right) \right. \\
& + \int_0^x p'(x) \frac{\partial v^*}{\partial x} dx - \left(q(x) v^* - q(0) v^*(0, \lambda) \right) \\
& \left. + \int_0^x q'(x) v^* dx - \int_0^x r(x) v^* dx \right\} d\lambda.
\end{aligned}$$

Die Vertauschung der Integrationsfolge ist nämlich wegen der gleichmäßigen Konvergenz der auftretenden Integrale gestattet und nach den §§ 8, 9, 10 hat man:

$$\int_{\sigma-i\infty}^{\sigma+i\infty} \frac{1}{\lambda} \frac{\partial v^*}{\partial x} d\lambda = 0.$$

Nun genügt

$$v(x, \lambda) = \frac{u_0(x)}{\lambda} + \frac{u_1(x)}{\lambda^2} - v^*(x, \lambda)$$

der Gleichung

$$\begin{aligned}
& (a(x)\lambda^2 + b(x)\lambda) v - \left(p(x) \frac{\partial v^2}{\partial x^2} + q(x) \frac{\partial v}{\partial x} + r(x) v \right) \\
& = \left(a(x)\lambda + b(x) \right) u_0(x) + a(x) u_1(x),
\end{aligned}$$

also ergibt sich durch Einsetzung und mittelst partieller Integration nach x

$$\begin{aligned} & \frac{1}{\lambda} \left\{ \int_0^x b(x) u_1(x) dx - \left(p(x) u_0'(x) - p(0) u_0'(0) \right) + \int_0^x p'(x) u_0'(x) dx \right. \\ & \quad - \left(q(x) u_0(x) - q(0) u_0(0) \right) + \int_0^x q'(x) u_0(x) dx - \int_0^x r(x) u_0(x) dx \left. \right\} \\ & \quad - \frac{1}{\lambda^2} \left\{ \left(p(x) u_1'(x) - p(0) u_1'(0) \right) - \int_0^x p'(x) u_1'(x) dx \right. \\ & \quad + \left(q(x) u_1(x) - q(0) u_1(0) \right) - \int_0^x q'(x) u_1(x) dx + \int_0^x r(x) u_1(x) dx \left. \right\} \\ & - \left\{ \int_0^x (a(x) \lambda^2 + b(x) \lambda) v^* dx - \left(p(x) \frac{\partial v^*}{\partial x} - p(0) \frac{\partial v^*}{\partial x} \Big|_{x=0} \right) + \int_0^x p'(x) \frac{\partial v^*}{\partial x} dx \right. \\ & \quad \left. - \left(q(x) v^* - q(0) v^*(0, \lambda) \right) + \int_0^x q'(x) v^* dx - \int_0^x r(x) v^* dx \right\} = 0. \end{aligned}$$

Multipliziert man diese Gleichung mit $\frac{1}{2\pi i} \frac{e^{\lambda t}}{\lambda}$ und integriert nach λ , so sieht man, daß der Ausdruck (62) gleich Null ist.

Wir schreiben nun (60) in der Form

$$\begin{aligned} (60_1) \quad & \int_0^x a(x) \frac{\partial u}{\partial t} dx - p(x) \int_0^t \frac{\partial u}{\partial x} dt + p(0) \int_0^t \frac{\partial u}{\partial x} \Big|_{x=0} dt = \\ & = - \int_0^t dt \int_0^x p'(x) \frac{\partial u}{\partial x} dx + \int_0^x a(x) u_1(x) dx - \int_0^x b(x) (u - u_0(x)) dx \\ & \quad + \int_0^t (q(x) u - q(0) u(0, t)) dt - \int_0^t dt \int_0^x q'(x) u dx + \int_0^x r(x) dx \int_0^t u dt. \end{aligned}$$

Die rechte Seite dieser Gleichung ist nach x differentiierbar, also auch die linke, woraus folgt:

$$\begin{aligned} & a(x) \frac{\partial u}{\partial t} - p(x) \frac{\partial}{\partial x} \int_0^t \frac{\partial u}{\partial x} dt = \\ & = a(x) u_1(x) - b(x) (u - u_0(x)) \int_0^t q(x) \frac{\partial u}{\partial x} dt + r(x) \int_0^t u dt. \end{aligned}$$

Da die rechte Seite nach t differentiierbar ist, so ergibt sich weiter

$$(63) \quad \frac{\partial}{\partial t} \left\{ a(x) \frac{\partial u}{\partial t} - p(x) \frac{\partial}{\partial x} \int_0^t \frac{\partial u}{\partial x} dt \right\} = q(x) \frac{\partial u}{\partial x} - b(x) \frac{\partial u}{\partial t} + r(x) u.$$

Wir setzen nun

$$(64) \quad u^*(x, t) = \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} e^{\lambda t} v^*(x, \lambda) d\lambda$$

und weisen nach, daß $\frac{\partial^2 u^*}{\partial t^2}$ fast überall existiert für jedes x .

Wir haben

$$\begin{aligned} \int_{\sigma-i\infty}^{\sigma+i\infty} \lambda e^{\lambda t} v^*(x, \lambda) d\lambda &= \int_{\sigma-i\infty}^{\sigma+i\infty} e^{\lambda t} d\lambda \int_0^1 G(x, s; \lambda) \left\{ L(u_0(s)) - b(s) u_1(s) \right. \\ &\quad \left. + \frac{1}{\lambda} L(u_1(s)) \right\} ds. \end{aligned}$$

Setzen wir

$$(65) \quad \mathfrak{F}(x, \lambda) = \lambda^2 v^*(x, \lambda) = \lambda \int_0^1 G(x, s; \lambda) \left\{ L(u_0(s)) - b(s) u_1(s) + \frac{1}{\lambda} L(u_1(s)) \right\} ds$$

und $\lambda = \sigma + i\eta$, so gilt nach den früheren Betrachtungen

$$\mathfrak{F}(x, \lambda) = O\left(\frac{1}{|\lambda|}\right)$$

d. h. das Integral

$$\int_{-\infty}^{+\infty} |\mathfrak{F}(x, \lambda)|^2 |d\lambda|$$

konvergiert gleichmäßig in x für $0 \leq x \leq 1$. Wir haben daher

$$(66) \quad \int_{-\infty}^{+\infty} |\mathfrak{F}(x, \sigma + i\eta)|^2 d\eta < M,$$

wo M von x unabhängig ist.

Man hat nun

$$\int_{\sigma-i\infty}^{\sigma+i\infty} \lambda e^{\lambda t} v^*(x, \lambda) d\lambda = \int_{\sigma-i\infty}^{\sigma+i\infty} \frac{e^{\lambda t}}{\lambda} \mathfrak{F}(x, \lambda) d\lambda$$

und daraus schließt man^{10) 11)}:

$$(67) \quad \frac{\partial}{\partial t} \left(\frac{\partial u^*}{\partial t} \right) = \frac{\partial}{\partial t} \left(\frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} \lambda e^{\lambda t} v^*(x, \lambda) d\lambda \right)$$

existiert für jedes x fast überall, d. h. mit Ausnahme einer Menge vom Maß Null. Nach (55) gilt also das gleiche von $\frac{\partial^2 u}{\partial t^2}$.

Differentiiert man (60₁) zuerst partiell nach t und dann nach x , so folgt überall in x und t :

$$(68) \quad \frac{\partial}{\partial x} \left\{ \frac{\partial}{\partial t} \int_0^x a(x) \frac{\partial u}{\partial t} dx - p(x) \frac{\partial u}{\partial x} \right\} = -p'(x) \frac{\partial u}{\partial x} + q(x) \frac{\partial u}{\partial x} - b(x) \frac{\partial u}{\partial t} + r(x)u.$$

Nun gilt:

$$a(x) \frac{\partial^2 u^*}{\partial t^2} = a(x) \frac{\partial}{\partial t} \left(\frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} \lambda e^{\lambda t} v^*(x, \lambda) d\lambda \right)$$

fast überall in t für jedes x , wobei u^* durch (64) erklärt ist. Beachtet man, daß die Beziehung (66) besteht, so ergibt sich^{10) 11)}:

$$\int_0^x a(x) \frac{\partial}{\partial t} \int_{\sigma-i\infty}^{\sigma+i\infty} \lambda e^{\lambda t} v^*(x, \lambda) d\lambda dx = \frac{\partial}{\partial t} \int_{\sigma-i\infty}^{\sigma+i\infty} \lambda e^{\lambda t} d\lambda \int_0^x a(x) v^*(x, \lambda) dx.$$

Differentiiert man nach x , so erhält man, abgesehen von einer zweidimensionalen Punktmenge vom Maß Null,

$$a(x) \frac{\partial}{\partial t} \int_{\sigma-i\infty}^{\sigma+i\infty} \lambda e^{\lambda t} v^*(x, \lambda) d\lambda = \frac{\partial}{\partial x} \left(\frac{\partial}{\partial t} \int_{\sigma-i\infty}^{\sigma+i\infty} \lambda e^{\lambda t} d\lambda \int_0^x a(x) v^*(x, \lambda) dx \right)$$

d. h.

$$(69) \quad a(x) \frac{\partial^2 u^*}{\partial t^2} = \frac{\partial}{\partial x} \left\{ \frac{\partial}{\partial t} \int_0^x a(x) \frac{\partial u^*}{\partial t} dx \right\}.$$

Es muß nun $\frac{\partial}{\partial x} \left(p(x) \frac{\partial u}{\partial x} \right)$ fast überall in x existieren, d. h. fast überall in x gilt:

$$\frac{\partial}{\partial x} \left(p(x) \frac{\partial u}{\partial x} \right) = p'(x) \frac{\partial u}{\partial x} + p(x) \frac{\partial^2 u}{\partial x^2}.$$

Wenn man jetzt (63), (68), (69) kombiniert, so erhält man den

Satz V: Abgesehen von einer zweidimensionalen Punktmenge vom Maß Null genügt die Funktion $u(x, t) = \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} e^{\lambda t} v(x, \lambda) d\lambda$ der homogenen Differentialgleichung I_1).

§ 12. Lösung des inhomogenen Problems I

Für $f(x, t)$ seien die Voraussetzungen (7) erfüllt. Wir betrachten nun die Funktion

$$(70) \quad g(x, \lambda) = \int_0^\infty e^{-\lambda t} f(x, t) dt.$$

Nach einem Satz von *M. Lerch*⁸⁾ folgt, wegen der gleichmäßigen Konvergenz von $\int_0^\infty |f(x, t)| dt$ bezüglich x in $0 \leq x \leq 1$, daß $g(x, \lambda)$ für $\Re(\lambda) > 0$ analytisch in λ ist und stetig in x .

Wir haben:

$$(71) \quad g(x, \lambda) = \frac{1}{\lambda} f(x, 0) + \frac{1}{\lambda^2} \frac{\partial f}{\partial x} \Big|_{t=0} + \frac{1}{\lambda^2} \int_0^\infty e^{-\lambda t} \frac{\partial^2 f}{\partial x^2} dt.$$

Nun hat $G(x, s; \lambda)$ die Ordnung $\frac{1}{|\lambda|}$, also gilt:

$$(72) \quad \left| \int_0^1 G(x, s; \lambda) g(s, \lambda) ds \right| = O\left(\frac{1}{|\lambda|^2}\right)$$

d. h. das Integral

$$(73) \quad U(x, t) = -\frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} e^{\lambda t} d\lambda \int_0^1 G(x, s; \lambda) g(s, \lambda) ds \quad *)$$

konvergiert gleichmäßig in x für $0 \leq x \leq 1$ und in t für $0 \leq t \leq T < \infty$.

Es gilt also:

$$U(x, 0) = -\frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} d\lambda \int_0^1 G(x, s; \lambda) g(s, \lambda) ds.$$

Da $g(x, \lambda)$ analytisch ist für $\Re(\lambda) > 0$ und $G(x, s; \lambda)$ analytisch für $\Re(\lambda) \geq \sigma$, so folgt:

$$\int_{\Gamma_n^*} d\lambda \int_0^1 G(x, s; \lambda) g(s, \lambda) ds = 0,$$

wenn Γ_n^* der in Fig. 2 angegebene Integrationsweg ist. Da weiter die Ordnung des Integranden $\frac{1}{|\lambda|^2}$ ist, so strebt das Integral längs des Halbkreises K_n^* gegen Null und zwar gleichmäßig in x für $0 \leq x \leq 1$, woraus sich ergibt:

$$U(x, 0) = 0.$$

Wir betrachten jetzt das Integral

$$\begin{aligned} & \int_{\sigma-i\infty}^{\sigma+i\infty} \lambda e^{\lambda t} d\lambda \int_0^1 G(x, s; \lambda) g(s, \lambda) ds \\ &= \int_{\sigma-i\infty}^{\sigma+i\infty} e^{\lambda t} d\lambda \int_0^1 G(x, s; \lambda) \left\{ f(s, 0) + \frac{1}{\lambda} \frac{\partial f(s, 0)}{\partial t} + \frac{1}{\lambda^2} \int_0^\infty e^{-\lambda \tau} \frac{\partial^2 f(s, \tau)}{\partial \tau^2} d\tau \right\} ds. \end{aligned}$$

*) Wir verweisen auf die in § 8 gegebene Definition dieses Integrals.

Zur Konvergenzfrage für dieses Integral genügt es also, das Folgende zu betrachten

$$(74) \quad \int_{\sigma-i\infty}^{\sigma+i\infty} e^{\lambda t} d\lambda \int_0^1 G(x, s; \lambda) f(s, 0) ds.$$

Wegen der in (7) vorausgesetzten stetigen Ableitung $\frac{\partial f(s, 0)}{\partial s}$ liefert eine partielle Integration nach s zusammen mit der asymptotischen Darstellung von $G(x, s; \lambda)$, genau wie in § 9, für den Integranden die Ordnung $\frac{1}{|\lambda|^2}$, gleichmäßig in x für $0 \leq x \leq 1$ und in t für $0 \leq t \leq T < \infty$, woraus die gleichmäßige Konvergenz des Integrals (74) sich ergibt für $0 \leq x \leq 1$, $0 \leq t \leq T < \infty$.

Wir haben daher

$$\frac{\partial U(x, t)}{\partial t} = -\frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} \lambda e^{\lambda t} d\lambda \int_0^1 G(x, s; \lambda) g(s, \lambda) ds$$

mit

$$\left. \frac{\partial U(x, t)}{\partial t} \right|_{t=0} = 0,$$

das letztere wie in den §§ 8, 9. Die Betrachtungen des § 10 lassen unter den Voraussetzungen (7) sogleich die gleichmäßige Konvergenz des Integrals

$$\int_{\sigma-i\infty}^{\sigma+i\infty} e^{\lambda t} d\lambda \int_0^1 \frac{\partial G(x, s; \lambda)}{\partial x} g(s, \lambda) ds$$

erkennen, wobei wieder $0 \leq x \leq 1$, $0 \leq t \leq T < \infty$. Daraus folgt somit

$$\frac{\partial U(x, t)}{\partial x} = -\frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} e^{\lambda t} d\lambda \int_0^1 \frac{\partial G(x, s; \lambda)}{\partial x} g(s, \lambda) ds.$$

Da die Green'sche Funktion die Randbedingungen erfüllt, so gilt das gleiche von $U(x, t)$.

Nun bleibt noch der Nachweis, daß die Funktion $U(x, t)$ der Differentialgleichung I_1) genügt. Wir zeigen auch hier zunächst, daß $U(x, t)$ die integrierte Gleichung I_1) befriedigt.

Diese lautet:

$$\begin{aligned} & \int_0^x a(x) \frac{\partial u}{\partial t} dx + \int_0^x b(x) u dx - \int_0^t \left\{ p(x) \frac{\partial u}{\partial x} - p(0) \frac{\partial u}{\partial x} \Big|_{x=0} - \int_0^x p'(x) \frac{\partial u}{\partial x} dx \right\} dt \\ & - \int_0^t \left\{ q(x) u - q(0) u(0, t) - \int_0^x q'(x) u dx \right\} dt - \int_0^x r(x) dx \int_0^t u dt \\ & = \int_0^x dx \int_0^t f(x, t) dt. *) \end{aligned}$$

Setzen wir die Abkürzung

$$(75) \quad V(x, \lambda) = - \int_0^1 G(x, s; \lambda) g(s, \lambda) ds$$

und substituieren wir die rechte Seite von (73) in die linke Seite der integrierten Differentialgleichung, so folgt:

$$\begin{aligned} & \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} \frac{e^{\lambda t}}{\lambda} d\lambda \left\{ \int_0^x (a(x) \lambda^2 + b(x) \lambda) V dx - \left(p(x) \frac{\partial V}{\partial x} - p(0) \frac{\partial V}{\partial x} \Big|_{x=0} \right) \right. \\ (76) \quad & \left. - \int_0^x p'(x) \frac{\partial V}{\partial x} dx - (q(x) V - q(0) V(0, \lambda)) \right. \\ & \left. - \int_0^x q'(x) V dx - \int_0^x r(x) V dx \right\}. \end{aligned}$$

Da aber

$$(a(x) \lambda^2 + b(x) \lambda) V - p(x) \frac{\partial^2 V}{\partial x^2} - q(x) \frac{\partial V}{\partial x} - r(x) V = g(x, \lambda)$$

*) Wir bemerken, daß wir hier $u(x, 0) = 0$, $\frac{\partial u(x, t)}{\partial t} \Big|_{t=0} = 0$ voraussetzen dürfen.

ist, so liefert eine Integration nach x , daß der Ausdruck in der geschweiften Klammer von (76) gleich

$$\int_0^x g(x, \lambda) dx$$

ist. Damit ist (76) dem folgenden Integral gleich

$$\frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} \frac{e^{\lambda t}}{\lambda} d\lambda \int_0^x g(x, \lambda) dx.$$

Nun weisen wir nach, daß dieses Integral gleich ist

$$\int_0^x dx \int_0^t f(x, t) dt.$$

Setzen wir dazu

$$\Phi(x, t) = \int_0^x f(x, t) dx,$$

so konvergiert

$$\int_0^\infty |\Phi(x, t)| dt$$

gleichmäßig in x für $0 \leq x \leq 1$.

Denn wegen der vorausgesetzten gleichmäßigen Konvergenz von

$$\int_0^\infty |f(x, t)| dt$$

gilt:

$$\int_0^x dx \int_0^\infty |f(x, t)| dt = \int_0^\infty dt \int_0^x |f(x, t)| dx,$$

und da

$$\int_0^A |\Phi(x, t)| dt = \int_0^A dt \left| \int_0^x f(x, t) dx \right| \leq \int_0^A dt \int_0^x |f(x, t)| dx$$

ist, so konvergiert das Integral

$$\int_0^\infty |\Phi(x, t)| dt$$

gleichmäßig in x für $0 \leq x \leq 1$.

Wir setzen

$$(77) \quad F(x, t) = \int_0^x dx \int_0^t f(x, t) dt = \int_0^t dt \int_0^x f(x, t) dx$$

und zeigen, daß das Integral

$$\int_0^\infty e^{-\lambda t} F(x, t) dt$$

für $\Re(\lambda) \geq k > 0$ absolut konvergent ist.

Man hat

$$\begin{aligned} \int_0^A e^{-kt} dt \int_0^t dt \int_0^x |f(x, t)| dx &= -\frac{1}{k} \left\{ e^{-kt} \int_0^t dt \int_0^x |f(x, t)| dx \right\}_{t=0}^{t=A} \\ &\quad + \frac{1}{k} \int_0^A e^{-kt} dt \int_0^x |f(x, t)| dx, \end{aligned}$$

woraus für $A \rightarrow \infty$ folgt:

$$\int_0^\infty e^{-kt} dt \int_0^t dt \int_0^x |f(x, t)| dx = \frac{1}{k} \int_0^\infty e^{-kt} dt \int_0^x |f(x, t)| dx,$$

da ja das letztere Integral existiert. Die Konvergenz ist übrigens gleichmäßig in x für $0 \leq x \leq 1$. Wir haben also:

$$\int_0^{\infty} e^{-\lambda t} F(x, \lambda) d\lambda = \frac{1}{\lambda} \int_0^{\infty} e^{-\lambda t} dt \int_0^x f(x, t) dx.$$

Multiplizieren wir diese Gleichung mit $e^{\lambda t}$ und integrieren nach λ , so liefert die Laplace'sche Integraltransformation

$$\begin{aligned} F(x, t) &= \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} e^{\lambda t} d\lambda \int_0^{\infty} e^{-\lambda \tau} F(x, \tau) d\tau = \\ &= \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} \frac{e^{\lambda t}}{\lambda} d\lambda \int_0^{\infty} e^{-\lambda \tau} d\tau \int_0^x f(x, \tau) dx. \end{aligned}$$

Da wegen der gleichmäßigen Konvergenz bezüglich x die Gleichung besteht

$$\int_0^x g(x, \lambda) dx = \int_0^{\infty} e^{-\lambda \tau} d\tau \int_0^x f(x, \tau) dx$$

so ergibt sich

$$F(x, t) = \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} \frac{e^{\lambda t}}{\lambda} d\lambda \int_0^x g(x, \lambda) dx.$$

Nach (77) ist damit nachgewiesen, daß (73) Lösung der integrierten Gleichung ist.

§ 13. $U(x, t)$ erfüllt fast überall in x, t die Differentialgleichung 1.)

Wie in § 11 sieht man, daß überall in x, t die Gleichung gilt:

$$\frac{\partial}{\partial t} \left(a(x) \frac{\partial U}{\partial t} - p(x) \frac{\partial}{\partial x} \int_0^t \frac{\partial U}{\partial x} dt \right) = q(x) \frac{\partial U}{\partial x} - b(x) \frac{\partial U}{\partial t} + r(x) U + f(x, t).$$

Nach § 12, No. (71), (75), wissen wir, daß die Beziehung besteht

$$\lambda^2 V(x, \lambda) = O\left(\frac{1}{|\lambda|}\right),$$

woraus folgt^{10) 11)}, daß das Integral

$$\frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} \lambda e^{\lambda t} V(x, \lambda) d\lambda = \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} \frac{e^{\lambda t}}{\lambda} \lambda^2 V(x, \lambda) d\lambda$$

fast überall nach t differentiierbar ist für jedes x .

Der § 11 zeigt ferner, daß abgesehen von einer zweidimensionalen Punktmenge vom Maß Null die Gleichung besteht

$$a(x) \frac{\partial^2 U}{\partial t^2} = \frac{\partial}{\partial x} \left(\frac{\partial}{\partial t} \int_0^x a(x) \frac{\partial U}{\partial t} dx \right)$$

so daß man erhält:

Satz VI: Die Funktion $U(x, t)$ ist bis auf eine zweidimensionale Punktmenge vom Maß Null eine Lösung der Gleichung I_1). Sie erfüllt ferner die Randbedingungen I_3) und es gilt: $U(x, 0) = 0$, $\left. \frac{\partial U(x, t)}{\partial t} \right|_{t=0} = 0$.

Nach den §§ 8 bis 13 gilt ferner der

Satz VII: Bedeuten $v(x, \lambda)$ und $V(x, \lambda)$ die Lösungen der folgenden zwei Probleme,

$$\left\{ \begin{array}{l} (a(x)\lambda^2 + b(x)\lambda)v - L(v) = \\ = (a(x)\lambda + b(x))u_0 + a(x)u_1, \\ L_1(v) = 0 \\ L_2(v) = 0 \end{array} \right\} \quad \left\{ \begin{array}{l} (a(x)\lambda^2 + b(x)\lambda)V - L(V) = g(x, \lambda) \\ L_1(V) = 0 \\ L_2(V) = 0 \end{array} \right.$$

so ist, abgesehen von einer zweidimensionalen Punktmenge vom Maß Null, die Funktion

$$u(x, t) = \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} e^{\lambda t} v(x, \lambda) d\lambda + \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} e^{\lambda t} V(x, \lambda) d\lambda$$

eine Lösung des Problems I. σ bedeutet dabei eine positive Zahl, die größer ist als die obere Schranke der Realteile der Eigenwerte des Problems II.

Für $t > 0$ läßt sich die Funktion $u(x, t)$ noch in der Form darstellen:

$$u(x, t) = \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} e^{\lambda t} \mathcal{U}(x, \lambda) d\lambda,$$

wenn unter $\mathcal{U}(x, \lambda) = v(x, \lambda) + V(x, \lambda)$ die Lösung des Problems II verstanden wird.

Die zweite der Formeln (10) folgt aus der ersten mittels der Laplace'schen Integraltransformation.

§ 14. Der Entwicklungssatz

Es ist (§§ 8, 9, 10)

$$(78) \quad u(x, t) = \frac{1}{2\pi i} \oint_{\sigma-i\infty}^{\sigma+i\infty} e^{\lambda t} v(x, \lambda) d\lambda$$

$$(79) \quad \frac{\partial u(x, t)}{\partial t} = \frac{1}{2\pi i} \oint_{\sigma-i\infty}^{\sigma+i\infty} \lambda e^{\lambda t} v(x, \lambda) d\lambda$$

$$(80) \quad \frac{\partial u(x, t)}{\partial x} = \frac{1}{2\pi i} \oint_{\sigma-i\infty}^{\sigma+i\infty} e^{\lambda t} \frac{\partial v(x, \lambda)}{\partial x} d\lambda,$$

wobei die Integrale als in x, t gleichmäßige Grenzwerte von Kurvenintegralen zu betrachten sind für $0 \leq x \leq 1, 0 \leq t \leq T < \infty$.

Sei nun $\lambda = \lambda_v$ ein Pol von $v(x, \lambda)$. Wir bezeichnen das zu ihm gehörige Residuum von $e^{\lambda t} v(x, \lambda)$ bzw. $e^{\lambda t} \lambda v(x, \lambda)$ mit $e^{\lambda_v t} R_v^{(0)}(x, t)$ bzw. $e^{\lambda_v t} R_v^{(1)}(x, t)$ *). Dann gilt:

$$(81) \quad u(x, t) = \lim_{A \rightarrow \infty} \sum_{|\lambda_v| < A} e^{\lambda_v t} R_v^{(0)}(x, t)$$

*) $R_v^{(0)}(x, t), R_v^{(1)}(x, t)$ sind im allgemeinen Polynome in t .

$$(82) \quad \frac{\partial u(x, t)}{\partial t} = \lim_{A \rightarrow \infty} \sum_{|\lambda_v| < A} e^{\lambda_v t} R_v^{(1)}(x, t)$$

$$(83) \quad \frac{\partial u(x, t)}{\partial x} = \lim_{A \rightarrow \infty} \sum_{|\lambda_v| < A} e^{\lambda_v t} \frac{\partial R_v^{(0)}(x, t)}{\partial x}$$

gleichmäßig in x, t für $0 \leq x \leq 1, 0 \leq t \leq T < \infty$. Wird ferner das zu $\lambda = \lambda_v$ gehörige Residuum von $v(x, \lambda)$ bzw. $\lambda v(x, \lambda)$ mit $r_v^{(0)}(x)$ bzw. $r_v^{(1)}(x)$ bezeichnet, also $r_v^{(0)}(x) = R_v^{(0)}(x, 0), r_v^{(1)}(x) = R_v^{(1)}(x, 0)$, so gilt der

Satz VIII: Ist $u_0(x)$ dreimal stetig differentiierbar, ist $u_1(x)$ zweimal stetig differentiierbar und erfüllen $u_0(x), u_1(x)$ die Randbedingungen I_3 des Problems I, so gelten gleichmäßig in x im Intervall $0 \leq x \leq 1$ die Reihenentwicklungen für das Funktionenpaar $u_0(x), u_1(x)$

$$u_0(x) = \lim_{A \rightarrow \infty} \sum_{|\lambda_v| < A} r_v^{(0)}(x),$$

$$u_1(x) = \lim_{A \rightarrow \infty} \sum_{|\lambda_v| < A} r_v^{(1)}(x).$$

Ferner ist, ebenfalls gleichmäßig in x für $0 \leq x \leq 1$,

$$u_0'(x) = \lim_{A \rightarrow \infty} \sum_{|\lambda_v| < A} \frac{d r_v^{(0)}(x)}{d x}.$$

Für den Fall *einfacher Pole* wollen wir noch die Struktur der oben auftretenden Residuen und Reihenentwicklungen genauer angeben.

Sei also $\lambda = \lambda_v$ ein einfacher Pol der Green'schen Funktion $G(x, s; \lambda)$. λ_v kann dann ein einfacher oder zweifacher Eigenwert des homogenen Problems II sein. Ist $v_v(x)$ eine zugehörige Eigenfunktion, so gilt:

$$\left. \begin{aligned} (a(x) \lambda_v^2 + b(x) \lambda_v) v_v(x) - L(v_v(x)) &= 0 \\ L_1(v_v) &= 0 \\ L_2(v_v) &= 0 \end{aligned} \right\}$$

oder

$$\left\{ \begin{aligned} L(v_v) - (a(x) \lambda^2 + b(x) \lambda) v_v &= -(\lambda - \lambda_v) \{a(x) (\lambda + \lambda_v) + b(x)\} v_v(x) \\ L_1(v_v) &= 0 \\ L_2(v_v) &= 0 \end{aligned} \right.$$

Wir haben daher, solange λ von einem Pol von $G(x, s; \lambda)$ verschieden ist

$$v_v(x) = -(\lambda - \lambda_v) \int_0^1 G(x, s; \lambda) \{a(s)(\lambda + \lambda_v) + b(s)\} v_v(s) ds.$$

Nun hängt aber $v_v(x)$ nicht von λ ab, so daß auch

$$v_v(x) = -\lim_{\lambda \rightarrow \lambda_v} (\lambda - \lambda_v) \int_0^1 G(x, s; \lambda) \{a(s)(\lambda + \lambda_v) + b(s)\} v_v(s) ds$$

ist und da wir voraussetzen $\lambda = \lambda_v$ sei ein einfacher Pol, können wir schreiben

$$(84) \quad v_v(x) = -\int_0^1 R_v(x, s) \{2\lambda_v a(s) + b(s)\} v_v(s) ds.$$

Das zu λ_v gehörige Residuum von $G(x, s; \lambda)$ ist dabei mit $R_v(x, s)$ bezeichnet worden. Schreibt man das homogene Differentialsystem II in der Form

$$(85) \quad \left\{ \begin{array}{l} \{ (a(x)\lambda_v^2 + b(x)\lambda_v) + (\lambda - \lambda_v)(2a(x)\lambda_v + b(x)) + \\ \quad + (\lambda - \lambda_v)^2 a(x) \} v - L(v) = 0 \\ L_1(v) = 0 \\ L_2(v) = 0 \end{array} \right.$$

und benützt man die Entwicklung

$$(86) \quad G(x, s; \lambda) = \frac{R_v(x, s)}{\lambda - \lambda_v} + P(x, s; \lambda - \lambda_v),$$

so liefert die Einsetzung*) von (86) in (85), daß $R_v(x, s)$ eine Lösung des homogenen Problems II ist. Daher muß $R_v(x, s)$ eine lineare Kombination der zu $\lambda = \lambda_v$ gehörigen Eigenfunktionen des homogenen Problems II sein.

Die Green'sche Funktion ist ferner als Funktion von s betrachtet Lösung des zu II adjungierten Differentialsystems^{3) 14)} und für beide

*) Für $x \neq s$ ist $G(x, s; \lambda)$ Lösung des homogenen Problems II, wenn λ von einem Pol verschieden ist.

Differentialsysteme ist der Eigenwert λ_v von gleicher Ordnung³⁾. Bezeichnet man daher mit $v_v^{(1)}(x)$, $v_v^{(2)}(x)$ die Eigenfunktionen des homogenen Problems II für $\lambda = \lambda_v$ und mit $w_v^{(1)}$, $w_v^{(2)}$ die des adjungierten Differentialsystems, so gilt:

$$(87) \quad R_v(x, s) = \sum_{\alpha, \beta=1, 2} c_{\alpha\beta} v_v^{(\alpha)}(x) w_v^{(\beta)}(s).$$

Unter Berücksichtigung von (84) erhält man somit

$$(88) \quad v_v^{(k)}(x) = - \sum_{\alpha, \beta=1, 2} c_{\alpha\beta} v_v^{(\alpha)}(x) \int_0^1 v_v^{(k)}(s) w_v^{(\beta)}(s) (2 \lambda_v a(s) + b(s)) ds$$

$$(k = 1, 2).$$

Nun kann man aber, in Verbindung mit (88), stets voraussetzen*), daß

$$(88_1) \quad \int_0^1 v_v^{(\alpha)}(s) w_v^{(\beta)}(s) (2 \lambda_v a(s) + b(s)) ds = \begin{cases} 0 & \alpha \neq \beta \\ 1 & \alpha = \beta \end{cases}$$

woraus sich nach (88) ergibt:

$$v_v^{(1)}(x) = - (c_{11} v_v^{(1)}(x) + c_{21} v_v^{(2)}(x)).$$

Wegen der linearen Unabhängigkeit von $v_v^{(1)}$, $v_v^{(2)}$ in $0 \leq x \leq 1$ folgt daher

$$c_{11} = -1$$

$$c_{21} = 0$$

und analog

$$c_{12} = 0$$

$$c_{22} = -1.$$

Man hat also

$$(89) \quad R_v(x, s) = - (v_v^{(1)}(x) w_v^{(1)}(s) + v_v^{(2)}(x) w_v^{(2)}(s)).$$

*) Vergleiche: *Goursat, E.:* Cours d'analyse, t. III, vierte Auflage (1927), S. 391.

Nun ergibt sich, nach (46) und (89), das zu $\lambda = \lambda_v$ gehörige Residuum $r_v^{(0)}(x)$ von $v(x, \lambda)$ im Fall eines einfachen Eigenwertes

$$(90) \quad r_v^{(0)}(x) = v_v(x) \int_0^1 w_v(s) \{ (\lambda_v a(s) + b(s)) u_0(s) + a(s) u_1(s) \} ds$$

und für den Fall eines zweifachen Eigenwertes

$$r_v^{(0)}(x) = \sum_{k=1}^{k=2} v_v^{(k)}(x) \int_0^1 w_v^{(k)}(s) \{ (\lambda_v a(s) + b(s)) u_0(s) + a(s) u_1(s) \} ds.$$

Das Residuum $r_v^{(1)}(x)$ von $\lambda v(x, \lambda)$ für $\lambda = \lambda_v$ wird

$$r_v^{(1)}(x) = \lambda_v r_v^{(0)}(x).$$

Hat man also nur einfache Pole und einfache Eigenwerte, so lauten die Reihenentwicklungen für $u(x, t)$, $\frac{\partial u(x, t)}{\partial t}$, nach (78), (79), (80),

$$u(x, t) = \lim_{A \rightarrow \infty} \sum_{|\lambda_v| < A} A_v e^{\lambda_v t} v_v(x),$$

$$\frac{\partial u(x, t)}{\partial t} = \lim_{A \rightarrow \infty} \sum_{|\lambda_v| < A} \lambda_v A_v e^{\lambda_v t} v_v(x),$$

gleichmäßig in x, t für $0 \leq x \leq 1$, $0 \leq t \leq T < \infty$,

mit

$$\boxed{\begin{aligned} u_0(x) &= \lim_{A \rightarrow \infty} \sum_{|\lambda_v| < A} A_v v_v(x) \\ u_1(x) &= \lim_{A \rightarrow \infty} \sum_{|\lambda_v| < A} \lambda_v A_v v_v(x) \end{aligned}}$$

gleichmäßig in x für $0 \leq x \leq 1$. $v_v(x)$ ist dabei die zu $\lambda = \lambda_v$ gehörige Eigenfunktion des homogenen Problems II und A_v berechnet sich nach (90) zu

$$A_v = \int_0^1 w_v(s) \{ (\lambda_v a(s) + b(s)) u_0(s) + a(s) u_1(s) \} ds.$$

Hierin bedeutet $w_\nu(s)$ die zu $\lambda = \lambda_\nu$ gehörige Eigenfunktion des zu II adjungierten Differentialsystems und über $v_\nu(x)$, $w_\nu(x)$ ist nach (88₁) so verfügt, daß

$$\int_0^1 v_\nu(s) w_\nu(s) (2\lambda_\nu a(s) + b(s)) ds = 1$$

ist.

Ist das Problem II speziell selbstadjungiert ³⁾, so kann man $v_\nu(x) = w_\nu(x)$ setzen und dabei $v_\nu(x)$ mit einer passenden Konstanten multiplizieren um

$$\int_0^1 v_\nu^2(s) (2\lambda_\nu a(s) + b(s)) ds = 1$$

zu erhalten. In diesem Fall gilt dann:

$$A_\nu = \int_0^1 v_\nu(s) \{ (\lambda_\nu a(s) + b(s)) u_0(s) + a(s) u_1(s) \} ds. *)$$

*) Man kann stets voraussetzen, eventuell nach Multiplikation mit $\frac{1}{p(x)} e^{\int_0^x \frac{q(x)}{p(x)} dx}$, daß (I) selbstadjungiert sei. Das zum homogenen Problem I gehörige Problem II lautet dann:

$$\left. \begin{aligned} (A(x)\lambda^2 + B(x)\lambda)v - \frac{d}{dx} \left(P(x) \frac{dv}{dx} \right) - R(x)v &= 0 \\ L_1(v) &= 0 \\ L_2(v) &= 0 \end{aligned} \right\}.$$

Sind λ_1, λ_2 zwei verschiedene Eigenwerte und v_1, v_2 zugehörige Eigenfunktionen, so gilt die verallgemeinerte Orthogonalitätsrelation:

$$(\lambda_1 + \lambda_2) \int_0^1 A(x) v_1 v_2 dx + \int_0^1 B(x) v_1 v_2 dx = 0 \quad 4)$$

sobald $\left\{ P(x)(v_2 v_1' - v_1 v_2') \right\}_{x=0}^{x=1} = 0$, d. h. das Differentialsystem ist dann selbstadjungiert.

Unter gewissen Voraussetzungen über die Funktionen $A(x), B(x)$ lassen sich daraus einige leicht herzuleitende Aussagen über die Lage der komplexen Eigenwerte in der λ -Ebene machen.

Literatur

- 1) *Birkhoff, G. D.*: On the asymptotic character of solutions of certain linear diff. equat. etc.
Trans. Am. Math. Soc. 9 (1908).
- 2) *Birkhoff, G. D.*: Boundary value and expansion problems etc.
Trans. Am. Math. Soc. 9 (1908).
- 3) *Bôcher, M.*: Leçons sur les méthodes de Sturm.
Collection de monographies sur la théorie des fonctions; Gauthier-Villars et Cie., Paris (1917).
- 4) *Bromwich, T. J. I' A.*: Normal coordinates in dynamical systems.
Proc. London Math. Soc. (2) 15 (1917).
- 5) *Doetsch, G.*: Probleme aus der Theorie der Wärmeleitung.
I., II., III. Mitteilung. (I. Mitteilung von *Bernstein, F.* und *Doetsch, G.*).
Math. Zeitschrift 22 (1925) und 25 (1926).
- 6) *Doetsch, G.*: Elektrische Schwingungen in einem anfänglich strom- und spannungslosen Kabel unter dem Einfluß einer Rand-erregung.
Festschrift der Techn. Hochschule Stuttgart zur Vollendung ihres ersten Jahrhunderts 1829—1929.
- 7) *Heaviside, O.*: Literaturangaben in dem Buche von Jeffreys, H.: Operational methods in mathematical physics.
Cambridge Tracts in Mathematics and Mathematical Physics; No. 23; second edition (1931).
- 8) *Lerch, M.*: Sur un point de la théorie des fonctions génératrices d'Abel.
Acta math. 27 (1903).
- 9) *Plancherel, M.*: Sur le développement d'un couple de fonctions arbitraires en séries de fonctions fondamentales d'un problème aux limites du type hyperbolique.
Atti del congresso Internazionale dei Matematici, Bologna (1928) VI.
- 10) *Plancherel, M.*: Contribution à l'étude de la représentation d'une fonction arbitraire par des intégrales définies.
Rend. di Palermo 30 (1910).
- 11) *Plancherel, M.*: Sur les formules d'inversion de Fourier et de Hankel.
Proc. London Math. Soc. (2) 24 (1925).
- 12) *Tamarkine, J.*: Some general problems of the theory of ordinary linear diff. equat. etc.
Math. Zeitschrift 27 (1927).
- 13) *Tamarkine, J.*: Ueber einige allg. Probleme aus der Theorie der gewöhnlichen Differentialgleichungen und über Reihenentwicklung von Funktionen.
Buch in russischer Sprache; Petrograd (1917).
- 14) *Westfall, W. D. A.*: Zur Theorie der Integralgleichungen.
Diss. Göttingen (1905).

(Eingegangen den 6. Juni 1932)

Ein allgemeines quadratisches Reziprozitätsgesetz in denjenigen algebraischen Zahlkörpern, worin 2 voll zerfällt.

Von Th. SKOLEM, Bergen.

Im folgenden betrachte ich in Zahlkörpern, worin 2 vollständig zerfällt, Jacobische Symbole, die in gewöhnlicher Art als Produkte von Legendreschen Symbolen definiert sein sollen. Dabei soll aber nicht wie gewöhnlich der Fall ausgeschlossen sein, daß ein Primfaktor l von 2 unter dem Strich (im „Nenner“) auftritt. Weil jede zu l prime Zahl $\equiv 1 \pmod{l}$ ist, wenn 2 voll zerfällt, und also quadratischer Rest mod. l ist, so soll immer

$$\left(\frac{\alpha}{l}\right) = +1$$

gesetzt werden; dabei ist also natürlich α nicht teilbar durch l . Es sollen nämlich immer der „Zähler“ und der „Nenner“ der Jacobischen Symbole relativ prim sein. Nach der gemachten Vereinbarung hat dann das Symbol

$$\left(\frac{\alpha}{a}\right)$$

immer denselben Wert wie

$$\left(\frac{\alpha}{\bar{a}}\right),$$

wenn $\bar{a}$ der zu 2 prime Bestandteil von a ist, d. h. das Produkt derjenigen Primfaktoren von a , die in 2 nicht aufgehen. Man erkennt sehr leicht, daß die fundamentalen Sätze über die Jacobischen Symbole noch gültig bleiben, nämlich

$$1) \left(\frac{\alpha}{ab}\right) = \left(\frac{\alpha}{a}\right)\left(\frac{\alpha}{b}\right) \qquad 2) \left(\frac{\alpha\beta}{a}\right) = \left(\frac{\alpha}{a}\right)\left(\frac{\beta}{a}\right)$$

$$3) \text{ Aus } \alpha \equiv \beta \pmod{a} \text{ folgt } \left(\frac{\alpha}{a}\right) = \left(\frac{\beta}{a}\right).$$

Diese Erweiterung der Bedeutung des Jacobischen Symbols wird nun zuerst im absoluten Körper k gemacht. Ich setze, wenn $\left(\frac{a, k}{b}\right)$ solche Symbole im erweiterten Sinne sind,

$$\left(\frac{a, k}{b}\right) \left(\frac{b, k}{a}\right) (-1)^{\frac{\text{sign } a-1}{2} \frac{\text{sign } b-1}{2}} = (a, k, b).$$

Dann ist $(a, k, b) = (b, k, a)$ und $(a, k, bc) = (a, k, b) (a, k, c)$, wie man leicht konstatiert. Folgende Tatsachen können bemerkt werden:

Satz 1. *a) Sind a und b ungerade, so ist (a, k, b) bestimmt durch die Restklassen, wozu a und b gehören mod. 4.*

Beweis. Es ist für a und b ungerade

$$(a, k, b) = (-1)^{\frac{a-1}{2} \frac{b-1}{2}}.$$

b) Ist a teilbar durch genau 2^n , $n > 0$, b ungerade, so ist (a, k, b) bestimmt durch die Restklassen, wozu a und b bezw. mod. 2^{n+2} und 2^3 gehören.

Beweis. Es ist

$$(a, k, b) = \left(\frac{a, k}{b}\right) \left(\frac{b, k}{a}\right) (-1)^{\frac{\text{sign } a-1}{2} \frac{\text{sign } b-1}{2}} = \left(\frac{2, k}{b}\right)^n \left(\frac{a}{2^n}, k, b\right),$$

und $\left(\frac{2, k}{b}\right)$ ist bestimmt durch den Rest von b mod. 8, während nach

a) $\left(\frac{a}{2^n}, k, b\right)$ bestimmt ist durch die Reste von $\frac{a}{2^n}$ und b mod. 4.

Weiter betrachte ich einen beliebigen Körper K vom Grade n , worin 2 voll zerfällt. Ich setze

$$(2) = I_1 I_2 \dots I_n.$$

Die Zahlen aus K sollen durch kleine griechische Buchstaben, die Ideale aus K durch kleine deutsche Buchstaben bezeichnet werden, während absolut rationale Zahlen durch kleine lateinische Buchstaben bezeichnet werden.

Indem $\left(\frac{\alpha, K}{\beta}\right)$ das Jacobische Symbol im erweiterten Sinne ist, setze ich

$$\left(\frac{\alpha, K}{\beta}\right) \left(\frac{\beta, K}{\alpha}\right) (-1)^{s \frac{\text{sign } \alpha-1}{2} \frac{\text{sign } \beta-1}{2}} = (\alpha, K, \beta)^{-1}.$$

¹⁾ Augenscheinlich ist wieder $(\alpha, K, \beta) = (\beta, K, \alpha)$ und $(\alpha, K, \beta\gamma) = (\alpha, K, \beta) (\alpha, K, \gamma)$.

Das Ziel dieses kleinen Aufsatzes ist nun zu zeigen, daß jedes Symbol (α, K, β) als ein Produkt von n „Komponentensymbolen“ (α, k, β) ausgedrückt werden kann, wobei die n Komponentensymbole den n Primfaktoren von 2 entsprechen. Wie dies zu verstehen ist, wird aus dem folgenden hervorgehen. In den Beweisen hierfür benutze ich keine höheren Mittel als die elementare Idealtheorie; freilich muß ich aber das quadratische Reziprozitätsgesetz für Zahlen, die prim zu 2 sind, als bekannt voraussetzen, nämlich ²⁾

$$\left(\frac{\alpha}{\beta}\right)\left(\frac{\beta}{\alpha}\right) = (-1)^{s \frac{\alpha-1}{2} \frac{\beta-1}{2} + s \frac{\text{sign } \alpha-1}{2} \frac{\text{sign } \beta-1}{2}}.$$

Bisweilen werde ich die bekannte „Nenner“transformationsformel

$$\left(\frac{\alpha, K}{\alpha}\right) = \left(\frac{\alpha, k}{N\alpha}\right)$$

benutzen, deren Richtigkeit in fast trivialer Weise beweisbar ist ³⁾. Daß diese Formel auch gültig bleibt bei der gemachten Erweiterung des Begriffes „Jacobisches Symbol“ ist ohne weiteres klar.

Zuerst sollen 2 Hilfssätze bewiesen werden.

Hilfssatz 1. Es sei ω eine ganze primitive Zahl in K und $\equiv 1 \pmod{I'_1}$, dagegen $\equiv 0 \pmod{I'_i}$, wenn $i > 1$ ist. Dann ist $S\omega \equiv 1 \pmod{2^f}$.

Beweis. Die irreduzible Gleichung, der ω genügt, sei $f(\omega) = 0$. Sobald e hinreichend groß ist, gilt die Funktionenkongruenz

$$f(x) \equiv (x - a_1) \dots (x - a_n) \pmod{2^e}$$

und zwar derart, ⁴⁾ daß für jedes i I_i^f aufgeht in $\omega - a_i$. Also gelten dann die Kongruenzen

$$a_1 \equiv 1 \pmod{I_1^f}, \quad a_2 \equiv 0 \pmod{I_2^f}, \quad \dots, \quad a_n \equiv 0 \pmod{I_n^f}$$

und also auch

$$a_1 \equiv 1, \quad a_2 \equiv a_3 \equiv \dots \equiv a_n \equiv 0 \pmod{2^f}.$$

²⁾ Vgl. *H. Hasse*, Bericht über neuere Untersuchungen und Probleme aus der Theorie der algebraischen Zahlkörper, Teil II, S. 77.

³⁾ *H. Hasse*, Über das allgemeine Reziprozitätsgesetz etc., Journal für Math. 154, S. 107.

⁴⁾ Vgl. *R. Fricke*, Lehrbuch der Algebra, Bd. 3, Kap. 2, § 14.

Folglich wird, da $e \geq f$,

$$S\omega \equiv a_1 + a_2 + \dots + a_n \equiv 1 \pmod{2^f},$$

w. z. bw. w.

Hilfssatz 2. Es sei α eine ganze Zahl in K und für jedes i teilbar durch genau $l_i^{f_i}$ und $\equiv a_i \pmod{l_i^{f_i+2}}$. Dann ist ⁵⁾ $N\alpha \equiv \prod_i a_i \pmod{2^{\sum f_i + 2}}$.

Beweis. Zuerst zeige ich, daß der Satz gilt für jede Wahl der Zahlen a_i , falls er für eine solche gilt. Es sei nämlich für jedes i $a_i \equiv b_i \pmod{2^{f_i+2}}$. Dann ist b_i genau durch 2^{f_i} teilbar wie a_i , so daß man

$$a_i = 2^{f_i} a'_i, \quad b_i = 2^{f_i} b'_i, \quad a'_i \equiv b'_i \equiv \pm 1 \pmod{4}$$

schreiben kann. Dann wird

$$\prod_i a'_i \equiv \prod_i b'_i \pmod{4},$$

woraus

$$\prod_i a_i = 2^{\sum f_i} \prod_i a'_i \equiv 2^{\sum f_i} \prod_i b'_i = \prod_i b_i \pmod{2^{\sum f_i + 2}}.$$

Ich betrachte erstens eine primitive Zahl α . Die Gleichung für α in k sei $f(\alpha) \equiv 0$. Dann gilt für beliebiges e eine Funktionenkongruenz

$$f(x) \equiv (x - b_1) \dots (x - b_n) \pmod{2^e},$$

und wenn e hinreichend groß ist ⁴⁾, so ist $\alpha - b_i$ teilbar durch $l_i^{f_i+2}$. Außerdem soll $e \geq \sum f_i + 2$ gewählt werden. Dann ist

$$N\alpha \equiv \prod_i b_i \pmod{2^e} \quad \text{und also auch} \quad \pmod{2^{\sum f_i + 2}}.$$

Ist für jedes i dann $\alpha \equiv a_i \pmod{l_i^{f_i+2}}$, so ist $a_i \equiv b_i \pmod{2^{f_i+2}}$. Nach der zuerst gemachten Bemerkung folgt daraus $N\alpha \equiv \prod_i a_i \pmod{2^{\sum f_i + 2}}$.

Zweitens sei α imprimitiv. Dann kann man sicher eine primitive Zahl β finden, welche $\equiv \alpha \pmod{2^{\sum f_i + 2}}$ ist. Für jedes i ist dann auch $\beta \equiv a_i$

⁵⁾ Überall im folgenden bedeuten $\prod_i$ und $\sum_i$ dasselbe wie $\prod_{i=1}^n$ bzw. $\sum_{i=1}^n$.

$(\text{mod. } l_i^{f_i+2})$, woraus man nach dem schon bewiesenen $N\beta \equiv \prod_i a_i \text{ mod. } 2^{\sum f_i+2}$ bekommt. Da aber offenbar $N\alpha \equiv N\beta \pmod{2^{\sum f_i+2}}$, folgt $N\alpha \equiv \prod_i a_i \text{ mod. } 2^{\sum f_i+2}$.

Satz 2. Es seien α und β ganze Zahlen in K , relativ prim und prim zu 2. Es sei

$$\alpha \equiv a_i, \beta \equiv b_i \pmod{l_i^2}.$$

Dann ist

$$(\alpha, K, \beta) = \prod_i (a_i, k, b_i).$$

Beweis: Man kann n ganze und primitive Zahlen $\omega_1, \dots, \omega_n$ in K finden derart, daß

$$\omega_i \equiv 1 \pmod{l_i^2}, \omega_i \equiv 0 \pmod{l_j^2}, \text{ wenn } j \neq i.$$

Nach dem Hilfssatz 1 ist dann für alle i $S\omega_i \equiv 1 \pmod{4}$. Außerdem hat man

$$\alpha \equiv a_1 \omega_1 + \dots + a_n \omega_n, \beta \equiv b_1 \omega_1 + \dots + b_n \omega_n \pmod{4};$$

denn jede dieser Kongruenzen gilt ja mod. l_i^2 für alle i , und deshalb wird

$$\frac{\alpha-1}{2} = \sum_i \frac{a_i-1}{2} \omega_i, \quad \frac{\beta-1}{2} = \sum_j \frac{b_j-1}{2} \omega_j \pmod{2}$$

und

$$\frac{\alpha-1}{2} \frac{\beta-1}{2} = \sum_h \frac{a_h-1}{2} \frac{b_h-1}{2} \omega_h^2 \pmod{2};$$

denn so oft $i \neq j$ ist, ist $\omega_i \omega_j \equiv 0 \pmod{4}$. Nach dem Hilfssatz wird aber auch $S\omega_h^2 \equiv 1 \pmod{4}$, und deshalb bekommt man

$$S \frac{\alpha-1}{2} \frac{\beta-1}{2} = \sum_h \frac{a_h-1}{2} \frac{b_h-1}{2} \pmod{2}.$$

Da

$$(\alpha, K, \beta) = (-1)^{S \frac{\alpha-1}{2} \frac{\beta-1}{2}} \quad \text{und} \quad (\alpha, k, b) = (-1)^{\frac{\alpha-1}{2} \frac{b-1}{2}},$$

so ist hierdurch Satz 2 bewiesen.

Satz 3. Es seien α und β zwei relativ prime ganze Zahlen, $\beta \equiv 1 \pmod{2}$, α teilbar durch genau 1_1^e und $\equiv 1 \pmod{\frac{2}{1_1}}$. Ist $\alpha \equiv a_1 \pmod{1_1^{e+2}}$ und für jedes $i > 1 \equiv a_i \pmod{1_i^2}$, während $\beta \equiv b_1 \pmod{1_1^3}$ und für jedes $i > 1 \equiv b_i \pmod{1_i^2}$, so ist

$$(\alpha, K, \beta) = \prod_i (\alpha_i, k, b_i).$$

Beweis: Zuerst soll der Satz bewiesen werden für eine beliebige Zahl β aber gewisse spezielle Zahlen α , nämlich

$$\alpha = \gamma^2 - \beta r,$$

wo r eine solche der Zahlen 1, 3, 5, 7 ist, daß $\beta r \equiv 1 \pmod{1_1^3}$ ist, während γ so bestimmt wird, daß

$$\gamma^2 - \beta r \text{ teilbar durch genau } 1_1^e \text{ wird und } \gamma \equiv 0 \pmod{\frac{2}{1_1}},$$

was offenbar alles stets möglich ist. Man kann auch γ prim zu β wählen, so daß auch α prim zu β wird.

Der zum Zahlenpaare (α, β) gehörende Signaturfaktor ist $= +1$; denn ist eine Konjugierte zu β negativ, so ist die entsprechende Konjugierte zu $\gamma^2 - \beta r$ augenscheinlich positiv. Also wird, wenn man bemerkt, daß

$$\left(\frac{\alpha, K}{r}\right) = \left(\frac{\gamma^2, K}{r}\right) = +1$$

ist,

$$\begin{aligned} (I) \quad (\alpha, K, \beta) &= \left(\frac{\alpha, K}{\beta}\right) \left(\frac{\beta, K}{\alpha}\right) = \left(\frac{\gamma^2, K}{\beta}\right) \left(\frac{\gamma^2 r, K}{\alpha}\right) = \\ &= \left(\frac{r, K}{\alpha}\right) \left(\frac{\alpha, K}{r}\right) = \left(\frac{r, k}{N\alpha}\right) \left(\frac{N\alpha, k}{r}\right) = (r, k, N\alpha). \end{aligned}$$

Nun ist $(N\alpha, k, r) = \left(\frac{2, k}{r}\right)^e \left(\frac{N\alpha}{2^e}, k, r\right)$ und nach dem Hilfssatze 2

$$\frac{N\alpha}{2^e} \equiv \frac{a_1}{2^e} a_2 a_3 \dots a_n \pmod{4},$$

so daß nach Satz 1

$$\left(\frac{N\alpha}{2^e}, k, r\right) = \left(\frac{a_1}{2^e}, k, r\right) \prod_{i=2}^n (a_i, k, r).$$

Da $r \equiv b_1 \pmod{8}$, so ist

$$\left(\frac{a_1}{2^e}, k, r\right) = \left(\frac{a_1}{2^e}, k, b_1\right).$$

Wenn $i > 1$ ist, so hat man $a_i \equiv -b_i r \pmod{4}$, woraus augenscheinlich

$$(a_i, k, r) = (a_i, k, b_i).$$

Aus den drei letzten Gleichungen folgt

$$\left(\frac{N\alpha}{2^e}, k, r\right) = \left(\frac{a_1}{2^e}, k, b_1\right) \prod_{i=2}^n (a_i, k, b_i).$$

Da aber $r \equiv b_1 \pmod{8}$, so ist $\left(\frac{2, k}{r}\right) = \left(\frac{2, k}{b_1}\right)$, so daß

$$(2) \quad (N\alpha, k, r) = \left(\frac{2, k}{r}\right)^e \left(\frac{N\alpha}{2^e}, k, r\right) = \\ \left(\frac{2, k}{b_1}\right)^e \left(\frac{a_1}{2^e}, k, b_1\right) \prod_{i=2}^n (a_i, k, b_i) = \prod_i (a_i, k, b_i).$$

Aus (1) und (2) folgt die Richtigkeit des Satzes in diesem Falle.

Um Satz 3 vollständig zu beweisen, braucht dann weiter nur folgendes gezeigt zu werden: Satz 3 gelte für $(\bar{\alpha}, \beta)$, während α eine beliebige Zahl von der im Satze erwähnten Beschaffenheit ist. Dann gilt Satz 3 für (α, β) .

Man hat, daß $\frac{\alpha}{\bar{\alpha}} = \frac{\delta}{\bar{\delta}}$ geschrieben werden kann, wo δ und $\bar{\delta}$ beide prim zu 2 und zu β sind. Ich setze $\bar{\alpha} \equiv \bar{a}_1 \pmod{l_1^{e+2}}$ und $\equiv \bar{a}_i \pmod{l_i^2}$ für $i > 1$, während $\alpha \equiv a_1$ bzw. a_i nach diesen Moduln ist. Ebenso schreibe ich δ bzw. $\bar{\delta}$ für jedes $i \equiv d_i$ bzw. $\bar{d}_i \pmod{l_i^2}$. Aus $\alpha \bar{\delta} = \bar{\alpha} \delta$ folgt

$$(\alpha \bar{\delta}, K, \beta) = (\bar{\alpha} \delta, K, \beta),$$

und nach Satz 2 und der gemachten Annahme hat man

$$(\bar{\alpha}\delta, K, \beta) = (\bar{\alpha}, K, \beta) (\delta, K, \beta) = \prod_i (\bar{a}_i, k, b_i) \prod_i (d_i, k, b_i) = \prod_i (\bar{a}_i d_i, k, b_i).$$

Hieraus wieder nach Satz 2

$$(3) \quad (\alpha, K, \beta) = (\bar{\alpha}\delta, K, \beta) (\bar{\delta}, K, \beta) = \prod_i (\bar{a}_i d_i, k, b_i) \prod_i (\bar{d}_i, k, b_i).$$

Nun ist

$$\alpha\bar{\delta} \equiv a_1 \bar{d}_1 \text{ und } \bar{\alpha}\delta \equiv \bar{a}_1 d_1 \pmod{1_1^{\epsilon+2}},$$

wie man leicht findet. Aus der Gleichung $\alpha\bar{\delta} = \bar{\alpha}\delta$ folgt deshalb

$$a_1 \bar{d}_1 \equiv \bar{a}_1 d_1 \pmod{2^{\epsilon+2}},$$

woraus nach Satz 1

$$(a_1 \bar{d}_1, k, b_1) = (\bar{a}_1 d_1, k, b_1).$$

Ebenso hat man für $i > 1$

$$a_i \bar{d}_i \equiv \bar{a}_i d_i \pmod{4},$$

woraus nach Satz 1

$$(a_i \bar{d}_i, k, b_i) = (\bar{a}_i d_i, k, b_i).$$

Aus den beiden letzten Gleichungen folgt

$$\prod_i (a_i \bar{d}_i, k, b_i) = \prod_i (\bar{a}_i d_i, k, b_i),$$

woraus

$$(4) \quad \prod_i (a_i, k, b_i) = \prod_i (\bar{a}_i d_i, k, b_i) \prod_i (\bar{d}_i, k, b_i).$$

Aus (3) und (4) folgt

$$(\alpha, K, \beta) = \prod_i (a_i, k, b_i).$$

Hierdurch ist Satz 3 vollständig bewiesen.

Satz 4. Die ganzen Zahlen α und β seien relativ prim, α für jedes i teilbar durch genau $l_i^{e_i}$, β dagegen prim zu 2. Ist $\alpha \equiv a_i \pmod{l_i^{e_i+2}}$ und $\beta \equiv b_i \pmod{l_i^{g_i}}$, wo $g_i = 2$ oder $= 3$ ist, je nachdem $e_i = 0$ oder > 0 ist, so ist

$$(\alpha, K, \beta) = \prod_i (a_i, k, b_i).$$

Beweis: Den Satz beweise ich zuerst für beliebiges β , aber nur gewisse α . Es sei nämlich für jede Zahl j der Reihe 1, 2, ... n α_j prim zu β und $\frac{2}{l_j}$, aber teilbar durch genau $l_j^{e_j}$. Dann ist $\alpha_1 \alpha_2 \dots \alpha_n$ eine spezielle Zahl⁶⁾ α der im Satze erwähnten Art, und der Satz soll zuerst für das Zahlenpaar $(\prod_j \alpha_j, \beta)$ bewiesen werden.

Ich setze

$$\alpha_j \equiv a_{j,j} \pmod{l_j^{e_j+2}} \text{ und } \equiv a_{j,i} \pmod{l_i^2}, \text{ wenn } i \neq j.$$

Nun ist

$$\left(\prod_j \alpha_j, K, \beta \right) = \prod_j (\alpha_j, K, \beta)$$

und nach Satz 3 für alle j

$$(\alpha_j, K, \beta) = \prod_i (a_{ji}, k, b_i);$$

denn für alle i ist ja $\beta \equiv b_i \pmod{l_i^2}$ und falls $e_j > 0$ ist, β sogar $\equiv b_j \pmod{l_j^3}$.

Also wird

$$(5) \quad \left(\prod_j \alpha_j, K, \beta \right) = \prod_j \prod_i (a_{ji}, k, b_i) = \prod_i \left(\prod_j a_{ji}, k, b_i \right).$$

Außerdem ist

$$\prod_{j \neq i} \alpha_j \equiv \prod_{j(\neq i)} a_{ji} \pmod{l_i^2},$$

⁶⁾ Ist die Zahl der Idealklassen in K gleich 1, so kann freilich jede Zahl α in dieser Form geschrieben werden.

woraus, da α_i teilbar ist durch $l_i^{e_i}$ und $\equiv a_{ii} \pmod{l_i^{e_i+2}}$ ist,

$$(6) \quad \prod_j \alpha_j \equiv \prod_j a_{ji} \pmod{l_i^{e_i+2}}.$$

Aus (5) und (6) folgt die Wahrheit des Satzes in diesem Falle.

Danach soll eine beliebige Zahl α der im Satz erwähnten Art betrachtet werden. Ich setze der Kürze halber $\prod_j \alpha_j = \bar{\alpha}$, $\prod_j a_{ji} = \bar{a}_i$, und man hat

$$\frac{\alpha}{\bar{\alpha}} = \frac{\delta}{\bar{\delta}},$$

wo δ und $\bar{\delta}$ prim zu 2 und β sind, und es soll für jedes i

$$\delta \equiv d_i, \bar{\delta} \equiv \bar{d}_i \pmod{l_i^2}$$

sein. Da $\alpha\bar{\delta} = \bar{\alpha}\delta$ ist, bekommt man unter Benutzung des schon bewiesenen

$$(\alpha\bar{\delta}, K, \beta) = (\bar{\alpha}\delta, K, \beta) = (\bar{\alpha}, K, \beta)(\delta, K, \beta) = \prod_i (\bar{a}_i, k, b_i) \prod_i (d_i, k, b_i) = \prod_i (\bar{a}_i d_i, k, b_i)$$

und

$$(7) \quad (\alpha, K, \beta) = (\bar{\alpha}\delta, K, \beta)(\bar{\delta}, K, \beta) = \prod_i (\bar{a}_i d_i, k, b_i) \prod_i (\bar{d}_i, k, b_i).$$

Nun ist

$$\alpha\bar{\delta} \equiv a_i \bar{d}_i \text{ und } \bar{\alpha}\delta \equiv \bar{a}_i d_i \pmod{l_i^{e_i+2}}$$

und also

$$a_i \bar{d}_i \equiv \bar{a}_i d_i \pmod{2^{e_i+2}},$$

woraus nach Satz 1

$$(a_i \bar{d}_i, k, b_i) = (\bar{a}_i d_i, k, b_i)$$

oder

$$(8) \quad (a_i, k, b_i) = (\bar{a}_i d_i, k, b_i)(\bar{d}_i, k, b_i).$$

Aus (7) und (8) folgt wiederum

$$(\alpha, K, \beta) = \prod_i (\alpha_i, k, b_i).$$

Hierdurch ist Satz 4 bewiesen.

Satz 5. Die ganzen Zahlen α und β seien relativ prim und für jedes i α genau teilbar durch $\mathfrak{l}_i^{e_i}$, β teilbar durch genau $\mathfrak{l}_i^{f_i}$. Weiter sei für jedes i $\alpha \equiv a_i \pmod{\mathfrak{l}_i^{g_i}}$, $\beta \equiv b_i \pmod{\mathfrak{l}_i^{h_i}}$, wobei $g_i = e_i + 2$ oder $= 3$ ist, je nachdem $f_i = 0$ oder > 0 ist, und ebenso $h_i = f_i + 2$ oder $= 3$, je nachdem $e_i = 0$ oder > 0 ist. Dann gilt die Gleichung

$$(\alpha, K, \beta) = \prod_i (\alpha_i, k, b_i).$$

Beweis: Indem β beliebig ist, beweise ich den Satz zuerst für gewisse spezielle α , nämlich

$$\alpha = \gamma^2 - \beta \xi.$$

So oft $e_i > 0$ sein soll, wird ξ so gewählt, daß

$$(9) \quad \beta \xi \equiv 1 \pmod{\mathfrak{l}_i^3}$$

wird, und danach γ so bestimmt, daß α teilbar durch genau $\mathfrak{l}_i^{e_i}$ wird. Ist $f_i > 0$, setze man $\gamma \equiv 1 \pmod{\mathfrak{l}_i}$; dann wird ja $e_i = 0$. Soll endlich $e_i = 0$ sein, indem auch $f_i = 0$ ist, so setze man $\gamma \equiv 0 \pmod{\mathfrak{l}_i}$. Außerdem kann man ξ total positiv und überhaupt prim zu 2 wählen. Dann hat man, weil der Signaturfaktor offenbar $= +1$ wird,

$$(10) \quad (\alpha, K, \beta) = \left(\frac{\alpha, K}{\beta} \right) \left(\frac{\beta, K}{\alpha} \right) = \left(\frac{\xi, K}{\alpha} \right) = \left(\frac{\xi, K}{\alpha} \right) \left(\frac{\alpha, K}{\xi} \right) = (\alpha, K, \xi).$$

Nach Satz 4 ist nun

$$(11) \quad (\alpha, K, \xi) = \prod_i (\alpha_i, k, x_i),$$

wenn $\xi \equiv x_i \pmod{\mathfrak{l}_i^3}$ ist, wenn $e_i > 0$, und $\pmod{\mathfrak{l}_i^2}$ sonst. Nun wird aber nach (9) $b_i x_i \equiv 1 \pmod{8}$, so oft $e_i > 0$ ist, woraus folgt $(\alpha_i, k, x_i) = (\alpha_i, k, b_i)$. Ist $e_i = f_i = 0$, so ist $a_i \equiv -b_i x_i \pmod{4}$, woraus $(\alpha_i, k, x_i) = (\alpha_i, k, b_i)$. Es bleibt der Fall $f_i > 0$. Dann ist $\beta \equiv b_i \pmod{\mathfrak{l}_i^{f_i+2}}$ und

also in jedem Falle $\text{mod. } l_i^8$. Weiter ist $\gamma \equiv 1 \pmod{l_i}$, woraus bekanntlich $\gamma^8 \equiv 1 \pmod{l_i^8}$. Da außerdem $\alpha \equiv a_i \pmod{l_i^8}$, bekommt man

$$a_i \equiv 1 - b_i x_i \pmod{8},$$

und b_i ist eine gerade Zahl, nämlich teilbar durch 2^{f_i} . Nun wird

$$(a_i, k, b_i) = \left(\frac{2, k}{a_i} \right)^{f_i} \left(a_i, k, \frac{b_i}{2^{f_i}} \right).$$

Ist $f_i \geq 3$, so ist $a_i \equiv 1 \pmod{8}$, woraus $(a_i, k, b_i) = 1 = (a_i, k, x_i)$. Ist $f_i = 2$, so ist $a_i \equiv 1 \pmod{4}$, woraus wiederum $(a_i, k, b_i) = 1 = (a_i, k, x_i)$. Ist endlich $f_i = 1$, so hat man

$$\text{entweder } \frac{b_i x_i}{2} \equiv 1 \quad \text{oder} \quad \equiv -1 \pmod{4}.$$

Im ersten Falle ist $a_i \equiv -1 \pmod{8}$ und deshalb

$$(a_i, k, b_i) = \left(a_i, k, \frac{b_i}{2} \right) = (a_i, k, x_i).$$

Im zweiten Falle ist $a_i \equiv 3 \pmod{8}$, woraus

$$(a_i, k, b_i) = - \left(a_i, k, \frac{b_i}{2} \right) = (a_i, k, x_i).$$

Man hat also für alle i , daß $(a_i, k, x_i) = (a_i, k, b_i)$ ist. Dies in Verbindung mit (10) und (11) zeigt die Richtigkeit von Satz 5 für die speziellen α .

Um die Allgemeingültigkeit des Satzes zu zeigen gehe ich in derselben Weise wie früher vor. Der Satz sei also gültig für $(\bar{\alpha}, \beta)$, und α sei eine beliebige Zahl der im Satze erwähnten Beschaffenheit. Ich setze

$$\frac{\alpha}{\bar{\alpha}} = \frac{\delta}{\bar{\delta}},$$

wo δ und $\bar{\delta}$ prim zu 2β sind. Weiter sollen δ und $\bar{\delta} \equiv d_i$ bzw. $\bar{d}_i \pmod{l_i^{t_i}}$ sein, wo $t_i = 2$ oder $= 3$ ist, je nachdem $f_i = 0$ oder > 0 ist. Wie früher hat man

$$(\alpha, K, \beta) = (\bar{\alpha}, K, \beta) (\delta, K, \beta) (\bar{\delta}, K, \beta),$$

und nach der gemachten Voraussetzung und nach Satz 4 erhält man

$$(12) \quad (\alpha, K, \beta) = \prod_i (\bar{a}_i d_i, k, b_i) \cdot \prod_i (\bar{d}_i, k, b_i).$$

Andererseits folgt aus der Gleichung $\alpha \bar{\delta} = \bar{\alpha} \delta$, daß wenn zuerst die i , für welche $f_i = 0$ ist, betrachtet werden

$$a_i \bar{d}_i \equiv \bar{a}_i d_i \pmod{2^{e_i+2}};$$

denn $\alpha \bar{\delta} \equiv a_i \bar{d}_i$ und $\bar{\alpha} \delta \equiv \bar{a}_i d_i \pmod{l_i^{e_i+2}}$, wie leicht einzusehen. Da $a_i \bar{d}_i$ und $\bar{a}_i d_i$ teilbar sind durch 2^{e_i} , aber nicht 2^{e_i+1} , so folgt nach Satz I, daß

$$(a_i \bar{d}_i, k, b_i) = (\bar{a}_i d_i, k, b_i),$$

woraus

$$(13) \quad (a_i, k, b_i) = (\bar{a}_i d_i, k, b_i) (\bar{d}_i, k, b_i).$$

Ist $f_i > 0$, so hat man $\alpha \bar{\delta} \equiv a_i \bar{d}_i$ und $\bar{\alpha} \delta \equiv \bar{a}_i d_i \pmod{l_i^3}$, woraus

$$a_i \bar{d}_i \equiv \bar{a}_i d_i \pmod{8};$$

außerdem sind $a_i \bar{d}_i$ und $\bar{a}_i d_i$ ungerade.

Aber dann ist offenbar wiederum

$$(a_i \bar{d}_i, k, b_i) = (\bar{a}_i d_i, k, b_i)$$

oder

$$(14) \quad (a_i, k, b_i) = (\bar{a}_i d_i, k, b_i) (\bar{d}_i, k, b_i).$$

Aus (12), (13) und (14) folgt die Richtigkeit des Satzes 5, der hierdurch vollständig bewiesen ist.

Beispiel: Im Körper $k(\sqrt{-7})$ ist $2 = \lambda_1 \lambda_2$, wobei

$$\lambda_1 = \frac{1 + \sqrt{-7}}{2}, \quad \lambda_2 = \frac{1 - \sqrt{-7}}{2},$$

sind. Es wird z. B.

$$\alpha = \frac{-5 + 3\sqrt{-7}}{2}, \quad \beta = 1 + 2\sqrt{-7}$$

gewählt. Hier ist $\beta \equiv 1 \pmod{2}$; dagegen ist α prim zu λ_2 , aber teilbar durch λ_1 , indem

$$\alpha = (2 + \sqrt{-7}) \lambda_1$$

ist. Man findet

$$\lambda_1^3 = \frac{-5 - \sqrt{-7}}{2}, \quad \lambda_2^3 = \frac{-3 - \sqrt{-7}}{2}.$$

Mod. λ_1^3 sind dann

$$\alpha \equiv \frac{-5 - 15}{2} = -10 \equiv 6 \text{ und } \beta \equiv 1 - 10 = -9 \equiv 7.$$

Mod. λ_2^3 sind

$$\alpha \equiv \frac{-5 - 9}{2} = -7 \equiv 1 \text{ und } \beta \equiv 1 - 6 = -5 \equiv 3.$$

Deshalb wird (dies folgt schon aus Satz 3 oben)

$$(\alpha, k(\sqrt{-7}), \beta) = (6, k, 7)(1, k, 3).$$

Nun ist $(1, k, 3) = 1$, während

$$(6, k, 7) = \left(\frac{2, k}{7}\right)(3, k, 7) = +1 \cdot -1 = -1.$$

Also, da der Signaturfaktor in einem total imaginären Körper immer $= +1$ ist,

$$\left(\frac{\alpha, k(\sqrt{-7})}{\beta}\right) \left(\frac{\beta, k(\sqrt{-7})}{\alpha}\right) = -1.$$

Nun findet man leicht, daß $\alpha \equiv 4$ ist mod. β ; also ist $\left(\frac{\alpha, k(\sqrt{-7})}{\beta}\right) = +1$.

Weiter ist nach der gemachten Vereinbarung

$$\left(\frac{\beta, k(\sqrt{-7})}{\alpha}\right) = \left(\frac{\beta, k(\sqrt{-7})}{2 + \sqrt{-7}}\right).$$

Mod. $2 + \sqrt{-7}$ ist aber $\beta \equiv -3$ und $\left(\frac{-3, k(\sqrt{-7})}{2 + \sqrt{-7}}\right) = \left(\frac{-3, k}{11}\right) = \left(\frac{11, k}{3}\right) = \left(\frac{-1, k}{3}\right) = -1$. Dadurch ist das obige Ergebnis bestätigt.

(Eingegangen den 15. November 1932.)

Ein Satz über die Ring- und Strahlklassenzahlen in algebraischen Zahlkörpern

Von RUDOLF FUETER, Zürich

Nach *Hurwitz*¹⁾ ist die Klassenzahl eines algebraischen Zahlkörpers gleich der Anzahl nicht äquivalenter Zahlbrüche des Körpers, wobei die Gruppe aller ganzzahligen unimodularer Substitutionen des Körpers zugrunde gelegt wird. Für reguläre Ringe und Strahlen ist meines Wissens kein solcher Satz bekannt. Im folgenden gebe ich für diesen Fall das sehr einfach zu beweisende Resultat. Es scheint mir bemerkenswert, da es auf die Untergruppen der genannten Gruppe ein neues Licht wirft.

1. Es sei s ein Strahl in einem gegebenen algebraischen Zahlkörper k . Derselbe sei aus allen $(\text{mod } f)$ kongruenten Zahlen einer Untergruppe bezüglich der Multiplikation von Restklassen $(\text{mod } f)$ gebildet²⁾. f ist sein Führer. Die gebrochenen Zahlen, deren Zähler und Nenner zu f teilerfremd sein müssen, und die in den betreffenden Kongruenzklassen liegen, werden ebenfalls zu s hinzugenommen. *In diese Definition fallen auch die regulären Ringe mit dem Führer f^3 .*

Es sei $S = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ irgend eine der Substitutionen in k , die folgenden Bedingungen genügen:

a) $\alpha, \beta, \gamma, \delta$ sind ganze Zahlen von k , die die Gleichung:

$$\alpha\delta - \beta\gamma = 1$$

befriedigen.

b) α, δ sind im Strahl s , und γ ist eine Zahl von f .

Alle S bilden eine Gruppe in Bezug auf Multiplikation, die mit $\mathfrak{G}(s, f)$ bezeichnet werde. Denn nach dem Multiplikationsgesetz ist:

$$S' S'' = S, \quad S^{(r)} = \begin{pmatrix} \alpha^{(r)} & \beta^{(r)} \\ \gamma^{(r)} & \delta^{(r)} \end{pmatrix}, \quad r = 1, 2; \quad S = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix},$$

¹⁾ *A. Hurwitz*: Die unimodularen Substitutionen in einem algebraischen Zahlkörper. Nachr. Ges. Wiss. Göttingen, Math. phys. Klasse 1895, S. 332.

²⁾ Siehe etwa die Ausführungen von *H. Hasse*: Bericht über neue Untersuchungen und Probleme aus der Theorie der algebraischen Zahlkörper, Jahresber. der D. M. V. Bd. 35 (1926), S. 5 u. ff., wo s als $H(f)$ bezeichnet wird.

³⁾ Siehe *Rud. Fueter*: Die Klassenkörper der komplexen Multiplikation und ihr Einfluß auf die Entwicklung der Zahlentheorie, Jahresber. D. M. V. Bd. 20 (1911), S. 9—10.

wo S wieder ganze Koeffizienten hat. Anderseits ist:

$$\gamma = \gamma' \alpha'' + \delta' \gamma'',$$

also wieder durch f teilbar, da es γ', γ'' sind. Außerdem ist:

$$\begin{aligned} \alpha &= \alpha' \alpha'' + \beta' \gamma'' \equiv \alpha' \alpha'' \pmod{f}, \\ \delta &= \gamma' \beta'' + \delta' \delta'' \equiv \delta' \delta'' \pmod{f}, \end{aligned}$$

sind also beide wieder in s . Ferner ist $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ und die Inverse $\begin{pmatrix} \delta & -\beta \\ -\gamma & \alpha \end{pmatrix}$ wieder in $\mathfrak{G}(s, f)$.

2. Jedes zu f teilerfremde Ideal $i = (\sigma, \tau)$ kann als größter gemeinsamer Teiler einer Zahl σ von s und einer zu f teilerfremden Zahl τ aufgefaßt werden. Ist $i' = (\sigma', \tau')$ ein zweites Ideal, wo σ', τ' dieselben Bedingungen erfüllen, und sollen i, i' in der gleichen Strahlklasse liegen, so muß es eine Zahl μ von s geben, für die:

$$(\mu) i' = i$$

ist. $\mu \tau'$ ist dann sicherlich eine ganze Zahl in k . Wir nehmen jetzt eine Zahl φ von f mit der Eigenschaft, daß sie zu σ und σ' teilerfremd ist. Dann ist auch:

$$i = (\sigma, \varphi \tau), \quad i' = (\sigma', \varphi \tau'),$$

und es muß eine Substitution $S = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ mit ganzen Zahlen $\alpha, \beta, \gamma, \delta$ in k geben⁴⁾, für die:

$$\begin{aligned} \mu \varphi \tau' &= \alpha \varphi \tau + \beta \sigma, \\ \alpha \delta - \beta \gamma &= 1, \\ \mu \sigma' &= \gamma \varphi \tau + \delta \sigma, \end{aligned} \tag{1}$$

ist. Aus der ersten Gleichung folgt, daß β durch φ teilbar ist, da σ zu φ teilerfremd ist und $\mu \tau'$ ganz ist: $\beta = \beta' \varphi$. Somit erhält man, wenn man $\gamma' = \varphi \gamma$ setzt, die neuen Gleichungen:

$$\begin{aligned} \mu \tau' &= \alpha \tau + \beta' \sigma, \\ \alpha \delta - \beta' \gamma' &= 1, \\ \mu \sigma' &= \gamma' \tau + \delta \sigma, \end{aligned} \tag{2}$$

⁴⁾ D. Hilbert: Zahlbericht, Satz 53, oder Hurwitz, a. a. O.

$S' = \left(\frac{\alpha \beta'}{\gamma' \delta} \right)$ liegt aber in $\mathfrak{G}(s, f)$. Denn γ' ist durch f teilbar, und $\mu \sigma' \equiv \delta \sigma \pmod{f}$; woraus folgt, daß δ in s liegt, da μ, σ', σ in s liegen nach Annahme. Wegen $\alpha \delta \equiv 1 \pmod{f}$ muß dann auch α in s liegen. Wenn also $i = (\sigma, \tau)$ und $i' = (\sigma', \tau')$ in derselben Strahlklasse liegen, und σ, σ' in s sind, so gibt es eine Substitution von $\mathfrak{G}(s, f)$, so daß die Zahlbrüche $\tau/\sigma, \tau'/\sigma'$ äquivalent sind.

3. Sind umgekehrt zwei Zahlbrüche, deren Zähler und Nenner zu f teilerfremd sind, nach $\mathfrak{G}(s, f)$ äquivalent, so kann man die beiden Zahlbrüche stets in der Form $\tau/\sigma, \tau'/\sigma'$ darstellen, wo σ, σ' in s liegen. Man ordnet den Zahlbrüchen die Ideale $(\sigma, \tau) = i, (\sigma', \tau') = i'$ zu. Verschiedenen Darstellungen eines Zahlbruches in dieser Form entsprechen zugeordnete Ideale (σ, τ) derselben Strahlklasse. Aus (2) folgt dann, da γ' durch f teilbar ist, daß $\mu \sigma' \equiv \delta \sigma \pmod{f}$, d. h. μ muß in s liegen, da es σ', δ, σ nach Annahme tun. Somit sind die den Zahlbrüchen zugeordneten Ideale i, i' in derselben Strahlklasse.

Aus 2. und 3. folgt der

I. Satz: Die Strahlklassenzahl von s ist gleich der Anzahl der in Bezug auf $\mathfrak{G}(s, f)$ inäquivalenten Zahlbrüche von k , deren Zähler und Nenner zu f teilerfremd sind.

Man kann dieses Resultat mit Hilfe eines entsprechenden Beweises auch anders aussprechen, was für die Formentheorie wichtig ist: Es sei $\overline{\mathfrak{G}}(s, f)$ die gleiche Gruppe wie $\mathfrak{G}(s, f)$, nur daß statt der γ die β durch den Führer teilbar sind. Dann gilt der

II. Satz: Die Strahlklassenzahl von s ist gleich der Anzahl der in bezug auf $\overline{\mathfrak{G}}(s, f)$ inäquivalenten Zahlbrüche von k , deren Zähler durch den Führer f teilbar, deren Nenner zu f teilerfremd sind.

4. Nimmt man als Anwendung die gewöhnlichen primitiven binären quadratischen Formen im Bereiche der ganzen rationalen Zahlen mit der Diskriminante $f^2 m$, wo m negativ und quadratfrei sei:

$$F \equiv Ax^2 + 2Bxy + Cy^2, B^2 - AC = f^2 m,$$

so gibt es, wie man sofort sieht, in jeder Klasse eine Form, in der B durch f , und C durch f^2 teilbar ist. Man kann sich also auf die Frage der Äquivalenz der Formen:

$$F \equiv Ax^2 + 2Bxfy + C(fy)^2, B^2 - AC = m,$$

beschränken. Dieselben gehen aber durch:

$$\begin{aligned} y &= \alpha y' + \beta' x', \\ \alpha \delta - \beta' \gamma' &= 1, \\ x &= \gamma' y' + \delta x' \end{aligned}$$

dann und nur dann ineinander über, wenn $\gamma' \equiv 0 \pmod{f}$ ist, also durch die Substitutionen von $\mathfrak{G}(s, f)$. Setzt man $f\beta' = \beta$, $\gamma' = f\gamma$, so geht dann:

$$F' \equiv Ax^2 + 2Bxy + Cy^2, B^2 - AC = m,$$

durch $S = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$, $\beta \equiv 0 \pmod{f}$ in eine entsprechende Form der Diskriminante m über. Die Ringklassenzahl ist also gleich der Zahl inäquivalenter Formen der Diskriminante m in bezug auf die Untergruppe der Modulgruppe, für die $\beta \equiv 0 \pmod{f}$.

(Eingegangen den 1. März 1933)